Fairness properties of compensation schemes*

Dietmar Fehr[1,2], Hannes Rau[1,3], Stefan T. Trautmann[1,4], Yilong Xu[5] and Christoph Becker[1]

[1]*University of Heidelberg, Germany*
[2]*CESifo, Germany*
[3]*Karlsruhe Institute for Technology, Germany*
[4]*Tilburg University, Netherlands*
[5]*Utrecht University, Netherlands*

January 23, 2023

Abstract
How do different characteristics of pay-for-performance schemes affect fairness perceptions? In two studies, we systematically consider three major classes of incentives schemes: continuous piece rate incentives, discrete bonus schemes, and tournament incentives. We find that pay inequality has a strong negative effect on perceived fairness. Controlling for pay inequality, people consider piece rate schemes fairer than those with a discrete bonus and especially a tournament design. Adding resource advantages or handicaps to vary the level of the playing field negatively influences perceived fairness. Distinguishing between procedural and outcome fairness, we find that procedural fairness judgments are more positive, and influence overall judgments more strongly than outcome fairness judgments.

KEYWORDS: incentives, merit, contract design, fairness, inequality.
JEL CLASSIFICATIONS: C90, D63, J31, J41

---

* We are very grateful for helpful comments by participants of the HEIRS 2021 Illusion of Merit Workshop.
Trautmann (corresponding author): Alfred-Weber-Institute for Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany; Phone: +49 6221 54 2952, Fax: +49 6221 54 3592; email: trautmann@uni-hd.de. Fehr: Alfred-Weber-Institute for Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany; email: dietmar.fehr@awi.uni-heidelberg.de. Rau: Department of Economics, KIT, Fritz-Erler-Str. 1-3, 76133 Karlsruhe, Germany; email: hannes.rau@kit.edu. Xu: Utrecht School of Economics, Utrecht University, Kriekenpitplein 21-22, Utrecht, The Netherlands; email: y.xu3@uu.nl.

1. Introduction

Pay-for-performance contracts are a key feature of any modern governance structure, and there is a long tradition in management research to investigate the optimal design of incentive contracts. This research has mostly focused on how bonus and tournament incentive schemes can mitigate the agency problem (Milgrom and Roberts, 1992; Prendergast, 1999; Bell et al., 2021) but less on the more intangible issues, such as fairness perceptions of incentive schemes. However, these issues gain importance given the steady rise in wage inequality, mainly as a result of the increasing use of pay-for-performance contracts (Lemieux et al., 2009; Pan et al., 2022).

Theoretical contributions typically account for agents' utility of monetary consequences and some formal aspects of social preferences (e.g., inequality aversion; Kőszegi, 2014). More general factors of fairness play a minor role when optimizing incentive schemes to induce a certain level of employee effort. However, such generalized fairness perceptions in employment relations are relevant for softer governance aspects, such as organizational culture and worker morale (Greenberg, 1987; Ambrose et al., 2002; Bénabou and Tirole, 2016). Assuming a framework in which the features of a performance contract affect fairness perceptions, and these fairness perceptions subsequently affect employee behavior, pay-for-performance contracts may thus have unintended consequences, leading to lower productivity, more cheating, and sabotage, or less cooperation and social cohesion in the organization (Harbring et al., 2007; Card et al., 2012; Murayama and Elliot, 2012; Cohn et al. 2015; Fehr, 2018; Breza, Kaur, and Shamdasani, 2018; Fehr et al. 2020; Xu and Ginevra, 2023). Therefore, from a governance perspective, understanding how different aspects of contracts affect fairness perceptions is critical to minimize the possible adverse side effects of incentive contracts.

We set out to systematically examine the impact of contract incentive features on fairness perceptions, which has received little attention in the literature so far.[1] In doing so, we focus on a broad array of contractual features that firms use to mitigate agency problems that arise through asymmetric information about employee types and behavior. Depending on the observability of agents' types and behavior and the contractibility of outcomes, firms use a variety of contract features such as piece-rate incentives, discrete bonus schemes, and tournament competition (e.g. Lazear, 2018). The overarching research question we address is how individuals perceive these three classes of payment schemes and, in particular, how they weigh specific features of the contract and the resulting pay inequality in their fairness perceptions. That is, we aim to test how different contract features affect fairness perceptions directly, for example, when they are considered inappropriate or undesirable (e.g., fierce worker competition) and indirectly through their influence on the level and distribution of monetary payoffs.

Observing fairness perceptions of incentive contracts in the field is complicated by three obstacles. First, randomized experimental variation of several contract features within a fixed organizational context is typically impossible (our study implements systematic variation in 16 different conditions). For similar reasons, using observational field data does not allow us to compare different schemes without confounding unobserved factors of the situation or the workplace. The presence of selection effects further complicates the interpretation of field data: Because people may select into occupations with specific incentive contracts, fairness judgment may not be unbiased

[1] While there is a large strand of literature that provides compelling evidence that perceptions of (un)fairness influence employee behavior (e.g., Greenberg, 1987; Ambrose et al., 2002; Cohn et al. 2015; Kao et al., 2018; Fehr et al. 2020), very few papers investigate the effect of different aspects of the incentive contract on fairness perceptions (Jawahar and Stone, 2011; Gupta and Shaw, 2014; see also Cappelen et al., 2020). An exception is an experimental study by Kleinlercher and Stöckel (2018), in which they vary the 'salience' of subjects' final payoffs (dependence of payoffs on own or other players' actions and the institutional framework in an understandable fashion) and measure the fairness perceptions of the corresponding incentive scheme.

(Dillard and Fisher, 1990; Bartling et al., 2018; Fulmer and Shaw, 2018).

To overcome these challenges, we conduct a controlled vignette study in which respondents rate the fairness of specific payment schemes. A vignette offers a contextually detailed description of a hypothetical situation and systematically varies the features of interest.[2] This makes it particularly well suited for research on perceptions of fairness and justice, as researchers can easily incorporate information that is potentially relevant for fairness perceptions, such as details about the sources of pay inequality (Konow, 2003).[3] In particular, vignettes allow us to address the two problems described above. First, we can avoid selection problems by randomly assigning respondents to payment schemes. We also use an abstract real-effort task, which makes it less likely that respondents' perceptions are driven by uncontrolled aspects of payment schemes in real-world labor relationships. Second, it also allows us to use a controlled manipulation of the different properties of the payment schemes within one conceptual framework. In other words, we can present different payment schemes holding the background situation fixed and, thus, avoid problems of omitted variable bias.

In two studies, we recruit diverse samples of respondents from two different platforms (MTurk and Prolific) and provide them with the description of a repeated, time-constrained real-effort task adapted from Fehr, Rau, Trautmann, and Xu (2020). The vignette also includes information on the empirically realized skill distribution in that task.

[2] While vignette studies have a long tradition in other social sciences, they only recently gained traction in economics and management, for example in research on labor markets (e.g., Finseraas et al., 2016; Kübler, Schmid, and Stueber, 2018), ethical judgments (Ambuehl et al., 2015; Ambuehl and Ockenfels, 2017), and financial literacy (Samek, Kapteyn, and Gray, 2020). Earlier pioneering work include Kahneman, Knetsch, and Thaler (1986) and Dahl and Ransom (1999).

[3] A common concern is the hypothetical nature of vignettes. However, evidence suggests that vignette responses reflect actual behavioral choices (e.g., Lanza et al. 1997; Hangartner et al. 2015). While the scenarios in our setting are hypothetical, we elicit non-incentivized *and* incentivized fairness judgments.

Across raters we then vary the description of the payment scheme, considering the three classes of payment schemes discussed above: continuous piece rate incentives, discrete bonus schemes that combine a piece rate with a bonus when surpassing a performance target level, and tournaments with two workers pitted against each other. Within each class, we vary the degree of pay inequality (steepness of the incentives) and consider time advantages and handicaps as a function of earlier performance in the discrete bonus and tournament payment schemes. Time advantages aim to capture the fact that successful employees often receive more resources or opportunities to improve future performance. Time handicaps, on the other hand, capture ratchet effects, where better performers are subsequently confronted with higher demands. They also aim to level the playing field, which is particularly relevant in tournament settings.

Given this setup, in Study 1, we elicited respondents' overall fairness judgments in a non-incentivized way. Study 2 replicates Study 1, and, in addition, we elicit respondents' incentivized fairness judgments using the coordination method proposed by Houser and Xiao (2011). Using both incentivized and non-incentivized fairness judgments gives us a comprehensive perspective on fairness perceptions of incentive contracts. The non-incentivized fairness judgments directly measure how individuals independently perceive different contractual features. In contrast, the incentivized fairness judgments represent the management perspective as they measure a shared understanding of fairness perceptions and thus reveal how the target population perceives certain contractual features. Moreover, to get a grip on how the fairness judgment depends on specific fairness concepts, in Study 2, we elicit judgments about the fairness of the allocation process and the allocation outcomes of the payment schemes. This distinction matters, given that a payment scheme can be considered procedurally fair even if it results in unequal outcomes (Trautmann, forthcoming). Again, we elicit these additional measures in both an incentivized and a non-incentivized way. We then test if the incentive mechanism affects the two concepts differently and how these concepts contribute to the overall fairness assessment.

We find that the different features of performance-based compensation schemes profoundly impact fairness perceptions across different samples, timing, and incentives. In particular, we show that, in general, people view payment schemes that lead to higher inequality as more unfair. To the extent that pay-for-performance contracts cannot be avoided, this result highlights that organizations and firms can accommodate fairness concerns through different contractual arrangements.

Zooming into the three classes of payment schemes, we delineate how different features of these schemes affect fairness perceptions. Controlling for the degree of inequality, people consider payment schemes including discontinuous payoffs, i.e., discrete bonuses and competitive tournaments, less fair than piece rates. Furthermore, time advantages or handicaps render a compensation scheme less fair. Advantages are perceived as especially unfair under a tournament scheme where competition between workers is relevant. Handicaps are generally perceived as unfair. This is interesting because the latter feature aims to level the playing field in subsequent interactions, leading to more equal outcomes, which seems desirable from a normative point of view (Roemer, 1998; Konow, 2003). However, a handicap appears at odds with fundamental conceptions of merit, and we find that it significantly reduces procedural fairness judgments. That is, merit is clearly part of a fair procedure.

Decomposing the fairness perceptions into procedural and outcome fairness, we find that procedural fairness judgments are consistently more positive than outcome fairness judgments in 13 of the 16 payment schemes and have a stronger influence on the overall fairness judgments than outcome fairness does. This suggests that firms can benefit from emphasizing procedural fairness considerations when designing payment schemes.

Finally, we test whether the results are sensitive to variation in socioeconomic characteristics, such as income, political orientation, and gender. We observe some specific patterns across high and low-income, liberal and conservative, and male and female respondents. However, these differences are not robust across our two studies and subject pools. This suggests that the pattern of fairness effects or the different payment scheme properties are more generally valid.

2. General Vignette Design

Study 1 and Study 2 are based on the same vignette design and scenarios, which we describe in this section. They differ with respect to the collected fairness judgments, the incentivization and the subject pools. These aspects are described separately for each study in Sections 3 and 4.

2.1. Task Scenario

In all conditions, respondents first received the same general description of an incentivized real-effort task, including a screenshot of the task. This task consisted of a screen showing 48 slider bars. The initial position of each slider was at the far left of a bar and had to be moved to the middle position of a bar within a given time frame (setup adapted from Gill and Prowse, 2012). Moreover, respondents learned that two workers were completing the real-effort task, that the task was repeated four times, and that the workers were paid depending on their success in the task, according to a payment scheme described to them later. It was made clear that, independently of the payment scheme, both workers in a group knew the other worker's performance and earnings after each repetition of the task (relative performance feedback; Eriksson et al., 2009). Respondents received a description of the specific payment scheme on the next screen. We implemented sixteen payment schemes, explained in detail below. Each respondent saw exactly one of these payment scheme scenarios to avoid concerns that ratings are blended with relative comparisons.

After reading the task description and the description of the payment scheme, respondents answered four questions probing their understanding of the payment scheme. These questions related to core aspects of the payment scheme, namely the expected earnings for a high-skilled versus a low-skilled worker and the time available for the slider task (which varies across conditions and workers). The questions aimed to ensure respondents understood the task, the payment scheme, and its payoff implications for high and low-skilled workers. Below we discuss how we treat respondents who failed the comprehension check for each study separately. After answering the comprehension questions, respondents judged the payment scheme's fairness properties. Finally, we ask respondents to fill out a demographic questionnaire (see section Sociodemographic Questionnaire in the Online Supplement).

2.2. Payment schemes

We base the task and payment schemes on a laboratory experiment with real payments (Fehr et al., 2020). We consider three general payment schemes (piece rate, discrete bonus, and tournament) and vary them along two margins. First, we vary the extent of the resulting pay inequality from low to high. Second, we manipulate specific characteristics of the schemes by introducing time advantages or handicaps for successful workers.

We consider 16 payment schemes (see the Online Supplement for the exact wording of all conditions). Table 1 displays the implications of each payment scheme on earnings for the average, the top 10%, and the bottom 10% of the workers. We calibrated these numbers based on workers' performance distribution in Fehr et al. (2020), in which the top 10% correctly placed 29 sliders per round, the bottom 10% placed 11 sliders, and the average worker 21 sliders. We designed the incentive contracts such that an average-skilled worker would earn approximately the same payment in all conditions. We calculate the Gini coefficient based on the top and bottom percentile payoffs. Note that we do not consider the potential effects of time advantages or handicaps (described below) on later-round earnings in the calculation of the Gini because these effects depend on the specific performance of both workers. However, ceteris paribus, time advantages induce larger inequality by giving more time to already successful workers, while handicaps reduce inequality by giving less time to successful workers.

2.2.1. Piece Rates

The first four conditions are individual (no interaction) piece-rate schemes. The piece rates differed concerning the steepness of the incentives. The first scheme had a safe base payment and a low piece rate, the second had no base payment and a larger piece rate, and the third had an entrance fee and an even larger piece rate. The fourth scheme had a non-linear, exponential mapping of the number of sliders correctly placed to payments, punishing low and average performance and rewarding high performance. The amount of time to complete each round of sliders was identical and fixed for all piece rate conditions at 120 seconds.

