Durham University Business School

Department of Economics & Finance

Free Riding and Workplace Democracy – Heterogeneous Task Preferences and Sorting

Kenju Kamei and Thomas Markussen

Working Paper No. 1, 2020

Durham Business School working papers series ISSN 1749-3641

Department Economics and Finance
Durham University Business School
Mill Hill Lane
Durham DH1 3LB, UK
Tel: +44 (0)191 3345200
https://www.dur.ac.uk/business/research/economics/

Free Riding and Workplace Democracy

– Heterogeneous Task Preferences and Sorting

January 2020

Kenju Kamei[1]

Thomas Markussen[2]

[1] Durham University, Email: kenju.kamei@gmail.com; kenju.kamei@durham.ac.uk.

[2] University of Copenhagen, Email: Thomas.Markussen@econ.ku.dk.

Abstract: A novel laboratory experiment is used to show that mismatching between task preferences and task assignment undermines worker productivity and leads to free riding in teams. We elicit task preferences from all workers. Workers' endogenous sorting into tasks significantly improves productivity under individual-based remuneration (performance pay). Under team-based remuneration (revenue sharing), free riding is significant, but almost exclusively among those working on undesired tasks. Task selection by majority voting in teams alleviates free riding, but only partly so, because some workers are still assigned to undesired tasks.

Keywords: free riding, team, workplace democracy, experiment, real effort

JEL codes: C92, C91, H41, D82, J01

Acknowledgment: We thank Alexander Sebald and Marco Piovesan for inputs to the experimental design, Matthew Walker, Änne Hansmeyer and Kett Panther for excellent research assistance, Chris Thornton (ITV) and Colin Theakston for providing the "You've Been Framed!" videos, and Independent Research Fund Denmark (project DFF – 4182-00163) for funding.

1. Introduction

How to overcome moral hazard in teams is one of the most important and long studied problems in the social sciences. Firms often compensate workers, contingent on the outcome of team production as individual effort is not perfectly observable (e.g., Alchian and Demsetz, 1972; Holmstrom, 1982). However, in such "team-based remuneration" schemes, selfish workers have an incentive to free ride on their peers' work (e.g., Prendergast [1999] for a survey). For the last several decades, substantial efforts have been devoted to understanding free riding in teams, using not only theoretical analyses, but also observational and experimental data. Prior research has provided rich evidence that free riding in teams is often a significant issue and proposes a number of ways to provide incentives, such as monitoring, punishment, rewards (e.g., bonuses), contracts, and tournaments. This paper contributes to the literature by investigating workers' free riding in an environment of a) multiple, heterogeneous tasks and b) workplace democracy – which we define as worker influence on task assignment.

An important insight from prior studies is that workers' self-selection of environments (e.g., employer, remuneration scheme) contributes to explaining productivity differences across firms (see, e.g., Lazear [2000] and Barro and Beaulieu [2003] for evidence in the field; and Niederle and Vesterlund [2007], Cadsby, Song and Tapon [2007], Eriksson and Villeval [2008], Dohmen and Falk [2011], Buser, Niederle and Oosterbeek [2014] for experimental evidence).[1] However, the burgeoning literature on worker self-selection has largely neglected *within*-firm sorting between different tasks, and leaves two important questions unanswered. First, how does possible mismatching between workers' task preferences and employers' task assignment affect individual performance and free riding in teams? Second, does workplace democracy increase worker performance and/or limit free riding in teams? If workers differ significantly in terms of task-specific skills and effort costs, then optimal task assignment could be an important driver of productivity. In addition, if workers care about own control over the task assignment, those who are assigned undesired tasks contrary to their preferences may not contribute their best efforts (Dow, 2003). Since skills and effort costs are at least partly unobservable to employers, self-selection into tasks could contribute to solving agency problems, especially if workers are compensated based on performance. Despite its obvious importance, the neglect of this topic in the literature is not surprising, because field data usually lack information on workers' task preferences and productivity.[2]

[1] Thus, firms' selection of remuneration schemes serves as a method of screening workers (e.g., Stiglitz, 1975).

[2] While there is a rich body of the empirical literature in personnel economics that investigates effects of workers' involvement and innovative human resource management (e.g., Ichniowski and Shaw, 1995; Ichniowski, Shaw and Prennushi, 1997; Boning *et al.*, 2002), these studies rarely use experimental methods and have not focused much on the issue of selection between tasks.

Employers typically retain a wide-ranging set of formal management rights, such as the right to assign workers to tasks. In practical terms, however, worker participation in decision making about task allocation varies considerably between different organizations. Some sales representatives are given a specific list of clients to visit while others choose clients on their own. Teachers are sometimes presented with a set curriculum and sometimes decide one themselves, either individually or as a group (e.g., a module team in higher education in the United Kingdom). A research group might be directed to investigate a certain topic, or alternatively, select such topics autonomously. In the latter case, individual researchers sometimes select topics on their own, while in other cases a group of researchers choose a common focus, for example by investing in purpose-specific research infrastructure. Mismatching between worker characteristics (task preferences and productivity) and assigned tasks could easily happen since the characteristics are often not fully observed by employers.

This paper uses a real effort laboratory experiment to analyze the effects of mismatching in task assignment on productivity and free riding, and also to investigate whether workplace democracy is an effective means against mismatching. In particular, we study an environment where recruited workers are assigned one of two different tasks (adding up two-digit numbers vs. counting zeroes in a matrix of ones and zeroes) and allow some workers to participate in deciding which task they should work on. We compare environments with, respectively, individual-based remuneration (performance pay) and team-based remuneration (revenue sharing), aiming to understand how workplace democracy could improve worker-task matching and/or reduce free riding in teams.