2.2.2. Discrete Bonus Incentives

The following six conditions involve discrete bonus incentives that pay a small piece rate incentive and a fixed bonus if a worker reaches a performance target. We included these conditions because bonuses based on performance targets are essential in real-world payment schemes (Oyer, 2000). For example, in many companies, employees receive yearly bonuses for reaching specific performance targets. Moreover, the discrete bonus conditions form an important intermediate step to the tournament incentives discussed below. In the tournament, the other player's performance provides a threshold at which a worker's earnings discontinuously change, depending on whether she performed better or worse than the other worker. In the discrete bonus schemes, this happens at the exogenously set performance target level. The six discrete bonus conditions differed concerning the bonus size at the performance target (low vs. high bonus). We calibrated the payment schemes so that for an average worker, who attains the target in two of the four rounds of work, the discrete bonus schemes paid approximately the same earnings as the piece rate schemes. However, the payoff difference between high and low performers was amplified in the high bonus condition.

The time to complete each round of sliders was fixed at 120 seconds in the first two discrete bonus conditions. In the other four conditions involving either high or low bonus levels, we added either a time advantage or a time handicap, in later rounds, for successful performance in earlier rounds. In particular, with a time advantage (handicap), workers who surpassed the target level gained (lost) 6 seconds of completion time in the next round, and vice versa for those who did not reach the threshold. The idea behind these conditions is that successful workers may either receive more resources or confront higher requirements (related to the ratchet effect and higher demands on good performers), for future tasks.

2.2.3. Tournament Incentives

The last six conditions in Table 1 involve competitive tournament incentives for the two workers in a group. In each round, the better-performing worker receives a larger prize than the worse-performing worker. In half of the six conditions, there is a low difference, and in the other half, there is a high difference between the winner and loser prizes. Like piece rates and discrete bonus schemes, tournament incentives lead to income differences between skilled and unskilled workers. However, they also induce a situation where both workers are directly pitted against each other, and income differences are potentially very salient.

Similar to the discrete bonus conditions, two tournament conditions involved a fixed time limit of 120 seconds. Four conditions involved the low or high tournament prize differences but added either a time advantage or handicap in later rounds for the winner of earlier rounds. Specifically, with a time advantage (handicap), round winners gained (lost) 6 seconds of completion time in the next round and vice versa for round losers.

Again, the idea behind these conditions is that successful workers may receive more or fewer resources than less successful workers for future tasks.

Table 1. Summary of fairness consideration and earnings by payment scheme

	Payment Scheme	Gini	Avg. earnings	Top 10% earnings	Bottom 10% earnings	Fairness[a] Study 1 / Study 2
	Piece-rate:					
T1:	Low_with base	0.17	6.40	7.48	5.32	8.51 / 8.46
T2:	Medium_no base	0.45	6.40	9.28	3.52	8.47 / 7.96
T3:	High_with	0.90	6.00	11.40	0.60	7.02 / 6.73
T4:	Exponential	0.91	6.08	25.84	1.28	7.50 / 7.01
	Discrete bonus:					
T5:	Low	0.17	6.84	7.92	5.68	7.83 / 7.93
T6:	High	0.87	6.64	12.32	0.88	7.17 / 6.97
T7:	Low_time	0.17	6.84	7.92	5.68	7.19 / 6.86
T8:	High_time	0.87	6.64	12.32	0.88	7.17 / 6.40
T9:	Low_time	0.17	6.84	7.92	5.68	6.81 / 6.43
T10:	High_time	0.87	6.64	12.32	0.88	6.71 / 6.58
	Tournament:					
T11:	Low	0.18	6.80	8.00	5.60	8.41 / 7.85
T12:	High	0.82	6.60	12.00	1.20	6.59 / 6.67
T13:	Low_time	0.18	6.80	8.00	5.60	6.55 / 6.29
T14:	High_time	0.82	6.60	12.00	1.20	5.35 / 5.40
T15:	Low_time	0.18	6.80	8.00	5.60	7.17 / 6.82
T16:	High_time	0.82	6.60	12.00	1.20	6.55 / 6.37

Notes: Low and high refers to the degree of inequality. Earnings (in USD) correspond to the amount an average (or top 10%, or bottom 10%) worker would receive as payment under the respective compensation scheme. a: Non-incentivized fairness evaluation, measured on a scale from 0 (completely unfair) to 10 (completely fair).

3. Study 1

3.1. Sample and Fairness Assessments

3.1.1 Respondents

We conducted Study 1 using the online labor market Amazon Mechanical Turk (MTurk).[4] We posted our study on MTurk, including a short description of the task, the task requirement, and the expected payment for completing the task. After accepting our task, respondents were redirected to our survey that was programmed in oTree (Chen et al., 2016). To address concerns of response quality, we added a simple CAPTCHA (adding two numbers) for an initial bot screen and, in addition, required that respondents *pass all* four comprehension questions about the payment scheme to proceed to the fairness judgments.

We recruited 2,431 U.S. residents in Fall of 2019 to participate in our study with a payment of $0.50 for completing the task.[5] Respondents were randomly allocated to one of the sixteen experimental conditions (on average, 152 respondents per condition). In the Online Supplement (Table OS3), we show that the sample is balanced with respect to observable characteristics such as gender, education, income, and ethnicity, among others, across our 16 payment scheme conditions.

[4] MTurk provides a more diverse subject pool than typically student populations and several studies indicate that findings from laboratory studies replicate on MTurk (e.g. Horton, Rand, and Zeckhauser 2011; Arechar, Gächter, and Molleman 2018; Coppock and McClellan 2019; Snowberg and Yariv 2021).

[5] In total, 4,805 MTurkers clicked on our task. Of these respondents, 216 did not pass the simple CAPTCHA, another 815 did not finish reading the instructions (without seeing the vignette), 1,335 did not pass the control questions and were excluded from the study, and 8 did not answer the fairness question, leaving us with *n=2,431* observations (see the Online Supplement for more details about dropout rates). In the Online Supplement we show that there is no differential attrition across conditions.

3.1.2 Fairness Assessment

In Study 1, respondents judged the overall fairness of the payment scheme using a simple non-incentivized measure, indicating their judgment on an 11-point scale from completely unfair (0) to completely fair (10):

> Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).
>
> Completely unfair O O O O O O O O O O O Completely fair

3.2 Results

We first observe that all payment schemes score above 5 (on a scale from 0-10) and vary substantially across conditions, ranging from 5.35 in the high-incentive tournament with time advantage (T14) to 8.51 in the condition with a low piece rate (T1), see Table 1.[6] This relative comparison is consistent with findings for players' own fairness judgments in incentivized laboratory games in Fehr et al. (2020), which implemented similar versions of these two payment schemes (T1 and T14). Fehr et al. (2020) also show that subjects in the perceived unfair condition T14 are less trusting and trustworthy than those in the perceived fair condition T1. We interpret this as evidence that assessments of partial (in Fehr et al., 2020) and impartial (in the current design) judges are similar with respect to the relative fairness comparisons of payment schemes, which are the focus of our study. The results by Fehr et al. (2020) suggest that fairness perceptions are relevant predictors of downstream behavior. They do not disentangle the differential impact of different payment scheme properties on fairness judgments and downstream behavior, though.

---

[6] In the Online Supplement, we show histograms of the full distribution of fairness evaluations for all 16 treatments.

Table 2. Multivariate analyses of fairness judgments – Study 1

	Fairness judgment			
	(1)	(2)	(3)	(4)
Gini	-1.33***	-1.30***	-1.32***	-1.28***
	(0.17)	(0.17)	(0.16)	(0.17)
Discrete bonus incentives	-0.32**	-0.24	-0.52***	-0.45***
	(0.15)	(0.15)	(0.17)	(0.17)
Tournament	-0.74***	-0.70***	-0.55***	-0.51***
	(0.14)	(0.15)	(0.16)	(0.17)
Time advantage	-0.94***	-0.99***	-0.28	-0.28
	(0.16)	(0.17)	(0.22)	(0.23)
Time handicap	-0.69***	-0.68***	-0.73***	-0.72***
	(0.15)	(0.16)	(0.22)	(0.23)
Tournament x			-1.29***	-1.39***
Time advantage			(0.31)	(0.33)
Tournament x			0.07	0.10
Time handicap			(0.30)	(0.31)
Exponential scheme	-0.02	0.01	-0.02	-0.01
	(0.19)	(0.21)	(0.19)	(0.20)
Constant	8.72***	8.61***	8.72***	8.59***
	(0.10)	(0.32)	(0.10)	(0.32)
Controls	No	Yes	No	Yes
Observations	2,431	2,182	2,431	2,182
R-squared	0.08	0.12	0.09	0.13
F-statistics (Discrete Bonus=Tournament)	10.16***	11.69***	0.02	0.08
F-statistics (Time bonus = Time handicap)	2.27	3.27*	3.60*	3.26*

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regression on fairness judgments with robust standard errors in parentheses. Piece rate is the reference category. Controls include gender, level of education, personal income (categorical), indicators for ethnic groups and for living in rural area, political orientation (1: Liberal; 6: Conservative), self-perceived social status, and employment status.

Next, we study the contribution of the different properties of the payment schemes to fairness judgments. In particular, we analyze the effect of inequality in outcomes as measured by the Gini coefficient, the effect of discrete bonus and tournament incentives compared to piece rate contracts, as well as the effect of additional contract features such as handicaps, time advantages, and exponential schemes. Table 2 shows the results of regressing fairness judgments on inequality, treatment indicators, and indicators for time handicaps, time advantages, and highly skewed piece rates, using all 16 payment schemes. In some specifications, we also include sociodemographic controls and interaction terms for the time advantage and handicap to test if advantages and handicaps are perceived differently in individual discrete bonus schemes compared to tournament schemes. We find that all design features of contracts that affect fairness lead to lower fairness judgments compared to the low-incentive piece rate contract, which is rated as most fair (see Table 1). We observe a negative impact of the Gini coefficient on fairness views. The same is true for discrete bonuses and tournament incentives, with tournament incentives being judged less favorably than discrete bonus schemes. Both time advantages and time handicaps reduce fairness judgments.

The specifications in the last two columns suggest that time handicaps for successful workers are never considered as fair. Therefore it seems that a handicap is at odds with fundamental conceptions of merit, even though it aims at leveling the playing field.[7] In contrast, time advantages are judged neutrally in individual-level discrete bonuses but strongly negatively in the case of competitive tournament conditions. The latter effect is remarkable because tournament-like settings involve better resources for successful people (scholarships, grants, special training, networking opportunities) in many practical contexts. The results show that all typical design features of incentive contracts come at a cost in terms of reduced fairness perceptions.

---

[7] This result contrasts with findings from Schildberg-Hörisch et al. (2020). They show in a lab experiment that subjects do not judge affirmative action rules (quota rules for low performance due to (i) bad luck, (ii) low productivity, and (iii) short working time), as less fair than no affirmative action.

Table 3. Fairness view analyses by income, political view, and gender

	Income		Political view		Gender	
	(1)	(2)	(3)	(4)	(5)	(6)
	High	Low	Conservative	Liberal	Female	Male
Gini	-1.08***	-1.48***	-1.03***	-1.41***	-1.25***	-1.26***
	(0.24)	(0.25)	(0.28)	(0.21)	(0.25)	(0.24)
Discrete bonus	-0.12	-0.76***	-0.12	-0.62***	-0.50*	-0.45**
Incentives	(0.23)	(0.25)	(0.27)	(0.22)	(0.26)	(0.23)
Tournament	-0.19	-0.83***	-0.66**	-0.44**	-0.28	-0.68***
	(0.14)	(0.24)	(0.27)	(0.21)	(0.23)	(0.24)
Time advantage	-0.28	-0.35	-0.09	-0.40	-0.29	-0.27
	(0.32)	(0.32)	(0.35)	(0.29)	(0.33)	(0.31)
Time handicap	-1.14***	-0.29	-1.43***	-0.28	-0.61*	-0.75**
	(0.31)	(0.33)	(0.36)	(0.29)	(0.31)	(0.33)
Tournament x	-1.30***	-1.47***	-1.66***	-1.16***	-1.77***	-1.03**
Time advantage	(0.45)	(0.48)	(0.52)	(0.42)	(0.46)	(0.46)
Tournament x	0.39	-0.21	0.70	-0.31	-0.61*	0.45
Time handicap	(0.44)	(0.45)	(0.54)	(0.39)	(0.31)	(0.46)
Exponential	0.31	-0.34	0.13	-0.11	-0.21	0.10
Scheme	(0.27)	(0.31)	(0.31)	(0.27)	(0.34)	(0.25)
Constant	8.53***	9.03***	8.72***	9.38***	8.96***	8.44***
	(0.49)	(0.43)	(0.49)	(0.39)	(0.43)	(0.46)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,102	1,080	794	1,388	1,033	1,142
R-squared	0.11	0.15	0.11	0.12	0.14	0.12

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regression on fairness judgments with robust standard errors in parentheses. High/Low income, Conservative/Liberal are median split. Controls include gender, level of education, personal income (categorical, high: ≥1.500$; not included in regression (1) and (2)), indicators for ethnic groups and for living in a rural area, political orientation (1: Liberal; 6: Conservative, not included in regression (3) and (4)), self-perceived social status, and employment status. Results are robust with or without control variables.

We next probe whether these patterns are robust for different subgroups of the population or whether they differ across demographic groups. Such differences can be driven, for example, by individual differences in personality (Fulmer and Shaw, 2018) or differences in intrinsic and non-monetary motivators (Benabou and Tirole, 2003, Erkal, Gangadharan, and Koh, 2018). Table 3 presents the results of estimating the full model (see Table 2, column 4) separately for high- and low-income individuals, politically conservative and liberal individuals, and males and females.

We observe that the above-identified pattern of effects is remarkably consistent across different groups. Still, there are some noteworthy differences. For example, respondents with a high income are less averse to bonus and tournament payment schemes. This could be interpreted as a self-serving bias attributing their achievements to a greater extent to their merits (see, for example, Fehr and Vollmann, 2022). In a similar direction, negative views of handicaps for successful workers seem to be driven predominantly by high-income and conservative respondents. These people may perceive these design features as violating basic meritocratic principles. Furthermore, male subjects view tournament incentives more negatively. This association with gender is a bit surprising, given the literature showing that females, in contrast to men, often avoid competitive payment schemes (Niederle and Vesterlund, 2007; Ors et al., 2013; Buser et al., 2014).

4. Study 2

4.1. Sample and Fairness Assessments

4.1.1 Respondents

We conducted Study 2 on Prolific with the exact same 16 scenarios as in Study 1. Again, we provided a short description of the task, the task requirement, as well as the expected payment for completing the task. After reading the study material, respondents answered four comprehension questions. If they failed a question, they first got a notification that their solution was incorrect and if they then failed a second time, they learned the correct solution. In any case, respondents could then proceed to the study. On average, 16 percent of the respondents failed to answer some or all comprehension questions correctly before proceeding to the fairness judgments. In Table OS5 of the Online Supplement, we show that these respondents respond less strongly (though qualitatively in the same way) to the scenarios as those who correctly answer the comprehension checks on their first attempt. This is what we would expect if respondents do not fully and accurately understand the consequences of the mechanisms. In the main text, we report results including all respondents, thus providing a conservative lower bound on the fairness judgment.[8]

We recruited 2,423 US respondents on Prolific in 2022. That is, on average 151 respondents per condition. Payment included a fixed payment of £1 and the opportunity to earn another £1.20 through incentivized fairness judgments (see below for details). At the median, respondents earned £1.40 for a task that took about 8 minutes. In the Online Supplement, we show that the sample is balanced with respect to observable characteristics such as gender, education, income, and ethnicity, among others, across our 16 payment scheme conditions.

4.1.2 Fairness Assessment

In Study 2, we first replicate the basic fairness judgment task employed in Study 1. We then additionally collect the following measures: First, we extend the overall fairness judgment by including judgments regarding outcome and procedural fairness of the payment scheme. The two items are measured on the same 11-point scale as the overall measure and are described as follows:

> Let us now consider two specific aspects of fairness. First, a payment mechanism may involve fair and balanced procedures, although the resulting distribution of payments to different workers may not be fair. That is what we call *procedural fairness*. Second, a payment mechanism may result in a distribution of payments to different workers that can be considered fair, irrespective of whether the procedures that led to the outcomes were fair. This is what we call *outcome fairness*.

---

[8] Note that in real-life setting it may well be the case that some workers do not fully comprehend the incentive structure of their contract, but nevertheless form their subjective fairness judgments.

After collecting respondents' non-incentivized fairness judgments, we next collect incentivized measures of fairness views using the method of Houser and Xiao (2011). The method involves incentivizing respondents to coordinate on the most salient fairness judgment by paying them for correctly indicating the majority (modal) judgment of all other respondents. The task was described as follows (overall fairness measure).

> We are now interested in how you think the majority of participants in this study think about the mechanism's overall fairness. So please indicate to what extent you think that other people consider this payment mechanism as fair. Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair). Important: The other participants in this study make the same judgment as you do. If your judgment matches the rating that is given by the largest number of other participants, you will receive an additional payment of £0.40 (approximately $0.44) (and the same is true for all other participants).

Subsequently, respondents indicated their incentivized choices also for the outcome and procedural fairness measures. In the following, we denote the incentivized measures as *HX fairness judgments* (Houser-Xiao).

4.2. Results: Consistency of Different Fairness Measures

The rightmost column of Table 1 shows that the observed pattern of basic overall fairness judgments closely replicates in the new sample. Multivariate regression analysis (see Table A1 in the Appendix) confirms that pay inequality, measured by the Gini coefficient, and tournament incentives hurt overall fairness judgments. As in Study 1, the discrete bonus scheme is judged less negatively than the tournament, but in Study 2 its effect is not significant. Both time advantage and handicap reduce fairness judgments, and we replicate the negative interaction of time advantages with tournament incentives. Thus, Study 1 results on non-incentivized fairness judgments turn out very robust and replicate closely in Study 2.

We can also compare the Study-2 non-incentivized fairness measures (overall, procedural fairness, and outcome fairness) to its new, incentivized HX-fairness counterparts. We find that the HX fairness measures are somewhat lower than the non-incentivized judgments in 15 of the 16 treatments for the overall fairness measure, in 13 treatments for the outcome fairness measure, and in all 16 treatments for the procedural fairness measure (see Table OS6). However, the basic patterns of relative judgments remain very similar across incentivized and non-incentivized measures.[9]

Having established the consistency of the different fairness measures in Study 1 and Study 2 (non-incentivized and incentivized), in the remainder of the section we will focus on the analysis of the incentivized overall HX fairness measure and on the differential effects on the incentivized HX procedural fairness and HX outcome fairness measures.

4.3. Results: Analysis of Incentivized Fairness Measures

4.3.1 HX overall fairness

The overall HX fairness measure shows the same pattern as the fairness measure reported in Study 1. In particular, the low-incentive piece rate (T1) elicits the highest HX fairness judgments (8.05), and the tournament with high incentives and time advantages (T14) elicits the lowest judgments (5.49). In the Online Supplement, we show histograms of the full distribution of HX fairness evaluations for all 16 treatments. Multivariate analyses for the overall HX fairness measure in Table 4 closely replicate the findings of Study 1 reported in Table 2. Here again, we find a significant negative effect for discrete bonus schemes, which is, however, lower than the negative effect of tournaments. Both time handicaps, and more strongly so time advantages, reduce the HX fairness judgment.

---

[9] Vesely (2015) shows that incentivized methods elicit fairness judgments that are indistinguishable from non-incentivized measures. Our results suggest that while relative fairness judgments across conditions are robust with respect to incentivization, levels might be affected by incentives.

These direct effects emerge consistently over all specifications. The interaction between tournament incentives and time advantages is again negative, but the standard errors are large.

Table 4. Multivariate analyses of HX fairness judgment – Study 2

	HX Fairness judgment			
	(1)	(2)	(3)	(4)
Gini	-1.03***	-1.03***	-1.03***	-1.03***
	(0.14)	(0.14)	(0.14)	(0.14)
Discrete bonus incentives	-0.24*	-0.25*	-0.28*	-0.28*
	(0.14)	(0.14)	(0.16)	(0.15)
Tournament	-0.49***	-0.50***	-0.46***	-0.46***
	(0.13)	(0.14)	(0.15)	(0.15)
Time advantage	-0.95***	-0.93***	-0.81***	-0.80***
	(0.13)	(0.13)	(0.18)	(0.18)
Time handicap	-0.62***	-0.63***	-0.65***	-0.65***
	(0.13)	(0.13)	(0.18)	(0.18)
Tournament x			-0.27	-0.26
Time advantage			(0.26)	(0.26)
Tournament x			0.05	0.04
Time handicap			(0.25)	(0.25)
Exponential scheme	-0.08	-0.09	-0.08	-0.09
	(0.22)	(0.22)	(0.22)	(0.22)
Constant	7.91***	8.21***	7.91***	8.20***
	(0.11)	(0.19)	(0.11)	(0.19)
Controls	No	Yes	No	Yes
Observations	2,423	2,423	2,423	2,423
R-squared	0.07	0.07	0.07	0.07
F-statistics (Discrete Bonus=Tournament)	5.56**	5.58**	1.04	1.11
F-statistics (Time bonus = Time handicap)	5.81**	5.22**	0.82	0.75

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regression on HX fairness judgment with robust standard errors in parentheses. Piece rate is the reference category. Controls include gender, level of education, personal income (categorical), indicators for ethnic groups and for living in a rural area, political orientation (1: Liberal; 6: Conservative), self-perceived social status, and employment status.

Conducting subsample analyses, as in Table 3, reveals that low-income (vs. high-income) and more liberal (vs. conservative) respondents are less concerned about time advantages and time handicaps (see Table A2 in the Appendix). Only the latter finding replicates Study 1 results. Females are less concerned about time advantages and more concerned about time handicaps. The subsample results suggest that, while the overall patterns are robust, they are not driven by certain groups holding specific fairness views.

4.3.2 HX Procedural and Outcome Fairness

We now analyze the role of procedural and outcome fairness in individuals' judgments. We have already seen that outcome inequality in the form of the Gini coefficient has a robust negative effect on fairness judgments. However, several other properties of the payment schemes have different effects on fairness judgments that are not captured by a purely outcome-based perspective.

We first observe that outcome fairness judgments are substantially and significantly lower than procedural fairness judgments in 13 of the 16 treatments and not significantly different in T9, T14 and T15 (see Table OS6). That is, overwhelmingly, the payment schemes are judged more positively from a procedural than from an outcome fairness perspective. Moreover, when regressing the overall HX fairness measure on the respective procedural and outcome measures, we find a significantly larger partial correlation with procedural than with outcome fairness (coefficients of 0.533 versus 0.370, $F=20.59$, $p<0.01$, see Table OS7). This suggests that the more positive procedural judgments also have a stronger effect on the overall judgment and supports the view that organizations can benefit from making procedural justice perspectives more salient among their employees (e.g., Trautmann and Wakker, 2010). This is particularly relevant when one cannot guarantee equal outcomes, for example, because of the need for payment schemes.

Table 5. Multivariate analyses of $\Delta_{\text{pro-out}}$

	$\boldsymbol{\Delta}_{\textbf{pro-out}}$			
	(1)	(2)	(3)	(4)
Gini	0.10**	0.10**	0.10**	0.10**
	(0.04)	(0.04)	(0.04)	(0.04)
Discrete bonus	-0.01	-0.01	-0.04	-0.04
	(0.04)	(0.04)	(0.05)	(0.05)
Tournament	-0.03	-0.03	0.01	0.01
	(0.04)	(0.04)	(0.04)	(0.04)
Time advantage	-0.09***	-0.10***	-0.01	-0.02
	(0.04)	(0.04)	(0.05)	(0.05)
Time handicap	-0.08**	-0.08**	-0.05	-0.05
	(0.04)	(0.04)	(0.05)	(0.05)
Exponential	-0.02	-0.02	-0.02	-0.02
	(0.05)	(0.05)	(0.05)	(0.05)
Tournament x			-0.16**	-0.16**
Time advantage			(0.07)	(0.07)
Tournament x			-0.06	-0.07
Time handicap			(0.07)	(0.07)
Constant	0.03	0.02	0.03	0.02
	(0.04)	(0.05)	(0.04)	(0.05)
Controls	No	Yes	No	Yes
Observations	2,394	2,394	2,394	2,394
R-squared	0.01	0.01	0.01	0.01
F-statistics (Discrete Bonus=Tournament)	0.54	0.73	1.09	0.99
F-statistics (Time bonus = Time handicap)	0.09	0.12	0.57	0.47

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. OLS regression on $\Delta_{\text{pro-out}}$ with robust standard errors in parentheses. Piece rate is the reference category. Controls include gender, level of education, personal income (categorical), indicators for ethnic groups and for living in a rural area, political orientation (1: Liberal; 6: Conservative), self-perceived social status, and employment status.

Next, we assess how the properties of the payment scheme differentially affect the two fairness properties. To this end, we define the relative distance between procedural and outcome fairness as:

$$\Delta_{pro\text{-}out} = \frac{HX\ procedural\ fairness - HX\ outcome\ fairness}{HX\ procedural\ fairness}$$

Note that $\Delta_{pro\text{-}out}$ increases if outcome fairness is judged relatively lower compared to procedural fairness, and decreases if procedural fairness is judged relatively lower compared to outcome fairness, for a certain payment scheme.

Table 5 presents the results. This analysis confirms the positive effect of the Gini coefficient on the relative difference between procedural and outcome fairness judgments, which means that higher inequality affects outcome fairness judgments more negatively. We do not observe significant differences between the discrete bonus and tournament scheme indicators. That is, they influence both fairness judgments similarly. In contrast, time advantages and time handicaps both strongly negatively affect the difference: they are perceived as more procedurally unfair. For the time advantage, this effect is particularly strong for the tournament condition. The result explains our finding that handicap schemes (i.e., those payment schemes that constrain the resources of more successful workers to level the playing field) are judged negatively, despite their desirable effect on equality. Although they positively impact equality, which is good for outcome fairness, they are seen as negative from a procedural fairness point of view. The latter effect is more substantial because procedural fairness more strongly affects respondents' overall fairness judgments.

5. Conclusion

Incentive contracts are an important feature in many employment settings where asymmetric information prevents the payment of fixed, output-independent wages.

Incentive contracts lead to inequality, success or failure, and often to competition between winners and losers (Verhaeghe, 2014).

To the best of our knowledge, the current paper is the first to systematically show how these issues affect fairness perceptions of different pay-for-performance schemes. In practical managerial and organizational settings, fairness perceptions will be relevant in determining employee satisfaction, theft, cooperation, and turnover. They need to be carefully monitored by the management to prevent unanticipated inefficiencies in employment relations. Furthermore, incentive benefits need to be traded off against unintended side effects due to violation of employees' fairness norms. Such trade-offs are not trivial, with potentially contradicting effects of payment scheme features on fairness. For example, contradicting the idea of merit-based payments, steeper incentives are uniformly judged negatively because of their implications for inequality between high and low-skilled workers. On the other hand, despite handicaps making outcomes more equal, they are perceived as unfair. They contradict some basic notion of procedural fairness and merit that is stronger than the effect of inequality. Indeed, these views are also captured in popular culture, for example, Kurt Vonnegut's (1968) famous short story *Harrison Bergeron*, which ridicules the use of handicaps to induce equality at a societal level. Importantly, these results hold both in non-incentivized individual judgments and incentivized measures of a shared understanding of fairness views.

There are some limitations of our study. First, we assume a constant mapping of fairness to employee behavior. However, different situations may be perceived as differing in the extent that strong incentives are necessary (e.g., depending on the degree of asymmetric information or misaligned preferences). This perceived necessity may influence the effect of fairness on behavior. Second, suppose workers self-select into industries and positions typically affiliated with different types of incentive contracts (Kosfeld and von Siemens, 2011). In that case, this selection may affect both the average fairness perception of the relevant subgroup and the linkage from fairness perceptions to behavior. Our sample split according to income, politics, and gender did not reveal a simple picture of how observables might affect such selection. Future work may fruitfully extend the current design to study the effects of self-selection more carefully and to assess better the role of context and the perceived necessity of incentives on fairness perceptions.