While free riding in teams is an established phenomenon in some experimental games, such as public goods games, its importance has not been conclusively established in real effort environments (e.g., Erev, Bornstein and Galili, 1993; Corgnet, Henan-Gonzales and Rassenti, 2015; Gächter, Huang and Sefton, 2016; van Dijk, Sonnemans and van Winden, 2001; Dohmen and Falk, 2011; Hamilton, Nickerson and Owan, 2003; Hansen, 1997; Ichniowski *et al.*, 1996; Ichniowski, Shaw and Prennushi, 1997; Kruse, 1992). A potential methodological reason for these ambiguous results is that people's responses to incentives are often quite weak in computerized real effort experiments (e.g., Araujo *et al.*, 2016). To address this concern, we design an experimental featuring (a) long duration of a single task-solving phase (30 minutes) and (b) availability of an alternative activity (watching funny videos). Aspect (b) provides an additional motive to shirk, beyond the sheer tediousness of adding and counting. Subjects in the experiment can also add material, personal gains from engaging in the video watching activity. This mimics the real-world situation where workers in a firm can gain benefits, other than psychological satisfaction, from shirking through alternative activities (e.g., maintaining networks through emails,

Facebook, etc.).

Another two aspects of our experimental design are also worth discussing: first, to our knowledge, we present the first real effort experiment in which subjects choose between different effort tasks. Second, we elicit task preferences from all subjects, regardless of whether the individual- or team-based remuneration scheme is used. This allows us to provide new experimental evidence on the effect of task mismatching on free riding in teams. We also provide evidence on how the effects of workplace democracy differ by the nature of the effort task and by the remuneration scheme (individual- or team-based).

Workplace democracy may affect productivity through three different channels. First, participatory decision making (democracy) allows workers to select into the task they do best. This is what we call "productivity sorting," and is the main focus of our paper. Second, however, the voting process may also convey *information* about the preferences of other workers. With team-based remuneration, this may affect workers' effort provision. If a worker learns that her fellow team members voted in favor of the task the group is assigned to working on, she might expect them to perform well. If the worker values reciprocity or equity, this information might in turn incentivize her to exert more effort herself, in order to match the expected effort levels of others (e.g., Tyran and Feld, 2006; Jensen and Markussen, 2019; Kamei, 2019). Alternatively, if the worker's utility function is concave with respect to income, the expectation that other members work hard (and therefore generate high income for the team) may incentivize her to lower her efforts due to diminishing marginal returns. Third, democratic task assignment could increase worker *motivation* through increased experiences of autonomy and identification with the task and with one's team (e.g., Dal Bó, Foster and Putterman, 2010; Deci and Ryan, 1985, 2000). The use of a controlled laboratory experiment allows us to distinguish between these three potential channels of causation.

Our experimental results reveal significant free riding in teams. Performance in the task-solving activities is poorer, and video-watching is more prevalent, when pay is based on team performance than when it is based on individual output. A strong interaction between task assignment and free riding in teams is observed. Free riding is much more prevalent among those working on their *least preferred task* than among those working on their most preferred task. This result contrasts with our standard theoretical model, which predicts that highly productive workers decrease effort more strongly than less productive workers when revenue sharing is used instead of individual performance pay.

The results also suggest that workplace democracy is indeed potentially a remedy against free riding in teams, because it allows more people to work on the tasks they like best, thereby reducing

workers' tendency to rely on others to get the job done. While the literature has so far focused on sorting between different firms (for example based on the incentive schemes offered by different employers), these findings underscore the importance of also devising effective task assignment mechanisms *within* firms.

Our results further show that the positive selection effect of workplace democracy is stronger in the individual- than in the team-based remuneration scheme, because workers are always able to work on the task they prefer in the former environment. Hence, productivity sorting based on individual task selection is more effective than a team-based system based on majority voting. We find that effects of workplace democracy operate almost entirely through sorting. We uncover little evidence that effects of democracy working through information and motivation are important in the present setting.

The paper is structured as follows. Section 2 presents a brief literature review and Section 3 explains the experimental design and implementation. Section 4 discusses free riding in teams and possible effects of workplace democracy using a simple theoretical model. Section 5 presents experimental results. Section 6 concludes.

2. Literature

Three branches of the literature in economics are closely related to the present study: (a) teamwork and free riding, (b) productivity sorting, and (c) effects of democracy.

There is a large theoretical literature on free riding in the provision of public goods in general (e.g., Samuelson, 1954; Bergstrom, Blume and Varian, 1986) and in organizations (e.g., Alchian and Demsetz, 1972; Holmstrom, 1982). In empirical work, free riding has been intensively studied using experimental social dilemmas, such as public goods games. For example, in public goods experiments, human subjects in a laboratory receive an endowment of experimental currency and interact with each other by choosing how much to contribute to a public good. Experimental parameters are usually set such that the social optimum is different from the Nash Equilibrium and each individual has a material incentive to free ride on others' contributions. Free riding is a stylized fact in public goods experiments, although some individuals do contribute positive amounts even in one-shot play or in the earlier periods of repeated play (e.g., Ledyard, 1995; Chaudhuri, 2011). For the last few decades, scholars have put considerable efforts into studying teamwork and free riding using real effort laboratory and field experiments. Results are much more mixed in these experiments than in experimental games. Some studies indeed find evidence of free riding (e.g., Erev, Bornstein and Galili, 1993; Corgnet, Henan-Gonzales and Rassenti, 2015; Gächter, Huang and Sefton, 2016), while others do not (e.g., van Dijk, Sonnemans and van Winden, 2001; Dohmen

and Falk, 2011; Hamilton, Nickerson and Owan, 2003; Hansen, 1997; Ichniowski *et al.*, 1996; Ichniowski, Shaw and Prennushi, 1997; Kruse, 1992).[3] Our paper addresses this unsettled issue by providing clear evidence of free riding, likely generated by the availability of an entertaining, alternative activity, as described in Section 3. As pointed out by Corgnet, Henan-Gonzalez and Rassenti (2015), one possible reason for null results in real effort experiments is that the alternative to working on a tedious task is often an equally tedious activity, namely to simply sit and wait.[4] Araujo *et al.* (2016) recently showed that output elasticity of monetary incentives is extremely low for a real effort experiment. In addition, and equally important, all the above prior studies using effort tasks studied workers' free riding behavior without considering their task preferences. This study is, to our knowledge, the first to study how workers' task preferences affect free riding in teams.