Appendix

A.1. Multivariate Analyses of Study-2 Non-incentivized Overall Fairness Measure

Table A1. Multivariate analyses of fairness judgment – Study 2

	Fairness judgment			
	(1)	(2)	(3)	(4)
Gini	-1.18***	-1.15***	-1.17***	-1.15***
	(0.16)	(0.16)	(0.16)	(0.16)
Discrete bonus incentives	-0.21	-0.20	-0.26	-0.26
	(0.15)	(0.15)	(0.17)	(0.17)
Tournament	-0.52***	-0.49***	-0.47***	-0.43**
	(0.15)	(0.15)	(0.17)	(0.17)
Time advantage	-1.12***	-1.10***	-0.82***	-0.78***
	(0.14)	(0.14)	(0.19)	(0.19)
Time handicap	-0.81***	-0.81***	-0.94***	-0.94***
	(0.14)	(0.14)	(0.19)	(0.19)
Tournament x			-0.59**	-0.63**
Time advantage			(0.29)	(0.29)
Tournament x			0.28	0.25
Time handicap			(0.28)	(0.28)
Exponential scheme	-0.24	-0.25	-0.24	-0.25
	(0.24)	(0.23)	(0.24)	(0.23)
Constant	8.31***	8.25***	8.31***	8.24***
	(0.12)	(0.22)	(0.12)	(0.22)
Controls	No	Yes	No	Yes
Observations	2,423	2,423	2,423	2,423
R-squared	0.07	0.10	0.08	0.10
F-statistics (Discrete bonus = Tournament)	7.23***	6.22**	1.20	0.73
F-statistics (Time bonus = Time handicap)	4.64**	3.89**	0.38	0.63

Notes: *** p<0.01, ** p<0.05, * p<0.1.

OLS regression on fairness judgments with robust standard errors in parentheses. Piece rate is the reference category. Controls include gender, level of education, personal income (categorical), indicators for ethnic groups and for living in a rural area, political orientation (1: Liberal; 6: Conservative), self-perceived social status, and employment status.

A.2. Subsample Analyses of Study-2 incentivized HX Overall Fairness Measure

Table A2. HX fairness judgments by income, political view, and gender – Study 2

	HX Fairness judgment					
	Income		**Political view**		**Gender**	
	(1)	(2)	(3)	(4)	(5)	(6)
	High	Low	Conservative	Liberal	Female	Male
Gini	-0.98***	-1.00***	-1.14***	-0.89***	-0.88***	-1.07***
	(0.22)	(0.21)	(0.27)	(0.17)	(0.21)	(0.20)
Discrete bonus	-0.18	-0.32	-0.26	-0.34*	-0.28	-0.32
Incentives	(0.23)	(0.23)	(0.28)	(0.19)	(0.22)	(0.22)
Tournament	-0.46**	-0.46**	-0.51**	-0.42**	-0.50**	-0.45**
	(0.22)	(0.22)	(0.26)	(0.19)	(0.23)	(0.20)
Time advantage	-1.12***	-0.64**	-1.11***	-0.62***	-0.57**	-0.94***
	(0.27)	(0.27)	(0.34)	(0.22)	(0.26)	(0.26)
Time handicap	-0.96***	-0.49**	-1.01***	-0.36*	-0.94***	-0.36
	(0.28)	(0.24)	(0.33)	(0.21)	(0.26)	(0.24)
Tournament x	-0.23	-0.20	-0.03	-0.42	-0.50	-0.10
Time advantage	(0.39)	(0.38)	(0.49)	(0.32)	(0.38)	(0.37)
Tournament x	0.30	-0.08	0.55	-0.31	-0.04	-0.02
Time handicap	(0.39)	(0.36)	(0.46)	(0.31)	(0.39)	(0.34)
Exponential	-0.50	0.34	-0.04	-0.19	-0.25	0.06
Scheme	(0.33)	(0.32)	(0.39)	(0.27)	(0.31)	(0.31)
Constant	8.28***	8.11***	9.00***	7.96***	8.11***	8.21***
	(0.35)	(0.28)	(0.37)	(0.23)	(0.28)	(0.26)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,094	1,078	699	1,635	1,071	1,304
R-squared	0.09	0.07	0.10	0.07	0.08	0.08

Notes: *** p<0.01, ** p<0.05, * p<0.1

OLS regression on fairness judgment with robust standard errors in parentheses. High/Low income, Conservative/Liberal are median split. Controls include gender, level of education, personal income (categorical, high: ≥1.500$; not included in regression (1) and (2)), indicators for ethnic groups and for living in a rural area, political orientation (0: Liberal; 1: Conservative, not included in regression (3) and (4)), self-perceived social status, and employment status. Results are robust with or without control variables.

References

Ambrose, M.L., Seabright, M. A., & Schminke, M. (2002). Sabotage in the worksplace: The role of organizational injustice. *Organizational Behavior and Human Decision Processes* 89, 947–965.

Ambuehl, S., Niederle, M., & Roth, A. E. (2015). More money, more problems? can high pay be coercive and repugnant? *American Economic Review*, *105*(5), 357–360.

Ambuehl, S., & Ockenfels, A. (2017). The ethics of incentivizing the uninformed: A vignette study. *American Economic Review*, *107*(5), 91–95.

Arechar, A. A., Gächter, S., & Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, *21*(1), 99–131.

Arechar, A. A., Kraft-Todd, G. T., & Rand, D. G. (2017). Turking overtime: how participant characteristics and behavior vary over time and day on Amazon Mechanical Turk. *Journal of the Economic Science Association*, *3*(1), 1–11.

Bartling, B., Cappelen, A. W., Ekström, M., Sorensen, E., & Tungodden, B. (2018). Fairness in Winner-Take-All Markets. NHH Dept. of Economics Discussion Paper.

Bénabou, R. & Jean Tirole, J. (2003). Intrinsic and Extrinsic Motivation, *The Review of Economic Studies*, 70 (3), 489–520.

Bénabou, R., & Tirole, J. (2016). Bonus culture: Competitive pay, screening, and multitasking. *Journal of Political Economy*, *124*(2), 305–370.

Breza, E., Kaur, S., & Shamdasani, Y. (2018). The morale effects of pay inequality. *Quarterly Journal of Economics*, *133*(2), 611–663.

Bell, B., Pedemonte, S., Van Reenen, J. (2021). CEO Pay and the Rise of Relative Performance Contracts: A Question of Governance? *Journal of the European Economic Association*, *19*, 2513–2542.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics*, *129*(3), 1409–1447.

Card, D., Mas, A., Moretti, E., and Saez, E. (2012). Inequality at Work: the Effect of Peer Salaries on Job Satisfaction, *American Economic Review*, 102(6), 2981-3003.

Cappelen, A.W., Falch, R., & Tungodden, B. (2020). Fair and Unfair Income Inequality. In: Zimmermann, K. (ed.), *Handbook of Labor, Human Resources and Population Economics*. Springer, Cham.

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

Coppock, A., & McClellan, O. A. (2019). Validating the demographic, political, psychological, and experimental results obtained from a new source of online survey respondents. *Research and Politics*, *6*(1), 2053168018822174.

Dahl, G., and M. Ransom. (1999). "Does Where You Stand Depend on Where You Sit? Tithing Donations and Self-Serving Beliefs," *American Economic Review* 89, 703–27.

Dillard, J. F., & Fisher, J. G. (1990). Compensation schemes, skill level, and task performance: An experimental examination. *Decision Sciences*, 21(1), 121–137.

Eriksson, T., Poulsen, A., & Villeval, M.-C. (2009). Feedback and incentives: Experimental evidence. *Labour Economics*, 16, 679–688.

Erkal, N., Gangadharan, L., & Koh, B. H. (2018). Monetary and non-monetary incentives in real-effort tournaments. *European Economic Review*, *101*, 528-545.

Fehr, D. (2018). Is increasing inequality harmful? Experimental evidence. *Games and Economic Behavior*, *107*, 123–134.

Fehr, D., & Vollmann, M. (2022). Misperceiving Economic Success: Experimental Evidence on Meritocratic Beliefs and Inequality Acceptance, CESifo Working Paper 9983.

Fehr, D., Rau, H., Trautmann, S. T., & Xu, Y. (2020). Inequality, fairness and social capital. *European Economic Review*, *129*, 103566.

Finseraas, H., Johnsen, Å. A., Kotsadam, A., & Torsvik, G. (2016). Exposure to female colleagues breaks the glass ceiling—Evidence from a combined vignette and field experiment. *European Economic Review*, *90*, 363–374.

Fulmer, I. S., & Shaw, J. D. (2018). Person-Based Differences in Pay Reactions: A Compensation-Activation Theory and Integrative Conceptual Review. *Journal of Applied Psychology*, 103, 939–958.

Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review,* 102, 469–503.

Greenberg, J. (1987). Reactions to Procedural Injustice in Payment Distributions: Do the Means Justify the Ends? *Journal of Applied Psychology* 72, 55-61.

Gupta, N., & Shaw, J. D. (2014). Employee compensation: The neglected area of HRM research. *Human Resource Management Review* 24, 1–4.

Harbring, C., Irlenbusch, B., Kräkel, M., & Selten, R. (2007). Sabotage in corporate contests–an experimental analysis. *International Journal of the Economics of Business*, *14*(3), 367-392.

Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, *14(3)*, 399–425.

Houser, D. and Xiao, E. (2011). Classification of natural language messages using a coordination game. *Experimental Economics* 14, 1–14.

Jawahar, I. M. & Stone, T. (2011). Fairness Perceptions and Satisfaction with Components of Pay Satisfaction. *Journal of Managerial Psychology*, 26, 297–312.

Kahneman, D., Knetsch, J., and R. Thaler. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review* 76, 728–41.

Kao, A.C., Jager, A. J., Koenig, B. A., Moller, A. C., Tutty, M.A., Williams, G. C., & Wright, S. M. (2018). Physician Perception of Pay Fairness and its Association with Work Satisfaction, Intend to Leave Practice, and Personal Health. *Journal of General Internal Medicine, 33*, 812–817.

Kleinlercher, D., & Stöckl, T. (2018). On the provision of incentives in finance experiments. *Experimental Economics*, *21*(1), 154-179.

Konow, J. (2003). Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature*, *41*(4), 1188–1239.

Kosfeld, M., & Von Siemens, F. A. (2011). Competition, cooperation, and corporate culture. *RAND Journal of Economics*, *42*(1), 23–43.

Kőszegi, B. (2014). Behavioral Contract Theory. *Journal of Economic Literature, 52*, 1075–1118.

Kübler, D., Schmid, J., & Stüber, R. (2018). Gender Discrimination in Hiring Across Occupations: A Nationally-Representative Vignette Study. *Labour Economics*, *55*, 215–229.

Lazear, E. P. (2018). Compensation and Incentives in the Workplace. *Journal of Economic Perspectives*, *32*(3), 195-214.

Lemieux, T., Macleod, W., & Parent, D. (2009). Performance Pay and Wage Inequality. *The Quarterly Journal of Economics, 124*(1), 1-49.

Milgrom, P. & Roberts, J. (1992). *Economics, Organization & Management*. Pearson.

Mullinix, K., Leeper, T., Druckman, J., & Freese, J. (2015). The Generalizability of Survey Experiments. *Journal of Experimental Political Science*, 2(2), 109-138.

Murayama, K., & Elliot, A. J. (2012). The competition–performance relation: A meta-analytic review and test of the opposing processes model of competition and performance. *Psychological Bulletin*, *138*(6), 1035.

Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, *122*(3), 1067–1101.

Ors, E., Palomino, F., & Peyrache, E. (2013). Performance gender gap: does competition matter? *Journal of Labor Economics*, 31(3), 443–499.

Oyer, P. (2000). A theory of sales quotas with limited liability and rent sharing. *Journal of Labor Economics*, *18*(3), 405-426.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5(5), 411–419.

Pan, Y., Pikulina, E. S. Siegel, S. and Wang, T.Y. (2022). Do Equity Markets Care about Income Inequality? Evidence from Pay Ratio Disclosure. Journal of Finance 77, 1371–1411.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, 37, 7–63.

Roemer, J. E. (1998). *Equality of opportunity*. Harvard University Press.

Samek, A., Kapteyn, A., and Gray, A. (2020). Using Vignettes to Improve Understanding of Social Security and Annuities, *NBER Working Paper No. w26176.*

Schildberg-Hörisch, H., Trieu, C., & Willrodt, J. (2020). Perceived Fairness and Consequences of Affirmative Action Policies, *IZA Discussion Paper 13202*.

Snowberg, E., & Yariv, L. (2021). Testing the Waters: Behavior across Participant Pools. *American Economic Review*, *111*(2), 687–719.

Trautmann, S. T. (forthcoming). Procedural fairness and equality of opportunity. *Journal of Economic Surveys.*

Trautmann, S. T. and Wakker, P. P. (2010). Process Fairness and Dynamic Consistency. *Economics Letters* 109, 187–189.

Verhaeghe, P. (2014). *What about me? The struggle for identity in a market-based society*, Scribe, London, UK.

Vesely, S. (2015). Elicitation of normative and fairness judgments: Do incentives matter? *Judgment and Decision Making* 10, 191–197.

Vonnegut, K. (1968). *Welcome to the Monkey House*. Dial Press, NY.

Xu, Y., & Marandola, G. (2022). The (Negative) Effects of Inequality on Social Capital. *Working Paper*.

ONLINE SUPPLEMENT

Contents

OS.1 INSTRUCTIONS

Note: We show the correct answers to the comprehension questions in curly brackets after each question.

General Instructions

Dear participant,

Thank you for taking part in the questionnaire. Please first read all information carefully, and then answer the question.

Two workers have to perform a task on a computer. They will be paid according to some payment scheme, which will be described in detail later. The task requires the workers to place a slider in the central position of a slider bar. The slider initially appears at the far-left position in the bar. One has to use the mouse to move the slider to the target position.

Initial position:

Target position:

The keyboard has been disabled in order to make the task sufficiently challenging. The task consists of 48 sliders to be correctly placed. Placing 48 sliders within a time limit (described later) is called a round. The task is repeated for ***four*** rounds. The actual screen the workers have seen is like the following.

Payment scheme: Piece rate

Low inequality with base pay (Low_with base pay)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. For each correctly positioned slider within the time limit per round, workers receive \$0.03. Additionally, they also receive \$1 each round for participating in the task. Their payment per round can be summarized as \$1 + \$0.03 × (Number of sliders).

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. An average-skilled worker would, therefore, receive \$1.60 (=\$0.03*20 + \$1) per round. The total earnings after four rounds would be \$6.40 (=\$1.60*4) for an average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of \$1.87 (=\$0.03*29+\$1) per round, or \$7.48 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of \$1.33 (=\$0.03*11+\$1) per round, or \$5.32 in total. Thus, the difference in total earnings between the two workers is \$2.16 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in \$)? ____________ {1.39}
What are the round 1 earnings of worker B (in \$)? ____________ {1.81}

How many seconds does worker A have available for the task in round 2? _________ {120}
How many seconds does worker B have available for the task in round 2? _________ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Piece rate

Medium inequality without base pay (Medium_no base pay)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. For each correctly positioned slider within the time limit per round, workers receive \$0.08. Their payment per round can be summarized as \$0.08 × (Number of sliders).

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. An average-skilled worker would, therefore, receive \$1.60 (=\$0.08*20) per round. The total earnings after four rounds would be \$6.40 (=\$1.60*4) for an average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of \$2.32 (=\$0.08*29) per round, or \$9.28 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of \$0.88 (=\$0.08*11) per round, or \$3.52 in total. Thus, the difference in total earnings between the two workers is \$5.76 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in \$)? ______________ {1.04}
What are the round 1 earnings of worker B (in \$)? ______________ {2.16}
How many seconds does worker A have available for the task in round 2? ________ {120}
How many seconds does worker B have available for the task in round 2? ________ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Piece rate

High inequality without base pay but with an entrance fee (High_with entrance fee)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. The task requires a mandatory participation fee of $6 to be paid by each worker (or $1.50 per round). For each correctly positioned slider within the time limit per round, workers receive $0.15. Their payment per round can be summarized as $0.15×(Number of sliders).

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. An average-skilled worker would, therefore, receive $1.50 (=$0.15*20-$1.50) per round. The total earnings after four rounds would be $6 (=$1.50*4) for an average-skilled worker (already accounting for the participation fee).

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $2.85 (=$0.15*29-$1.50) per round, or $11.40 in total (already accounting for the participation fee). The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $0.15 (=$0.15*11-$1.50) per round, or $0.60 in total (already accounting for the participation fee). Thus, the difference in total earnings between the two workers is $10.80 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:
Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1. Please take into account the entrance fee of $1.50 per round.
What are the round 1 earnings of worker A (in $)? _____________ {0.45}
What are the round 1 earnings of worker B (in $)? _____________ {2.55}
How many seconds does worker A have available for the task in round 2? _____ {120}
How many seconds does worker B have available for the task in round 2? ______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Piece rate

Extreme inequality, no base pay, no entrance fee, exponential reward function (Extreme)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. The total earnings for the task in a round depend on the number of sliders workers can correctly place within the time limit, which is summarized in the table below. The upper rows show the number of correctly placed sliders and the lower rows show the total payoffs in dollar.

Sliders	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Total Payoffs	0.01	0.03	0.04	0.06	0.08	0.11	0.14	0.17	0.21	0.26	0.32	0.38
Sliders	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
Total Payoffs	0.46	0.55	0.66	0.78	0.92	1.09	1.29	1.52	1.79	2.11	2.48	2.91
Sliders	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
Total Payoffs	3.42	4.01	4.70	5.52	6.46	7.58	8.87	10.39	12.17	14.25	16.69	19.54
Total Sliders	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**
Payoffs	22.87	26.77	31.34	36.67	42.92	50.23	58.78	68.78	80.49	94.18	110.21	128.95

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. An average-skilled worker would therefore receive $1.52 per round for doing the task (see the table). The total earnings after four rounds would be $6.08 for an average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $6.46 (see table) per round, or $25.84 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $0.32 (see table) per round, or $1.28 in total. Thus, the difference in total earnings between the two workers is $24.56 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? _____________ {0.46}
What are the round 1 earnings of worker B (in $)? _____________ {4.70}
How many seconds does worker A have available for the task in round 2? _______ {120}
How many seconds does worker B have available for the task in round 2? _______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

Low inequality, no time advantage, no time handicap (Low)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $1.40 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $1.20.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $1.82 (=$0.02*21+$1.40) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $1.60 (=$0.02*20+$1.20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds. The total earnings after four rounds would then be $6.84 (=$1.82*2+$1.60*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $1.98 (=$0.02*29+$1.40) per round, or $7.92 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $1.42 (=$0.02*11+$1.20) per round, or $5.68 in total. Thus, the difference in total earnings between the two workers is $2.24 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:
Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ____________ {1.46}
What are the round 1 earnings of worker B (in $)? ____________ {1.94}
How many seconds does worker A have available for the task in round 2? _______ {120}
How many seconds does worker B have available for the task in round 2? _______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

High inequality, no time advantage, no time handicap (High)

The payment mechanism is as follows:
Both workers perform the task individually and independently. The time limit per round to work on the sliders is 120 seconds. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $2.50 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $0.00.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $2.92 (=$0.02*21+$2.50) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $0.40 (=$0.02*20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds. The total earnings after four rounds would then be $6.64 (=$2.92*2+$0.40*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $3.08 (=$0.02*29+$2.50) per round, or $12.32 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $0.22 (=$0.02*11) per round, or $0.88 in total. Thus, the difference in total earnings between the two workers is $11.44 after four rounds if one worker is skilled and the other is not. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ____________ {0.26}

What are the round 1 earnings of worker B (in $)? _____________ {3.04}
How many seconds does worker A have available for the task in round 2? _______ {120}
How many seconds does worker B have available for the task in round 2? _______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

Low inequality, with time advantage (Low_time bonus)

The payment mechanism is as follows:
Both workers perform the task individually and independently. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $1.40 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $1.20.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, if the worker correctly places 21 or more sliders, 6 seconds are added to his current time budget. If the worker correctly places less than 21 sliders, 6 seconds are subtracted from his current time budget. See the example for illustration.

Example: A worker places 20 sliders correctly in round 1. His time budget in round 2 will be 114 seconds. In round 2 he then places 21 sliders correctly. His time budget in round 3 will be 120 seconds. In round 3 he then places 22 sliders correctly. His time budget in round 4 will be 126 seconds.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place many sliders and to reach the threshold for the bonus payment, in subsequent rounds.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $1.82 (=$0.02*21+$1.40) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $1.60 (=$0.02*20+$1.20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds.

The total earnings after four rounds would then be \$6.84 (=\$1.82*2+\$1.60*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of \$1.98 (=\$0.02*29+\$1.40) per round, or \$7.92 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of \$1.42 (=\$0.02*11+\$1.20) per round, or \$5.68 in total. Thus, the difference in total earnings between the two workers is \$2.24 after four rounds if one worker is skilled and the other is not.

Note that these calculations do not account for increased or reduced time budgets due to managing or not managing to reach the threshold of 21 sliders in earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time after passing the threshold, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time after not passing the threshold. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.

What are the round 1 earnings of worker A (in \$)? ______________ {1.46}

What are the round 1 earnings of worker B (in \$)? ______________ {1.94}

How many seconds does worker A have available for the task in round 2? _______ {114}

How many seconds does worker B have available for the task in round 2? _______ {126}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

High inequality, with time advantage (High_time bonus)

The payment mechanism is as follows:
Both workers perform the task individually and independently. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $2.50 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $0.00.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, if the worker correctly places 21 or more sliders, 6 seconds are added to his current time budget. If the worker correctly places only less than 21 sliders, 6 seconds are subtracted from his current time budget. See the example for illustration.

Example: A worker places 20 sliders correctly in round 1. His time budget in round 2 will be 114 seconds. In round 2 he then places 21 sliders correctly. His time budget in round 3 will be 120 seconds. In round 3 he then places 22 sliders correctly. His time budget in round 4 will be 126 seconds.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place many sliders and to reach the threshold for the bonus payment, in subsequent rounds.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $2.92 (=$0.02*21+$2.50) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $0.40 (=$0.02*20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds. The total earnings after four rounds would then be $6.64 (=$2.92*2+$0.40*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $3.08 (=$0.02*29+$2.50) per round, or $12.32 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $0.22 (=$0.02*11) per round, or $0.88 in total. Thus, the difference in total earnings between the two workers is $11.44 after four rounds if one worker is skilled and the other is not.

Note that these calculations do not account for increased or reduced time budgets due to managing or not managing to reach the threshold of 21 sliders in earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time after passing the threshold, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time after not passing the threshold. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ____________ {0.26}
What are the round 1 earnings of worker B (in $)? ____________ {3.04}
How many seconds does worker A have available for the task in round 2? ______ {114}
How many seconds does worker B have available for the task in round 2? ______ {126}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

Low inequality, with time handicap (Low_time handicap)

The payment mechanism is as follows:
Both workers perform the task individually and independently. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $1.40 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $1.20.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, if the worker correctly places 21 or more sliders, 6 seconds are subtracted from his current time budget, a time handicap. If the worker correctly places less than 21 sliders, 6 seconds are added to his current time budget, a time benefit. See the example for illustration.

Example: A worker places 20 sliders correctly in round 1. His time budget in round 2 will be 126 seconds. In round 2 he then places 21 sliders correctly. His time budget in round 3 will be 120 seconds. In round 3 he then places 22 sliders correctly. His time budget in round 4 will be 114 seconds.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place many sliders and to reach the threshold for the bonus payment, in subsequent rounds.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $1.82 (=$0.02*21+$1.40) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $1.60 (=$0.02*20+$1.20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds. The total earnings after four rounds would then be $6.84 (=$1.82*2+$1.60*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $1.98 (=$0.02*29+$1.40) per round, or $7.92 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $1.42 (=$0.02*11+$1.20) per round, or $5.68 in total. Thus, the difference in total earnings between the two workers is $2.24 after four rounds if one worker is skilled and the other is not.

Note that these calculations do not account for increased or reduced time budgets due to managing or not managing to reach the threshold of 21 sliders in earlier rounds. This will lead, on average, to lower numbers of sliders compared to the number for 120 seconds for those with less time after passing the threshold, and on average larger numbers of sliders compared to the number for 120 seconds for those with more time after not passing the threshold. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.

What are the round 1 earnings of worker A (in $)? ______________ {1.46}

What are the round 1 earnings of worker B (in $)? ______________ {1.94}

How many seconds does worker A have available for the task in round 2? ________ {126}

How many seconds does worker B have available for the task in round 2? ________ {114}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism.

To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Discrete bonus target

High inequality, with time handicap (High_time handicap)

The payment mechanism is as follows:
Both workers perform the task individually and independently. For each correctly positioned slider within the time limit per round, workers receive $0.02. Additionally, in each round, there will also be a bonus payment of $2.50 if a worker correctly positioned 21 sliders or more. If the total number of sliders placed is below 21, then the bonus payment is $0.00.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, if the worker correctly places 21 or more sliders, 6 seconds are subtracted from his current time budget, a time handicap. If the worker correctly places less than 21 sliders, 6 seconds are added to his current time budget, a time benefit. See the example for illustration.

Example: A worker places 20 sliders correctly in round 1. His time budget in round 2 will be 126 seconds. In round 2 he then places 21 sliders correctly. His time budget in round 3 will be 120 seconds. In round 3 he then places 22 sliders correctly. His time budget in round 4 will be 114 seconds.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place many sliders and to reach the threshold for the bonus payment, in subsequent rounds.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. If an average-skilled worker just manages to place 21 sliders in a round, his payoff would be $2.92 (=$0.02*21+$2.50) this round. If, on the other hand, he just fails to get 21 sliders in a round, his payoff would be $0.40 (=$0.02*20) for this round. Suppose that he manages to correctly place 21 sliders in two rounds, but only places 20 sliders in the other two rounds. The total earnings after four rounds would then be $6.64 (=$2.92*2+$0.40*2) for the average-skilled worker.

The most skilled workers (top 10%) can place about 29 sliders in 120 seconds on average, amounting to a payment of $3.08 (=$0.02*29+$2.50) per round, or $12.32 in total. The least skilled workers (bottom 10%) can place about 11 sliders in 120 seconds on average, amounting to a payment of $0.22 (=$0.02*11) per round, or $0.88 in total. Thus, the difference in total earnings between the two workers is $11.44 after four rounds if one worker is skilled and the other is not.

Note that these calculations do not account for increased or reduced time budgets due to managing or not managing to reach the threshold of 21 sliders in earlier rounds. This will lead, on average, to lower numbers of sliders compared to the number for 120 seconds for those with less time after passing the threshold, and on average larger numbers of sliders compared to the number for 120 seconds for those with more time after not passing the threshold. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ______________ {0.26}
What are the round 1 earnings of worker B (in $)? ______________ {3.04}
How many seconds does worker A have available for the task in round 2? _______ {126}
How many seconds does worker B have available for the task in round 2? _______ {114}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

Low inequality, no time advantage, no time handicap (Low)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The time limit per round to work on the sliders is 120 seconds. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $2. The loser receives $1.40. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.80 (=$2*2+$1.40*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $8 (=$2*4), while his opponent earns $5.60 (=$1.40*4), a difference of $2.40 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:
Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ______________ {1.40}
What are the round 1 earnings of worker B (in $)? ______________ {2.00}
How many seconds does worker A have available for the task in round 2? _______ {120}
How many seconds does worker B have available for the task in round 2? _______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

High inequality, no time advantage, no time handicap (High)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The time limit per round to work on the sliders is 120 seconds. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $3. The loser receives $0.30. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.60 (=$3*2+$0.30*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $12 (=$3*4), while his opponent earns $1.20 (=$0.30*4), a difference of $10.80 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:
Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ____________ {0.30}
What are the round 1 earnings of worker B (in $)? ____________ {3.00}

How many seconds does worker A have available for the task in round 2? _______ {120}
How many seconds does worker B have available for the task in round 2? _______ {120}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

Low inequality, with time advantage (Low_time bonus)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $2. The loser receives $1.40. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, for the winner, 6 seconds are added to his current time budget. For the loser, 6 seconds are subtracted from his current time budget. See the example for illustration.