The second related area is a substantial recent literature on sorting effects of performance pay. Prior research consistently suggests that improving incentive schemes in firms not only magnifies existing workers' productivity, but also enhances productivity by attracting certain types of workers. For example, highly productive workers sort into firms with variable-pay, rather than fixed-wage, schemes (e.g., Cadsby *et al.*, 2007; Eriksson and Villeval, 2008; Dohmen and Falk, 2011; Leuven *et al.*, 2014). Overconfident workers sort into firms with tournament, rather than with piece-rate payment, schemes, thereby helping enhance efficiency by reducing variance in performance (e.g., Eriksson *et al.*, 2009). Niederle and Vesterlund (2007) and Buser, Niederle, and Oosterbeek (2014) point out that sorting is also affected by gender. These prior studies suggest that inter-firm performance differences are explained almost exclusively by sorting and difference in incentives, not by motivation effects driven by enhanced autonomy generated by the sorting process. Once incentive effects and sorting effects are controlled for, almost no performance differences are observed. Unlike all past studies, the focus of our paper is workers' sorting among tasks under *given* payment schemes. We examine sorting specific to workplace democracy: individual self-selection of a task, or collective selection of a task in teams.

The third related area is a recent experimental literature on democracy in social dilemmas. This literature has demonstrated that rules to induce pro-social behavior have a stronger effect when implemented democratically than when imposed from above, even after selection effects are controlled

[3] Team-based remuneration may even magnify productivity, compared with individual performance pay, if additional devices of control, such as monitoring, peer pressure, mutual learning between workers, formation of group norms and endogenous team formation, are in use (e.g., Román, 2009; Hamilton, Nickerson and Owan, 2003), or if problem-solving activities, such as having regular team meetings to deal with production issues, are combined (e.g., Boning *et al.*, 2007).

[4] Corgnet, Henan-Gonzales and Rassenti (2015) included an alternative activity in a real effort experiment by allowing subjects to browse the internet.

(e.g., Tyran and Feld, 2006; Dal Bó, Foster and Putterman, 2010; Sutter, Haigner and Kocher, 2010; Markussen, Putterman and Tyran, 2014; Kamei, 2016, 2019; Jensen and Markussen, 2019). This result is at odds with the second area of the literature mentioned just above. Dal Bó, Foster and Putterman (2010) interpret their results as mainly driven by factors other than selection and information, i.e., by what we refer to as "motivation" (see also Kamei [2016]). Tyran and Feld (2006) and Jensen and Markussen (2019), on the other hand, argue that information transmission (signaling) is likely to be the main factor behind the effects of democracy they document. While all the prior papers on democracy are based on experimental games, the closest papers to the present one are Melizzo, Carpenter and Matthews (2014) and Dal Bó, Foster and Kamei (2019), who considered the effects of democracy when members collectively select a remuneration scheme by voting in a real effort experiment. The choice was between revenue sharing and tournament in Melizzo *et al.*, and between fixed and piece-rate payment in Dal Bó *et al*. Aside from the difference in choices (remunerations versus tasks), the research questions of Melizzo, Carpenter and Matthews (2014) and, respectively, Dal Bó, Foster and Kamei (2019) are different from ours, since they do not focus on free riding in teams.[5] Hence, unlike in our paper, Melizzo *et al.* and Dal Bó *et al.* do not have treatments without group membership. Further, they did not elicit task preferences from all workers. In contrast, preference elicitation is an essential element of our set-up.

Experiments on the effects of democracy naturally focus on decision making in groups. To our knowledge, the present paper is the first to compare participatory decision making between individuals and groups. This is interesting because some of the mechanisms, which have been hypothesized as driving effects of democracy, function at the individual level (e.g., selection effects, motivation), while others are inherently group-level phenomena (e.g., transmission of information between voters).

3. Experimental design

The experiment revolves around two different real effort tasks. First, in the "addition" task, subjects are shown five two-digit numbers and must calculate the sum. When they submit an answer or skip the question (by clicking the "submit" button without entering an answer), a new set of numbers is randomly generated. A correct response is rewarded with two Danish kroner. This task has been used in a number of previous experiments, for example in Niederle and Vesterlund (2007), Sloof *et al.* (2010), Corgnet *et al.* (2011), Eckarz *et al.* (2012) and Lezzi *et al.* (2015). Second, in the "counting" task, subjects are presented

[5] Similar to the findings in the literature on productivity sorting, there were no effects of voting on motivations in Dal Bó *et al.* (2019). Melizzo *et al.* (2014), on the other hand, did find that democracy increases effort, but the effect was economically small.

with a 9×10 matrix of ones and zeroes and must count the zeroes. After an answer is submitted or the question is skipped (the submit button is pressed without an answer being entered), a new matrix is randomly generated (e.g., Falk *et al.*, 2006; Abeler *et al.*, 2009; Lezzi *et al.*, 2015). Each correct response is rewarded with two and a half Danish kroner.[6] Workers do not receive feedback on whether their responses are correct during the task-solving activities. They are provided information on their own performance (number of questions answered correctly) at the end of each part of the experiment. For screenshots of the task-solving interfaces, see the instructions in the appendix.