Example: Worker A places more sliders correctly in round 1 than worker B. His time budget in round 2 will be 126 seconds, worker B's time budget will be 114 seconds. In round 2 worker A again places more sliders correctly than worker B. His time budget in round 3 will be 132 seconds, worker B's time budget will be 108 seconds. In round 3, worker A now places less sliders correctly than worker B. His time budget in round 4 will be 126 seconds, worker B's time budget will be 114.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place more sliders than the opponent.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.80 (=$2*2+$1.40*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $8 (=$2*4), while his opponent earns $5.60 (=$1.40*4), a difference of $2.40 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

Note that these calculations do not account for increased or reduced time budgets due to winning or losing earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ______________ {1.40}
What are the round 1 earnings of worker B (in $)? ______________ {2.00}
How many seconds does worker A have available for the task in round 2? ________ {114}
How many seconds does worker B have available for the task in round 2? ________ {126}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

High inequality, with time advantage (High_time bonus)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $3. The loser receives $0.30. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, for the winner, 6 seconds are added to his current time budget. For the loser, 6 seconds are subtracted from his current time budget. See the example for illustration.

Example: Worker A places more sliders correctly in round 1 than worker B. His time budget in round 2 will be 126 seconds, worker B's time budget will be 114 seconds. In round 2 worker A again places more sliders correctly than worker B. His time budget in round 3 will be 132 seconds, worker B's time budget will be 108 seconds. In round 3, worker A now places less sliders correctly than worker B. His time budget in round 4 will be 126 seconds, worker B's time budget will be 114.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place more sliders than the opponent.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.60 (=$3*2+$0.30*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $12 (=$3*4), while his opponent earns $1.20 (=$0.30*4), a difference of $10.80 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

Note that these calculations do not account for increased or reduced time budgets due to winning or losing earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.
What are the round 1 earnings of worker A (in $)? ______________ {0.30}
What are the round 1 earnings of worker B (in $)? ______________ {3.00}
How many seconds does worker A have available for the task in round 2? _______ {114}
How many seconds does worker B have available for the task in round 2? _______ {126}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

Low inequality, with time handicap (Low_time handicap)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $2. The loser receives $1.40. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, for the winner, 6 seconds are subtracted from his current time budget, a time handicap. For the loser, 6 seconds are added to his current time budget, a time benefit. See the example for illustration.

Example: Worker A places more sliders correctly in round 1 than worker B. His time budget in round 2 will be 114 seconds, worker B's time budget will be 126 seconds. In round 2 worker A again places more sliders correctly than worker B. His time budget in round 3 will be 108 seconds, worker B's time budget will be 132 seconds. In round 3, worker A now places less sliders correctly than worker B. His time budget in round 4 will be 114 seconds, worker B's time budget will be 126.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place more sliders than the opponent.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.80 (=$2*2+$1.40*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $8 (=$2*4), while his opponent earns $5.60 (=$1.40*4), a difference of $2.40 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

Note that these calculations do not account for increased or reduced time budgets due to winning or losing earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.

What are the round 1 earnings of worker A (in $)? ______________ {1.40}

What are the round 1 earnings of worker B (in $)? ______________ {2.00}

How many seconds does worker A have available for the task in round 2? ________ {126}

How many seconds does worker B have available for the task in round 2? ________ {114}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism.

To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Payment scheme: Tournament

High inequality, with time handicap (High_time handicap)

The payment mechanism is as follows:
The two workers form a competitive group. That is, the goal of the workers in each round is to put a larger number of sliders to the target position within the given time limit than the other worker in the group. The worker who placed more sliders in a group is the winner of this round. It does not matter how large the difference actually is, it only matters who of the two workers positioned more sliders correctly. In each of the four rounds, the winner of the round receives $3. The loser receives $0.30. In case of a tie where both players have the same number of correctly positioned sliders, the winner for this round will be randomly determined.

The initial time limit in the first round is 120 seconds. Depending on performance, the time budget is reduced or increased in the subsequent round. In particular, for the winner, 6 seconds are subtracted from his current time budget, a time handicap. For the loser, 6 seconds are added to his current time budget, a time benefit. See the example for illustration.

Example: Worker A places more sliders correctly in round 1 than worker B. His time budget in round 2 will be 114 seconds, worker B's time budget will be 126 seconds. In round 2 worker A again places more sliders correctly than worker B. His time budget in round 3 will be 108 seconds, worker B's time budget will be 132 seconds. In round 3, worker A now places less sliders correctly than worker B. His time budget in round 4 will be 114 seconds, worker B's time budget will be 126.

In general, a shorter time budget to work on the task in a round makes it more difficult, and a longer time budget makes it less difficult, to correctly place more sliders than the opponent.

Our data containing more than 600 workers show that workers can correctly place about 20 sliders per round on average within a time limit of 120 seconds. How much a worker can earn per round depends on his own performance, the performance of his opponent, and in the event of a tie, a flip of a coin. If both workers in a group win two of the four rounds, then each would receive $6.60 (=$3*2+$0.30*2).

The most skilled workers (top 10%) can place about 29 sliders per round on average, while the least skilled workers (bottom 10%) can place about 11 sliders per round on average. If a worker wins four rounds in a row, then he can earn $12 (=$3*4), while his opponent earns $1.20 (=$0.30*4), a difference of $10.80 after four rounds. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

Note that these calculations do not account for increased or reduced time budgets due to winning or losing earlier rounds. This will lead, on average, to larger numbers of sliders compared to the number for 120 seconds for those with more time, and on average lower numbers of sliders compared to the number for 120 seconds for those with less time. Note also that both workers know their own and the other worker's performance and payoffs in each round and in total.

i) If all aspects of the task and the payment of workers are clear, please answer the following questions about the task:

Assume that worker A placed 13 sliders correctly and worker B placed 27 sliders correctly in round 1.

What are the round 1 earnings of worker A (in $)? ______________ {0.30}

What are the round 1 earnings of worker B (in $)? ______________ {3.00}

How many seconds does worker A have available for the task in round 2? _______ {126}

How many seconds does worker B have available for the task in round 2? _______ {114}

ii) Let us now consider the overall picture of the payment mechanism described above. Different payment mechanisms to reward people for their work may be considered more or less fair. We are interested in how people think about the above described mechanism. To what extent do you think this payment mechanism is fair? Please indicate your fairness judgment on a scale from 0 (completely unfair) to 10 (completely fair).

Completely unfair O O O O O O O O O O O Completely fair

Sociodemographic Questionnaire

1. Please indicate your gender:
    A. Female
    B. Male
    C. Prefer not to answer or others

2. What is your year of birth (YYYY): _______________.

3. What is your highest degree of completed education?
    A. Not completed high school.
    B. High school.
    C. 2-year college degree
    D. 4-year college degree
    E. Master's degree
    F. Doctoral degree
    G. Professional degree (JD, MD)

4. What is your ethnicity?
    A. White/European-American,
    B. Black/African-American.
    C. Asian/Asian-American/Pacific Islander.
    D. Hispanic/Latino.
    E. Other:______

5. On a continuum from liberal to conservative, how would you describe your political beliefs?
    A. Strongly liberal.
    B. Moderately liberal.
    C. Slightly liberal.
    D. Slightly conservative.
    E. Moderately conservative.
    F. Strongly conservative.

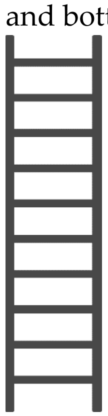
6. Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or something else?
    A. Republican.
    B. Democrat.
    C. Independent.
    D. Other: ___

7. Do you live in the United States? Yes/ No

8. In which state do you live? _____

9. What describes best the area where you currently live in?
   A. Urban area.
   B. Rural area.

10. This ladder shows where people in the U.S. stand in society. At the top of the ladder (marked "10") are those people who are doing best, those who have the most money, have the best education and are most respected. At the bottom of the ladder (marked "1") are those who are worst off, have the least money, have the worst education, and are the least respected. Where would you place yourself on this ladder relative to the people at the top and bottom? Please mark the box next to the rung you think you are standing on!



10
9
8
7
6
5
4
3
2
1

11. What is your current employment status?
   A. Currently employed.
   B. Currently unemployed (but employed before)
   C. Never employed

12. What is your approximate monthly disposable income in USD?
   A. Up to $1,500
   B. $1.500 to $4,500
   C. $4,500 or more
   D. Don’t know or prefer not to answer.

13. In the area where you live, your family financial situation is…
   A. far below the average.

B. below the average.
C. at the average.
D. higher than the average.
E. far higher than the average
F. Don’t know or prefer not to answer.

14. Consider all income sources in your family, including all sources from your family members, how easy it is to break-even?

A. Very Difficult.
B. Somewhat difficult.
C. Neither Difficult nor easy.
D. Somewhat Easy.
E. Very Easy.
F. Don’t know or prefer not to answer.

OS.2 STUDY 1: ATTRITION AND BALANCE STATISTICS

To rule out selective attrition after reading the scenarios as an explanation for our results, we analyze the pattern of attrition across the 16 conditions, and test whether the sample is still balanced across conditions with respect to the characteristics of the respondents. Table OS1 gives an overview of dropouts by treatment. As shown in the table, there is some variation in dropout rates. Table OS2 shows pair-wise comparisons of dropout rates between conditions, using the Fisher's exact test. Overall, there are few systematic differences, except for conditions T3 and T8.

We further examine whether the observed differences affected the sample composition in the conditions. We regress each treatment dummy separately on each of the 7 demographic variables we collected: $Treatment_i = \beta_0 + \beta_{i,j} Covariate_{i,j} + \epsilon_{i,j}\ \forall\ i = 1, \ldots, 16;\ j = Demographic\ variable$. The p-values of the estimated coefficients (OLS, robust standard errors) are reported in Table OS3. There are in total merely 5 instances where the p-value is smaller than 0.05 (not corrected for multiple testing), out of 112 regressions. If we correct for multiple testing using Bonferroni's correction, then none of the estimated coefficients is significantly different from zero. In the last row, we also report F-tests when including all covariates into the regression.

There is no instance in which the p-value is less than the critical value of 0.05. We thus conclude that there was no selective attrition affecting the sample composition. The sample is balanced with respect to the observed characteristics across all treatments.

Table OS1. Dropout overview

Treatment	Payment Scheme	Stayed	Dropped out	Total	Dropout rate
	Piece-rate:				
1	Low_with base pay	245	97	342	28%
2	Medium_no base pay	227	85	312	27%
3	High_with entrance fee	171	125	296	42%
4	Exponential	236	99	335	30%
	Discrete bonus:				
5	Low	147	68	215	32%
6	High	118	73	191	38%
7	Low_time advantage	132	80	212	38%
8	High_time advantage	102	89	191	47%
9	Low_time handicap	138	91	229	40%
10	High_time handicap	119	75	195	38%
	Tournament:				
11	Low	135	72	208	35%
12	High	140	65	205	32%
13	Low_time advantage	132	82	214	38%
14	High_time advantage	121	86	207	42%
15	Low_time handicap	122	83	205	40%
16	High_time handicap	146	73	219	33%
Total		2,431	1,343	3,776	

Table OS2. Study 1 – Pair-wise comparisons of the dropout rate (corrected for multiple testing)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
T1															
T2	1.0														
T3	**0.04**	**0.01**													
T4	1.0	1.0	0.14												
T5	1.0	1.0	1.0	1.0											
T6	1.0	1.0	1.0	1.0	1.0										
T7	1.0	1.0	1.0	1.0	1.0	1.0									
T8	**<0.01**	**<0.01**	1.0	**0.01**	0.27	1.0	1.0								
T9	0.59	0.35	1.0	1.0	1.0	1.0	1.0	1.0							
T10	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0						
T11	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0					
T12	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.33	1.0	1.0	1.0				
T13	1.0	0.93	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0			
T14	0.24	0.10	1.0	0.61	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0		
T15	0.57	0.25	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	
T16	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.77	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Notes: the entries report the Bonferroni corrected p-value of the Fisher's exact test comparing the dropout rate in treatment T_i to that in treatment T_j where $i \neq j$. When the p-value is greater than 1.0 after correction, we record it as 1.0. P-values below 0.05 are shown in bold

Table OS3. Study 1 – Sample Balance Check

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16
Male	0.42	0.22	0.39	**0.04**	0.69	0.57	0.35	0.77	0.94	0.42	0.19	**0.03**	0.12	0.52	0.53	0.99
Education	0.96	0.16	0.65	0.92	0.84	0.32	0.28	0.55	0.39	0.55	0.09	0.34	0.57	0.90	**0.03**	0.68
Income	0.35	0.66	0.93	0.42	0.65	0.71	0.10	0.08	0.75	0.28	0.35	0.37	0.34	**0.02**	0.06	0.44
Social status	0.79	0.38	0.41	0.26	0.89	0.49	0.58	0.94	0.67	0.55	0.22	0.57	0.90	0.09	0.37	0.66
Ethnicity	0.87	0.28	0.96	0.62	0.75	0.56	0.66	0.39	0.70	**0.00**	0.11	0.29	0.88	0.31	0.11	0.16
Rural area	0.76	0.70	0.75	0.33	0.90	0.59	0.06	0.41	0.74	0.42	0.12	0.40	0.77	0.73	0.28	0.06
Employment status	0.35	0.49	0.45	0.85	0.83	0.76	0.79	0.07	0.92	0.43	0.85	0.98	0.47	0.47	0.84	0.86
All	0.94	0.48	0.95	0.25	0.99	0.63	0.10	0.21	0.80	0.06	0.11	0.30	0.79	0.41	0.21	0.47

Notes: the table reports p-values of the estimated coefficient from OLS regressions with robust standard errors. The last row reports p-values of the F-test when including all demographic variables. All p-values reported here are not corrected for multiple-testing. P-values below 0.05 are shown in bold

OS.3 STUDY 1: HISTOGRAMS OF FAIRNESS JUDGMENTS

Figure OS1: Histogram of fairness judgments of piece rate schemes (T1-T4)

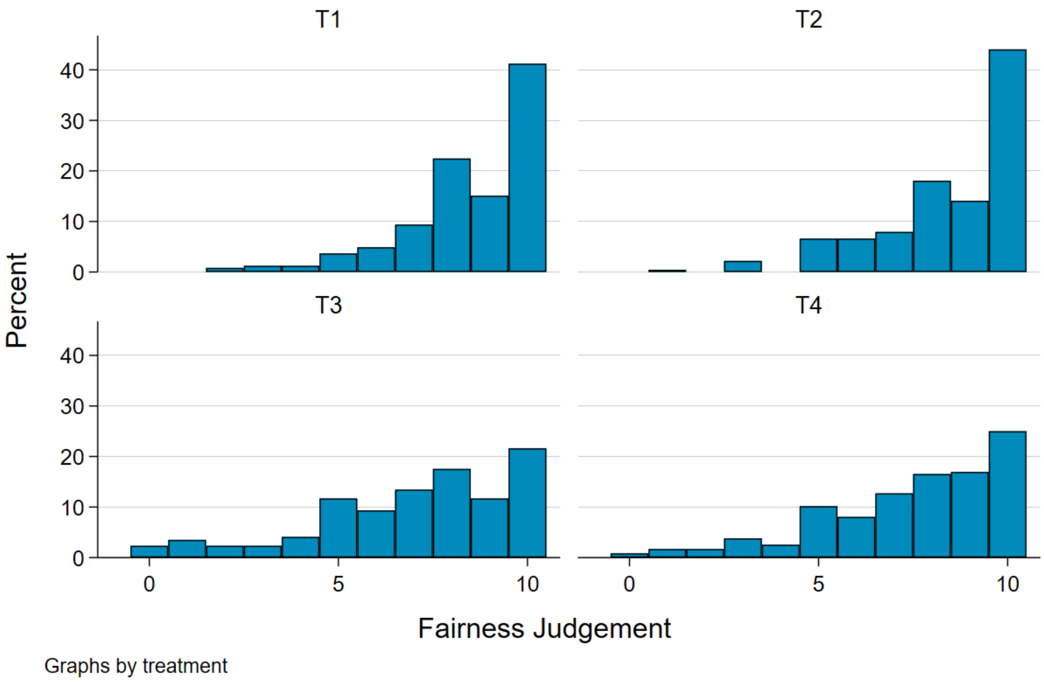


Figure OS2: Histogram of fairness judgments of discrete bonus schemes (T5-T10)

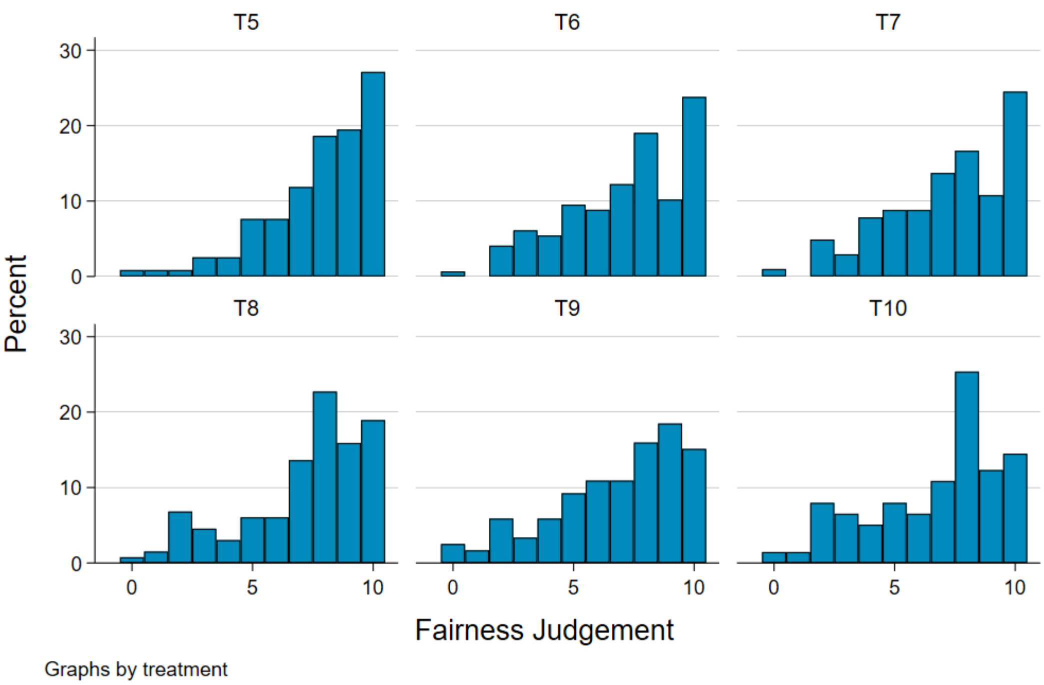
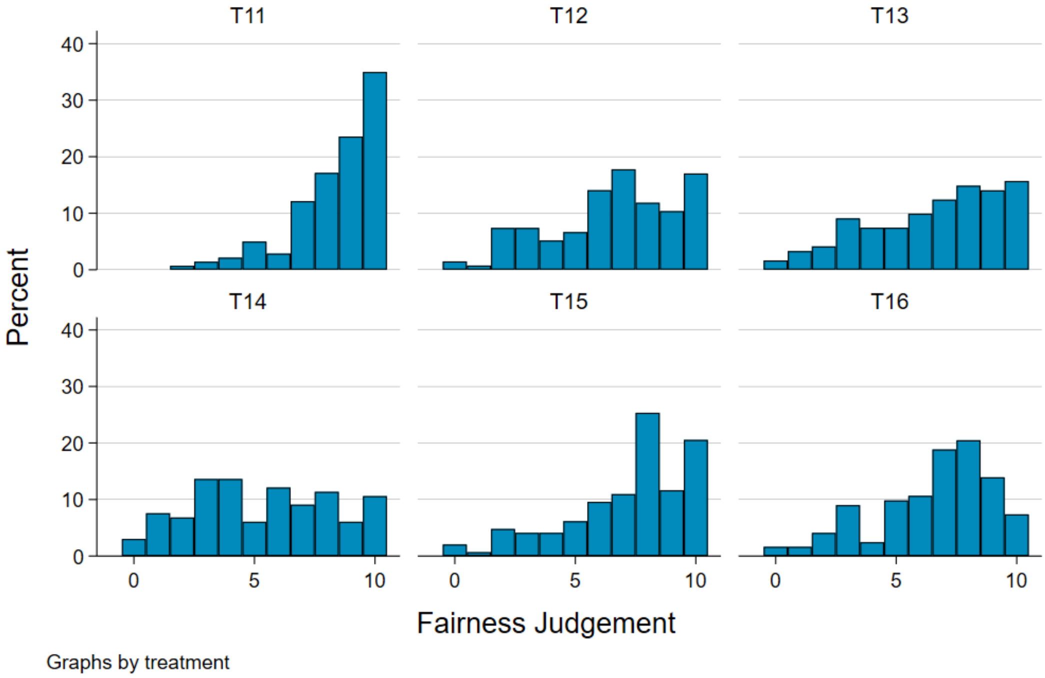


Figure OS3: Histogram of fairness judgments of tournament schemes (T11-T16)



OS.4 STUDY 2: ATTRITION AND BALANCE STATISTICS

Table OS4. Study 2. Sample Balance Check

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16
Gender	0.44	0.41	0.58	0.62	0.83	0.65	0.62	0.73	0.48	0.43	0.77	0.16	0.11	0.95	**0.01**	0.86
Education	0.08	0.65	0.35	0.99	0.98	0.72	0.55	0.73	0.66	0.94	0.53	0.42	0.95	0.60	0.61	0.37
Income	0.07	0.99	0.28	0.05	**0.02**	0.18	0.18	0.79	0.57	0.61	0.67	0.30	0.14	0.88	0.74	0.65
Social status	0.44	0.41	0.46	0.56	0.66	0.53	0.84	0.74	0.64	0.52	0.09	0.67	0.65	**0.03**	**0.02**	0.12
Ethnicity	0.72	0.72	0.21	0.27	0.88	0.35	0.42	0.91	0.80	0.74	0.89	0.18	0.44	0.96	0.42	0.30
Rural area	0.57	**0.04**	0.68	**0.03**	0.62	0.43	0.90	0.15	0.17	0.33	0.28	0.53	0.10	0.18	0.55	0.86
Employment status	**0.01**	0.76	0.96	0.96	0.67	0.19	0.75	0.13	0.91	**0.00**	0.24	0.28	0.94	0.86	0.17	0.13
All	0.15	0.51	0.67	0.21	0.45	0.46	0.73	0.64	0.86	**0.01**	0.74	0.53	0.25	0.31	**0.01**	0.54

Notes: the table reports p-values of the estimated coefficient from OLS regressions with robust standard errors. The last row reports p-values of the F-test when including all demographic variables. All p-values reported here are not corrected for multiple-testing. P-values below 0.05 are shown in bold

OS.5 STUDY 2: HISTOGRAMS OF OVERALL FAIRNESS JUDGMENTS AND HX FAIRNESS JUDGMENTS

Figure OS4: Histogram of fairness judgments of piece rate schemes (T1-T4)

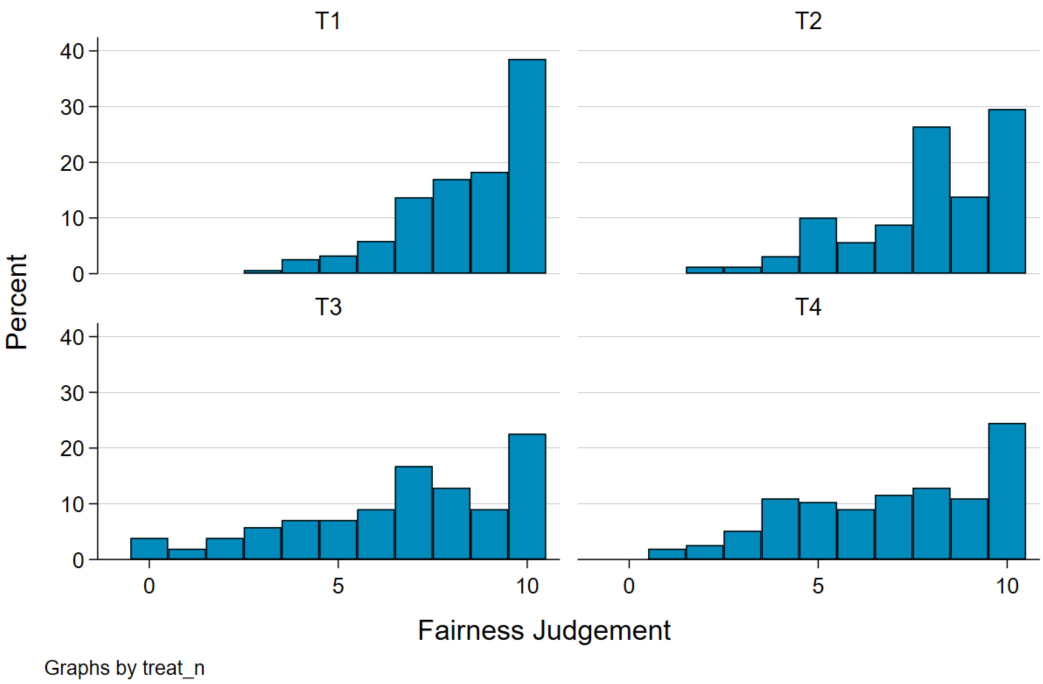


Figure OS5: Histogram of fairness judgments of discrete bonus schemes (T5-T10)

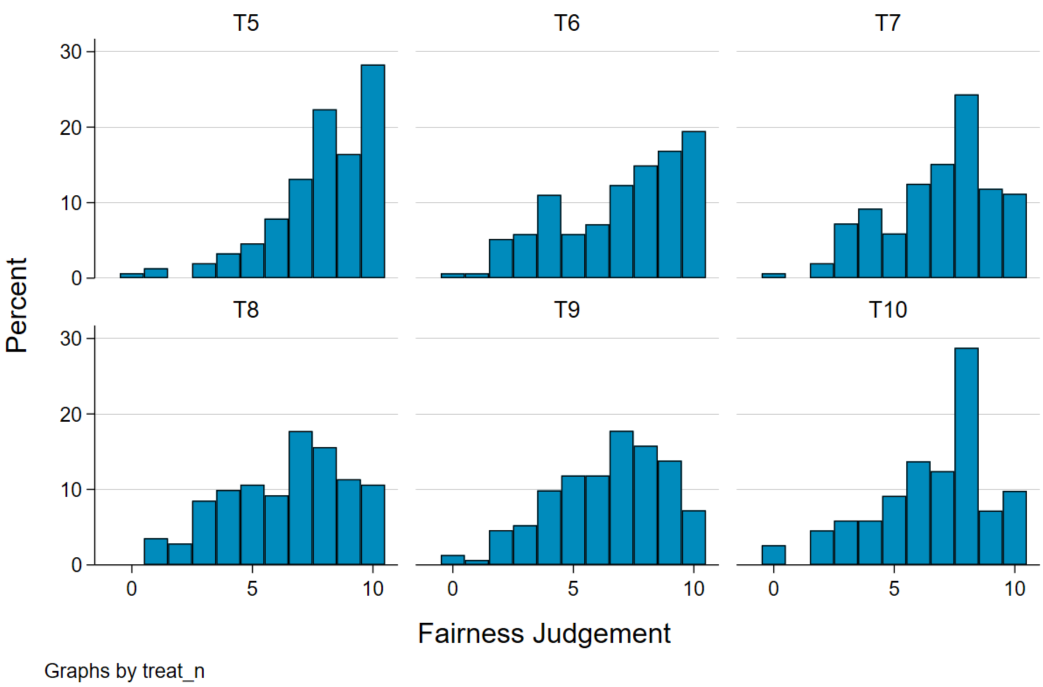


Figure OS6: Histogram of fairness judgments of tournament schemes (T11-T16)