We selected the addition and counting tasks because intrinsic motivation is assumed to be low in both tasks. They are tedious and labor-intensive assignments. The goal of this paper is to study free riding in teams, and the effects of democratic task assignment. We thus need to have an environment in which extrinsic monetary incentives (individual- vs. team-based remuneration) affect behavior. For tasks with high intrinsic motivation, there might be little room for such effects to play a role. At the same time, the two tasks are clearly different. In particular, the counting task is almost entirely effort-intensive. The skills required to solve the counting task are possessed by any university student. Performance is therefore mainly driven by effort. The addition task, on the other hand, requires skills as well as effort. Mathematical abilities are not evenly distributed and workers with such abilities have a clear advantage in the addition task. This implies that preferences for tasks are likely to differ across workers.

Subjects receive instructions for the entire experiment upon entering the laboratory. When all subjects have read the instructions and the experimenter has answered all questions in public, the experiment starts. The experiment has two parts. In Part 1, all subjects work on both the addition and counting tasks for three minutes of each task. The sequence of tasks is set randomly for each subject. Remuneration is based on individual performance. At the end of each three-minute period, subjects receive feedback on the number of correct responses and their own earnings. The decision to provide performance feedback at this stage was taken in order to render the ensuing task choice as meaningful and salient as possible for the subjects. Information about relative performance in the two tasks is likely to generate induced preferences for tasks, and thus form a basis for task selection.

Part 2 consists of a voting phase and the main task-solving phase. Once Part 1 is over, all subjects vote on which task they would prefer to use in the main task-solving phase. We design three treatments (Table 1). The treatments differ in terms of whether remuneration in the main task-solving phase is based

[6] Based on a few pilot experiments, payment and task parameters were set such that expected average earnings from the two tasks are about even.

on individual or team performance. The first treatment, called the "Individual" treatment, uses individual remuneration. Subjects in this treatment are paid on the basis of their own performance, as in Part 1. In the other two treatments, called "Team (Full feedback)" and "Team (Limited feedback)," subjects are randomly assigned to teams of three. Their remuneration is based on the aggregate performance in the team. In particular, each team member receives a third of the total earnings generated by the task-solving activities in the team. Therefore, an incentive for free riding is present.

In each treatment, we randomly generate the "democracy" condition (in which the voting outcome determines task assignment) and the "non-democracy" condition (in which the voting outcome is ignored and the computer randomly determines task assignment). This is the "vote override" procedure first used by Dal Bó, Foster and Putterman (2010). Our paper is the first to apply this method to a real effort environment. An important benefit of this methodology is that we learn the task preferences of all workers, including those in the non-democratic condition. This allows us to not only study how free riding in teams differ by task preference, but also to investigate the effects of workplace democracy while distinguishing three effects of workplace democracy: democratic selection effects, information effects, effects of motivation.

The voting outcome under the democracy condition is simply the individual's vote in the Individual treatment (hence, it is more accurately described as a "choice" rather than a "vote", but for consistency, we will label the choice as a "vote". We will also say that the condition where the individual choice decides task assignment is "democratic"). By contrast, the voting outcome in the two Team treatments is decided by the majority rule.[7] Hence, the task preferred by a majority of team members (at least two votes) is implemented in the democracy condition.

The Team treatments differ not only in terms of democracy vs. non-democracy, but also by the feedback provided to teams in the non-democracy condition. In the Team (Full feedback) treatment, team members receive information about the voting outcome (the majority decision, not the precise distribution of votes) even in the non-democratic condition. This means that workers in both the democratic and non-democratic teams have the same information. This setup allows us to isolate the effect of motivation in teams, namely an effect of democracy *beyond* the effects of selection and information. In the Team (Limited information) treatment, teams in the non-democracy condition do not learn the voting outcome. Using the Team (Limited information) treatment, therefore, we measure an

[7] We acknowledge that the size of the effect of democratic task assignment in a team may depend on voting rules. For instance, it could be weak if unanimity rule, rather than majority rule, is used, because voting may not provide a reliable signal of players' future action choices under the unanimity rule (e.g., Kamei, 2019).

effect of voting potentially driven by the information. As shown in Table 1, we have a total of six regimes. We label the six regimes "I-Dem" (Individual Remuneration with Democracy), "I-No" (Individual Remuneration with No Democracy), "T-Dem (Full)" (Team Remuneration with Democracy (Full feedback)), "T-No (Full)" (Team Remuneration with No Democracy (Full feedback)), "T-Dem (Limited)" (Team Remuneration with Democracy (Limited feedback)), and "T-No (Limited)" (Team Remuneration with No Democracy (Limited feedback)).

The main task-solving phase lasts for 30 minutes. During this period, subjects work on the same task without any performance feedback. They are reminded on the screen about the mode of task assignment (e.g., "your group voted by majority to solve counting zeroes problems" or "the computer decided that your group will solve addition problems"), since it is known that some subjects do not remember correctly the vote outcome in this kind of voting experiment.

Table 1. Summary of Treatments

		Remuneration		
		Individual	*Team (Full feedback)*	*Team (Limited feedback)*
Task assignment	*Democracy*	I-Dem (94)	T-Dem (Full) (117)	T-Dem (Limited) (60)
	No democracy	I-No (90)	T-No (Full) (105)	T-No (Limited) (51)

Note: Numbers in parentheses indicate the numbers of subjects. We set the number of observations for the Team (Full feedback) treatment around double than for the Team (Limited feedback) treatment because we expected that a large number of observations in the Full feedback treatment was necessary in order to detect potential motivation effects.

Alternative activity – watching funny videos:

At any time during the main phase, workers can watch funny videos by switching from the task-solving screen to the video viewing screen (they cannot work while watching the video). The videos are short sketches from the British "You've Been Framed!" show (produced by ITV Studios). The full video available is more than 30 minutes long, meaning that subjects could potentially watch the video for the entire 30-minute period without seeing the same sketches twice. Headphones are provided such that subjects can listen to the soundtrack. Subjects can return to solving tasks when they wish and they can watch the video as many times as they like and for any duration they prefer during the 30-minute phase. Subjects are paid one Danish krone per minute for watching the video. These earnings are *not* shared with the team, even in the treatments with team-based remuneration. Hence, video watching is clearly interpretable as free riding in the Team treatments. The video used is available upon request from the authors.

Subjects were instructed to stay silent and loud laughing at the funny videos was not observed. Nevertheless, we acknowledge that independence of observations could have been compromised had even a small ripple of laughter occurred.[8] To address this concern, we cluster standard errors at the session level when analyzing individual-level data. The benefit of using videos instead of, say, providing access to the internet, is that the content is controlled. On the internet, subjects could potentially have searched for answers to the addition tasks, or for information about problem-solving techniques.

As touched upon earlier, the alternative activity was included to address concerns about potentially inelastic responses to incentives in a real effort experiment. The presence of alternative activities increases the incentive to reduce task-solving effort. If the only alternative to, say, counting zeroes is to stare at an empty computer screen, even those prone to shirking might prefer to work on the task. With the videos available, on the other hand, an entertaining alternative is present, as it would be in a workplace where one could, for example, check Facebook or the sports news instead of working.

Implementation:

The experiment was conducted in the Laboratory for Experimental Economics at University of Copenhagen between April and October 2018, using student subjects from a broad range of fields (e.g., humanities, social science, natural science). The client computers were separated from each other by three tall partitions. The experiment, except the instructions, was programmed using the z-Tree software (Fischbacher, 2007). The total number of subjects was 517. The subjects were recruited using solicitation emails sent through the ORSEE (the Online Recruitment System for Economic Experiments), developed by Greiner (2015). 57 percent of subjects were women. The instructions were neutrally framed. The experiment (including payment) lasted about 80 minutes and subjects on average earned 154 Danish kroner (about 23 US dollars).

4. Theoretical Predictions and Discussions

We next discuss the problem of free riding in teams, a possible effect of democratic task assignment, and the interaction between democracy and remuneration structure, with the use of a simple model. Assume that a worker *i* has a quadratic cost function. We can then express *i*'s payoff as follows:

$$\pi_{ik}(e_i) = c_{ik}e_i + \zeta z_i - (h_{ik} - \delta_i D)e_i^2 \text{ under the individual-based remuneration.} \quad (1)$$

$$\pi_{ik}(e_i|e_{-i}) = \frac{1}{n}\sum_{j=1}^{n} c_{jk}e_j + \zeta z_i - (h_{ik} - \delta_i D)e_i^2 \text{ under the team-based remuneration.} \quad (2)$$

[8] While there was a session with only three subjects who at least once watched the video, there was in fact another one with 15 such subjects.

Here e_i is effort, i.e., the time spent working on effort task [minutes], z_i is the time spent watching the video [minutes], c_{ik} is *i*'s return per minute of exerting effort for task *k*, ζ is the return per minute of watching the video (= 1 in the experiment), *n* is the number of team members (= 3 in the experiment), h_{ik} is a positive task- and individual-specific cost of effort parameter, δ_i is a parameter measuring a possible increase in motivation due to democratic task assignment ($\delta_i < h_{ik}$), and *D* is a binary indicator for democracy being in place ($D \geq 0$). Worker *i* faces the time constraint as in Condition (3):

$$e_i + z_i = 30,\ e_i \geq 0, \text{ and } z_i \geq 0. \tag{3}$$

Under the team-based remuneration scheme, not only does *i*'s payoff depend on her peers' production, but *i* is also aware of the peers' payoffs at the end of the experiment (Equation (2)). In the standard self-interest model, *i*'s optimal effort provision under the team-based remuneration (e_i^{**}) is predicted to be lower than that under the individual-based remuneration (e_i^{*}). To see this, we can solve e_i^{*} and e_i^{**} by maximizing Equations (1) and (2), respectively, subject to time constraint (3):

$$e_i^{*} = 0 \text{ for } c_{ik} < 1;\ e_i^{*} = \frac{c_{ik}-1}{2(h_{ik}-\delta_i D)} \text{ for } c_{ik} \in [1, 60(h_{ik} - \delta_i D) + 1];$$
$$e_i^{*} = 30 \text{ for } c_{ik} \geq 60(h_{ik} - \delta_i D) + 1. \tag{4}$$

$$e_i^{**} = 0 \text{ for } c_{ik} < 3;\ e_i^{**} = \frac{\frac{c_{ik}}{3}-1}{2(h_{ik}-\delta_i D)} \text{ for } c_{ik} \in [3, 180(h_{ik} - \delta_i D) + 3];$$
$$e_i^{**} = 30 \text{ for } c_{ik} \geq 180(h_{ik} - \delta_i D) + 3. \tag{5}$$

These calculations show that both e_i^{*} and e_i^{**} are increasing in individual productivity (c_{ik}) but are decreasing in the unit cost of effort (h_{ik}). We also see that the optimal efforts increase with the introduction of democracy, if $D > 0$, under both the individual- and team-based remunerations. These two conditions suggest that e_i^{**} is smaller than e_i^{*}, at least for intermediate values of c_{ik}. In other words, free riding among workers is predicted in the Team treatments.