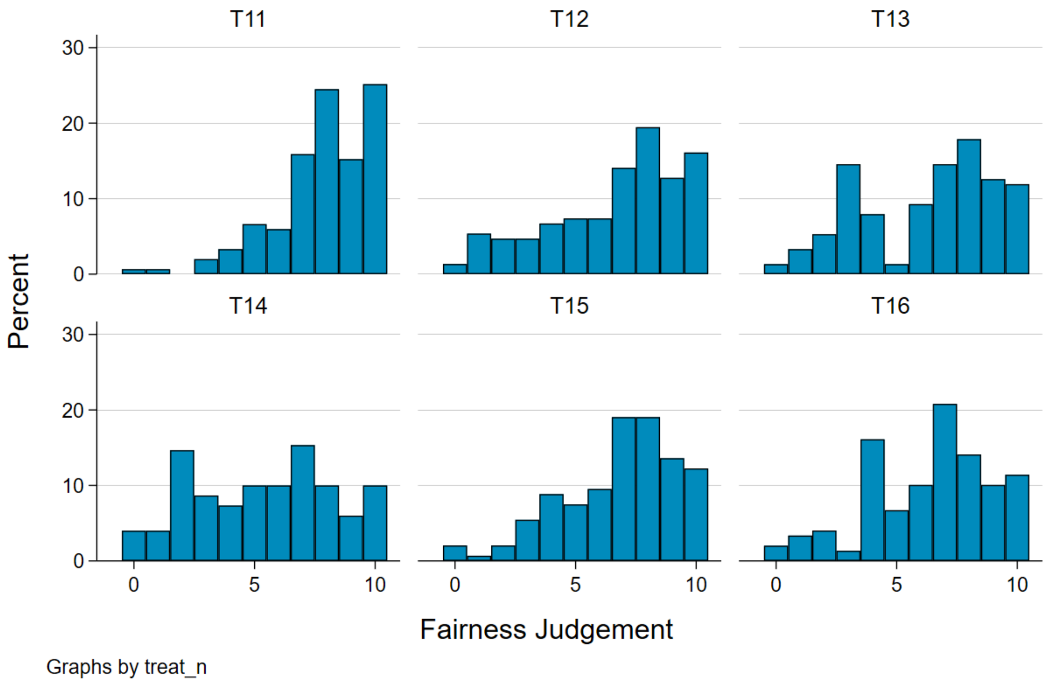


Figure OS7: Histogram of HX fairness judgments of piece rate schemes (T1-T4)

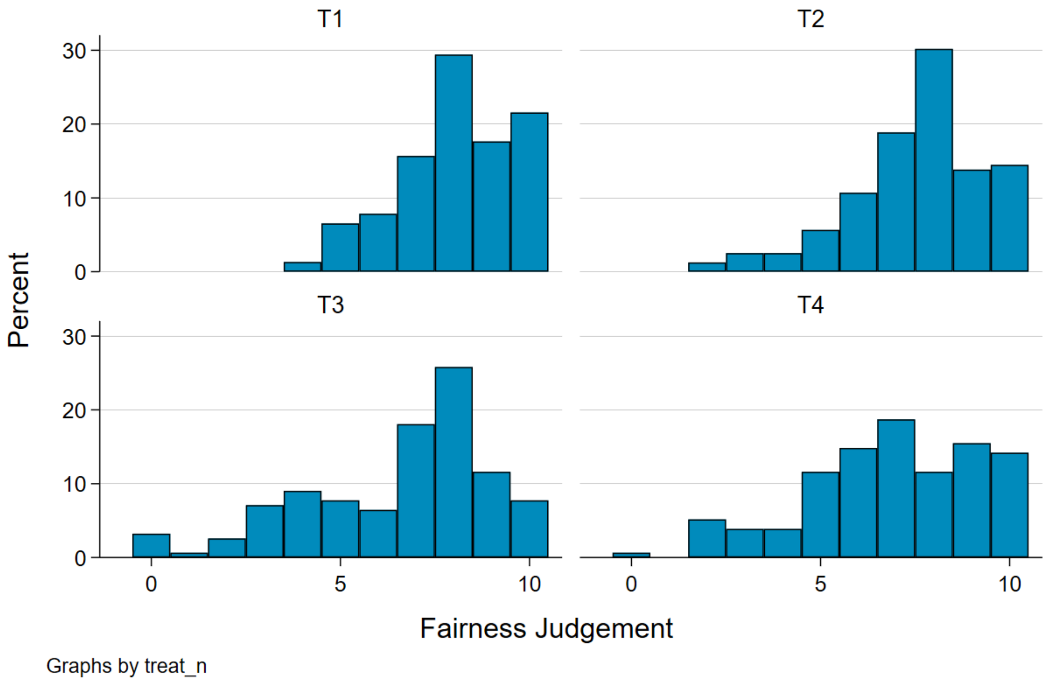


Figure OS8: Histogram of HX fairness judgments of discrete bonus schemes (T5-T10)

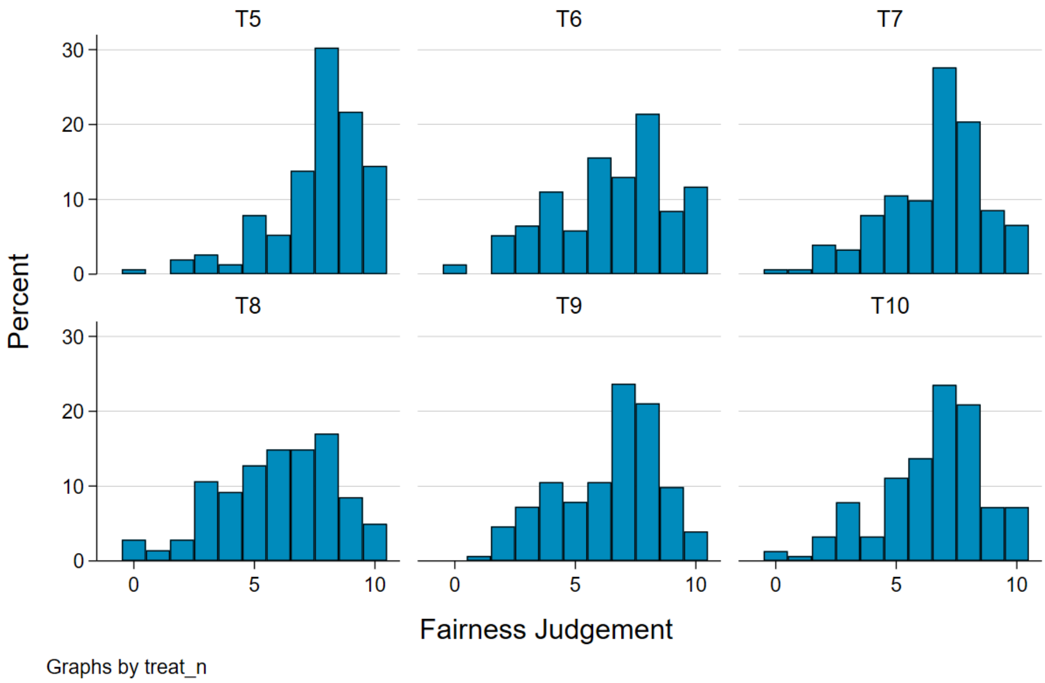
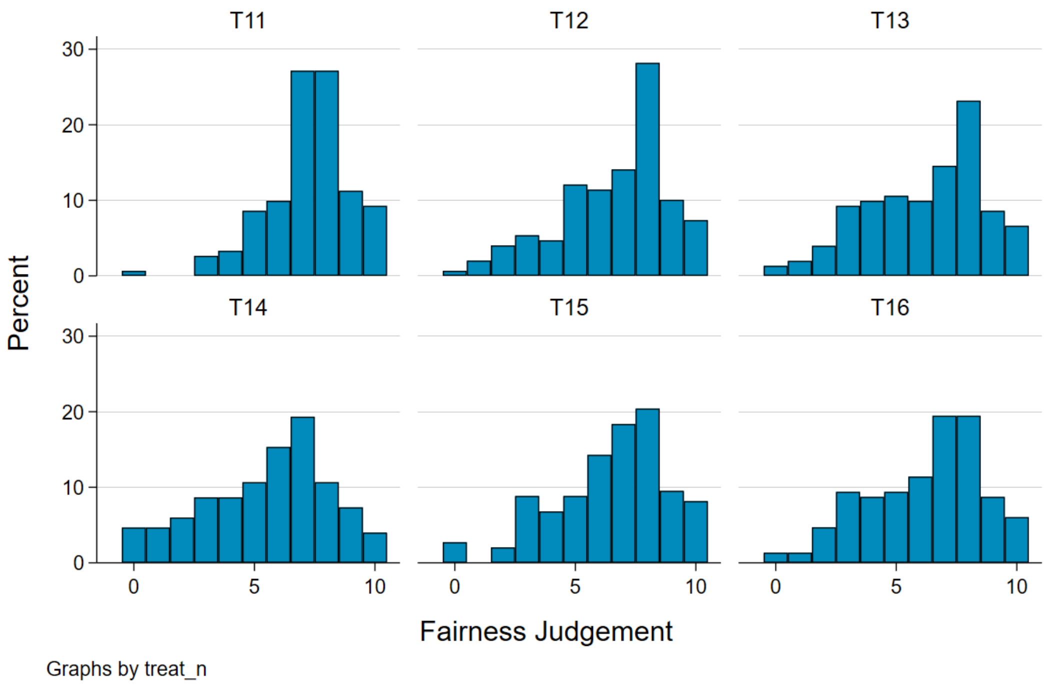


Figure OS9: Histogram of HX fairness judgments of tournament schemes (T11-T16)



OS6. STUDY 2: SUBSAMPLE ANALYSES BY TASK COMPREHENSION

Table OS5. Study 2 – Subsample analyses by task comprehension

	HX Fairness judgment			
	(1)	(2)	(3)	(4)
Gini	-1.14***	-0.44	-1.14***	-0.45
	(0.16)	(0.36)	(0.16)	(0.36)
Discrete bonus	-0.22	-0.14	-0.28*	0.03
	(0.15)	(0.41)	(0.17)	(0.45)
Tournament	-0.48***	-0.43	-0.42***	-0.70
	(0.14)	(0.43)	(0.16)	(0.51)
Time advantage	-1.00***	-0.66*	-0.79***	-0.90**
	(0.14)	(0.34)	(0.20)	(0.44)
Time handicap	-0.66***	-0.52	-0.66***	-0.77*
	(0.14)	(0.34)	(0.19)	(0.45)
Exponential	-0.07	0.12	-0.06	0.12
	(0.23)	(0.80)	(0.23)	(0.80)
Tournament x			-0.40	0.59
Time advantage			(0.28)	(0.70)
Tournament x			-0.00	0.62
Time handicap			(0.27)	(0.71)
Constant	8.40***	7.04***	8.39***	7.04***
	(0.20)	(0.60)	(0.20)	(0.61)
Task Comprehension	**Yes**	**No**	**Yes**	**No**
Controls	Yes	Yes	Yes	Yes
Observations	2,025	398	2,025	398
R-squared	0.09	0.06	0.09	0.06
F-statistics (Discrete bonus = Tournament)	5.36**	1.09	0.59	1.84
F-statistics (Time bonus = Time handicap)	5.17**	0.20	0.44	0.11

Notes: *** p<0.01, ** p<0.05, * p<0.1. OLS regression on HX fairness judgment with robust standard errors in parentheses. Sample split in subjects who passed the comprehension check (Task Comprehension "Yes") and subjects who answered some or all questions wrong in all attempts (Task Comprehension "No"). Piece rate is the reference category. Controls include gender, level of education, personal income (categorical), indicators for ethnic groups, dummy for living in a rural area, political orientation (0: Liberal; 5: Conservative), self-perceived social status, and employment status.

OS7. STUDY 2: SUMMARY OF FAIRNESS MEASURES

Table OS6. Study 2 – Summary of different fairness measures

	Payment Scheme	Fairness overall	Procedural fairness	Outcome fairness	HX fairness overall	HX procedural fairness	HX outcome fairness
	Piece-rate:						
T1:	Low_with base pay	8.46	8.56	7.01***	8.05	8.05	6.69***
T2:	Medium_no base pay	7.96	8.23	7.18***	7.55	7.72	6.69***
T3:	High_with entrance fee	6.73	7.35	6.39***	6.55	7.28	6.03***
T4:	Exponential	7.01	7.47	6.06***	6.90	7.34	6.10***
	Discrete bonus:						
T5:	Low	7.93	8.11	7.55***	7.67	7.74	7.08***
T6:	High	6.97	7.48	5.82***	6.52	6.88	5.69***
T7:	Low_time advantage	6.86	6.99	6.47**	6.61	6.78	6.30**
T8:	High_time advantage	6.40	6.94	5.79***	5.96	6.61	5.48***
T9:	Low_time handicap	6.43	6.58	6.47	6.39	6.57	6.33
T10:	High_time handicap	6.58	7.09	5.90***	6.50	6.80	5.61***
	Tournament:						
T11:	Low	7.85	7.77	6.82***	7.25	7.38	6.58***
T12:	High	6.67	7.59	5.93***	6.62	7.26	5.87***
T13:	Low_time advantage	6.29	6.66	6.42	6.21	6.40	5.97*
T14:	High_time advantage	5.40	6.01	5.57	5.49	5.97	5.59
T15:	Low_time handicap	6.82	7.00	6.86	6.47	6.49	6.44
T16:	High_time handicap	6.37	6.68	5.66***	6.20	6.58	5.66***

Notes: *** p<0.01, ** p<0.05, * p<0.1 indicate significant differences between procedural and outcome fairness judgment (two-sided t-test).

OS8. STUDY 2: COMPOSITION OF OVERALL FAIRNESS MEASURES

Table OS.7 Study 2 – Composition of overall fairness judments

	Fairness overall	**HX fairness overall**
Procedural fairness measure	0.51***	
	(0.02)	
Outcome fairness measure	0.41***	
	(0.02)	
HX procedural fairness measure		0.53***
		(0.02)
HX outcome fairness measure		0.37***
		(0.02)
Constant	0.59***	0.70***
	(0.13)	(0.12)
Observations	2,423	2,423
R-squared	0.58	0.59
F-statistics(Procedural fairness = Outcome fairness)	9.92***	20.59***

Notes: *** p<0.01,** p<0.05, * p<0.1. OLS regression on overall and HX fairness judgment with robust standard errors in parentheses.