Conditions (4) and (5) also suggest that relative material payoffs among team workers differ by the remuneration scheme. Under the individual-based remuneration, the higher ability (i.e., higher c_{ik} and lower h_{ik}) a worker *i* has, the higher payoff she can obtain by exerting a higher level of effort. However, under the team-based remuneration, as each worker's production ($c_{jk}e_j$) is equally shared in their team, the higher ability, relative to the other team workers, *i* has, the *lower* payoff she will earn.[9]

[9] For simplicity, suppose that (a) *i* has not only the highest *c*, but also the lowest *h*, in her team, and (b) δ_i is the same among the workers. Then, $\pi_{ik} - \pi_{jk} = e_j - e_i - (h_{ik} - \delta D)e_i^2 + (h_{jk} - \delta D)e_j^2 \leq$. Here, clearly $e_j - e_i \leq 0$. The rest of terms, $-(h_{ik} - \delta D)e_i^2 + (h_{jk} - \delta D)e_j^2$, is also negative because $-(h_{ik} - \delta D)e_i^2 + \left(h_{jk} - \delta D\right)e_j^2 = -\frac{\left(\frac{c_{ik}}{3}-1\right)^2}{4(h_{ik}-\delta D)} + \frac{\left(\frac{c_{jk}}{3}-1\right)^2}{4(h_{jk}-\delta D)} \leq -\frac{\left(\frac{c_{ik}}{3}-1\right)^2}{4(h_{jk}-\delta D)} + \frac{\left(\frac{c_{jk}}{3}-1\right)^2}{4(h_{jk}-\delta D)} \leq 0$ (the first [second] inequality holds as $h_{ik} \leq h_{jk}$ [$c_{ik} \geq c_{jk}$]).

Further, these two conditions suggest that the degree of free riding differs by worker's type because $e_i^* - e_i^{**} = \frac{c_{ik}}{3(h_{ik}-\delta_i D)}$ when the solutions are interior. This implies that workers' free riding is not only increasing in c_{ik}, but is also decreasing in h_{ik}. We can thus predict that highly productive workers free ride more strongly than less productive workers.

As member *i*'s payoff depends on her team members' effort provision e_{-i} (Equation (2)), a wider range of effort level can hold as an equilibrium outcome under the team-based remuneration scheme if we assume that workers have other-regarding preferences. For an illustrative purpose, let us assume that all workers have inequity-averse preferences (Fehr and Schmidt, 1999). We can then express *i*'s utility, $U_{ik}(e_i|e_{-i}^b)$, as follows:

$$U_{ik}(e_i|e_{-i}^b) = \pi_{ik}(e_i|e_{-i}^b) - \frac{1}{n-1}\sum_{j\neq i}\left[\alpha_i \cdot max\left\{\pi_{jk}\left(e_j^b\middle|e_i, e_{-i}^b\right) - \pi_{ik}\left(e_i\middle|e_{-i}^b\right), 0\right\} + \beta_i \cdot max\left\{\pi_{ik}\left(e_i\middle|e_{-i}^b\right) - \pi_{jk}\left(e_j^b\middle|e_i, e_{-i}^b\right), 0\right\}\right], \quad (6)$$

where $\alpha_i \geq \beta_i$ and $1 > \beta_i \geq 0$, and n = 3. e_{-i}^b is *i*'s belief regarding her two team members' effort provision. For simplicity, let us also assume that the belief formed by *i* is correct. In order to derive a benchmark of the team behavior, for simplicity let us further assume that individual productivity, cost functions and inequity-averse preferences are all homogeneous ($c_{ik} = c_k$, $\delta_i = \delta$, $h_{ik} = h_k$, $\alpha_i = \alpha$, $\beta_i = \beta$) and that interior solutions exist ($c_k > 3$). We prove in Appendix A that in this setup, (a) the three members in each team select the same effort level ($e_1 = e_2 = e_3 = e^{***}$) at optimum and (b) the team members face a coordination problem in that not only a greater but also a smaller per-worker effort provision is possible, compared with the effort level chosen by an individual without equity concerns, i.e., Equation (5). The set of equilibrium per-worker effort levels is given by Condition (7):

$$e^{***} \in \left[\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}, \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}\right]. \quad (7)$$

Condition (7) indicates that if β is large enough that $\beta > 2/3$, it could even be possible for a team to have a greater per-worker effort provision, compared with the effort provision e^* in the Individual treatment (Equation (4)).[10,11] We note, however, that such a high team performance is unlikely to occur since mean β in the population is likely to be less than 0.6. For example, Fehr and Schmidt (1999, 2010) estimate that

[10] When $\beta = 2/3$, the upper limit of Condition (7) is the same as the optimal effort provision under the individual-based remuneration scheme: i.e., $\frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)} = \frac{c_k-1}{2(h_k-\delta D)}$.

[11] Notice that if $\alpha = \beta = 0$ (the assumption of self-interest), e^{***} in Condition (7) becomes equal to $\frac{\frac{c_k}{3}-1}{2(h_k-\delta D)}$, the same as the interior solution case in Equation (5).

about 30%, 30% and 40% of individuals have $\beta = 0$, $\beta = 0.25$, and $\beta = 0.6$, respectively.

We remark that if we relax the assumption of homogenous productivity (e.g., c_{ik} differs by worker), the model predicts that each worker can exert a different level of effort unlike in the symmetric case. However, free riding in teams is expected. The effort provision by member *i* in equilibrium depends on three members' technologies and inequity-averse preferences ($\{c_{ik}, \delta_i, h_{ik}, \alpha_i, \beta_i\}_{i=1,2,3}$). *i* exerts the largest effort among all possible scenarios if her payoff given by Equation (2) is higher than that of any other team member. $\frac{\partial U_{ik}}{\partial e_i}$ is then calculated as in Equation (A1) of the online Appendix. This suggests that the highest possible effort exerted by *i* (e_i^{****}) is $\frac{\frac{c_{ik}}{3}-(1-\beta_i)}{2(1-\beta_i)(h_{ik}-\delta_i D)}$. Note, however, that $\frac{\frac{c_{ik}}{3}-(1-\beta_i)}{2(1-\beta_i)(h_{ik}-\delta_i D)} < \frac{c_{ik}-1}{2(h_{ik}-\delta_i D)}$ considering the distribution of β_i estimated in Fehr and Schmidt (1999, 2010).

We discussed earlier that highly productive workers would free ride more strongly than less productive workers. It should be noted here that introducing inequity-averse preferences would not change this prediction since the highly productive workers earn smaller π_{ik} without considering inequity concerns (see also footnote 9). This difference in the material payoff encourages them to shrink the income inequality by decreasing effort further. In contrast, the less productive workers would increase effort provision so as to shrink income inequality if they are concerned about the inequality with the highly productive workers.

In summary, even if we consider other-regarding preferences in workers, we expect effort levels to be lower in the team- than in the individual-based remuneration scheme.

Hypothesis 1 *(free riding)*: *(a) Workers exert less effort in teams than under the individual-based remuneration scheme. (b) Free riding is more serious for highly productive than for less productive workers.*

The model and our calculations above predict possible effects of democratic task assignment. First, and most importantly, a democratic selection effect is expected. As workers' effort provision depends on productivity (c_{ik}) and cost functions (h_{ik}, δ_i), individuals would choose/vote for the task with the highest advantage. Second, we may have two other forms for effects of democratic task assignment. The $\delta_i D$ term captures a possible motivation effect, i.e., that the unit cost of effort provision is arguably lower if a task is implemented democratically than if it is given exogenously (e.g., Deci and Ryan, 1985, 2000; Dal Bó, Foster and Putterman, 2010; Kamei, 2016). For example, workers may experience higher degrees of autonomy (and also build group identity under the two Team treatments) if they are allowed to participate in task assignment ($D > 0$), which could in turn motivate them to exert more effort. This effect could be present under both the individual- and team-based remuneration schemes. In addition to

the motivation effect, an information effect may also be present under the team-based remuneration. We described that the team members face a coordination problem if the members are concerned about within-team income inequality. Past experiments have consistently found positive effects of communication and cheap talk in improving cooperation in the context of both prisoner's dilemma games (e.g., Duffy and Feltovich, 2002 and 2006) and coordination games (e.g., DeJong, Forsythe, Ross, 1992; Charness, 2000; Blume and Ortmann, 2007). Voting could also have such a signaling value because team members learn that the peers can work on their preferred task and thus they may expect the peers to put in more effort (e.g., Tyran and Feld, 2006; Jensen and Markussen, 2019; Kamei, 2019).

Hypothesis 2: *(a) Democratic task assignment leads to workers' sorting based on expected performance between the two tasks. (b) Democratic task assignment also improves workers' effort provision through a motivation effect and/or a signaling effect.*

We can also hypothesize that democratic selection effects are stronger under the individual- than under the team-based remuneration scheme. Under the latter scheme, some team members may not get the task they desire even with democratic assignment if a majority of team members prefer a different task. Under the former scheme, on the other hand, subjects are always able to work on the task they prefer with self-selection.

Hypothesis 3: *Democratic selection effects are stronger under the individual- than under the team-based remuneration scheme.*

However, we cannot provide definite predictions regarding whether the total effect of democratic task assignment (i.e. the aggregate effect of selection, signaling and motivation) is stronger under the individual- than under the team-based remuneration. As discussed, the information effect may be present in the team-based remuneration scheme.[12] However, the motivation effect, if $D > 0$, may be either smaller or larger under the team- than under the individual-based remuneration scheme. On the one hand, people may care about self-determination, i.e. that they control their own environment (e.g., Deci and Ryan, 1985, 2000). The self-determination effect could be stronger under the individual- than under the team-based remuneration scheme, since subjects in the democratic condition have full control over task assignment under the former scheme, whereas members of teams may potentially be outvoted by others under the latter scheme. However, on the other hand, the act of voting potentially increases group

[12] This effect is possible only in the Team (Limited feedback) treatment, however. In the Team (Full feedback) treatment, the same information effect is expected even in the non-democratic environment as subjects are informed of the distribution of votes.

identity, thereby enhancing motivation to work for the team. This type of mechanism is only active in teams, not in isolated individuals.

5. Results

We first go over the subjects' behavior in Part 1 and their voting decisions at the beginning of Part 2 (Section 5.1). We then study the treatment differences in Part 2 task-solving performance and video watching in Sections 5.2 and 5.3.

5.1. Part 1 performance and voting

Table 2 shows subjects' average performance in Part 1 and their voting decisions at the onset of Part 2. The table confirms, first, that subjects' average earnings in the initial three-minute phase were similar for the two tasks, as intended. Subjects earned on average 10.7 Danish kroner in the addition task and 12.0 kroner in the counting task. Second, voting varies little across treatments, except for the I-No regime, where the vote share for the counting task was somewhat higher than average. None of the differences between this and other treatment regimes are statistically significant, however ($p > .100$, two-sided Fisher's exact tests). Overall, 55 percent of subjects voted for working on the counting task and 45 percent for the addition task. The almost balanced voting distribution is a methodological advantage because it provides us with a reasonably large number of subjects working on both tasks, and with people who work on the task they desire, and others who do not.

Table 2. Part 1 Earnings and Voting

	Individual		Team (Full feedback)		Team (Limited feedback)		All subjects
	No democ.	Democracy	No democ.	Democracy	No democ.	Democracy	
	I-No	I-Dem	T-No (Full)	T-Dem (Full)	T-No (Limited)	T-Dem (Limited)	
Earnings, addition task (DKK)	10.1	10.4	11.4	10.3	11.5	10.7	10.7
Earnings, counting task (DKK)	12.4	12.5	12.3	11.2	11.8	11.9	12.0
Voted for addition task (percent)	38.9	47.9	46.7	45.3	45.1	43.3	44.7
Voted for counting task (percent)	61.1	52.1	53.3	54.7	54.9	56.7	55.3

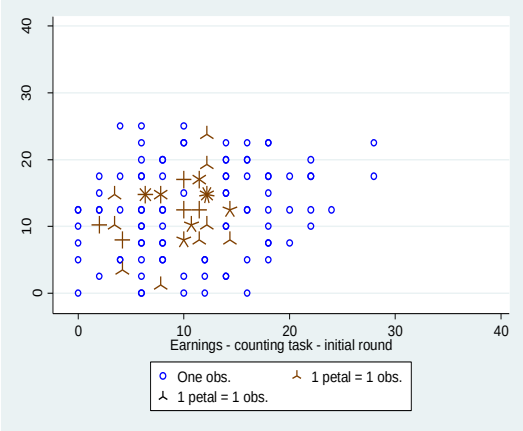
Figure 1.A shows the distribution of subjects' Part 1 earnings in the two tasks by treatment. The figure shows a positive but far from perfect correlation in each treatment. A large number of workers earned significantly more in one task than in the other. The large degree of heterogeneity is a crucial feature, which we intended to generate by carefully selecting two clearly different effort tasks (see discussion in Section 3). This provides workers with incentives to select a task based on their relative performance in the two tasks in Part 1. It also means that voting can serve as a signal, and that the vote

outcome potentially conveys important information about the abilities of fellow team members in the Team treatments.

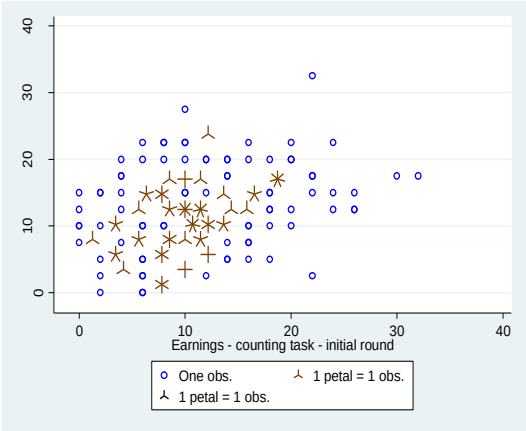
In order to check whether subjects do sort into tasks according to Part 1 performance, we next examine how subjects' Part 1 relative performance affected their voting. As shown in Figure 1.B, our data support Hypothesis 2(a). The subjects' relative performance in the two tasks in Part 1 was a clear predictor of voting in all three treatments. Appendix Figure C.1 reports the cumulative distribution of relative performance (i.e., the ratios of score in the addition task to score in the counting task) by voting in each treatment. It shows that the distribution for supporters of the addition task first-order statistically dominates the distribution for those who voted for the counting task in every treatment. Hence, we conclude that voting leads to productivity sorting within organizations. This is consistent with results from other studies, showing that workers sort into remuneration schemes according to productivity (e.g., Cadsby *et al.*, 2007; Eriksson and Villeval, 2008; Dohmen and Falk, 2011).

Result 1: *Subjects selected tasks based on their relative performance in Part 1, consistent with Hypothesis 2(a).*

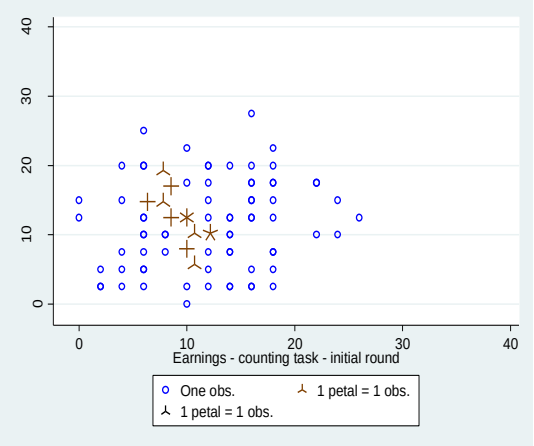
Figure 1. Distribution of Part 1 Performances, and Voting Decisions



(i) Individual treatment

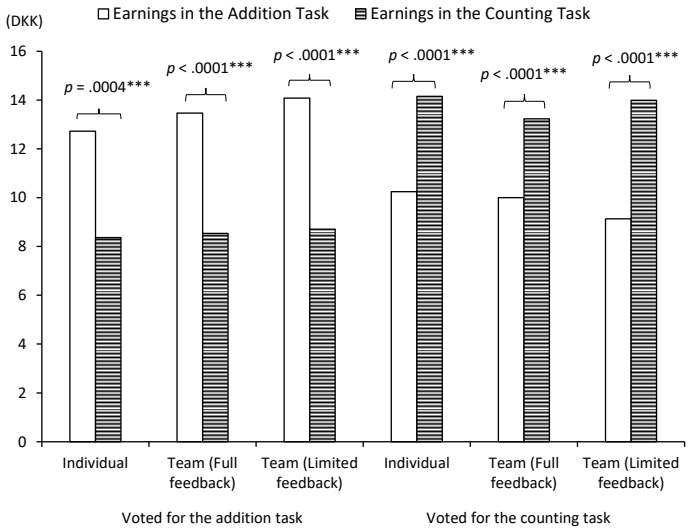


(ii) Team (Full feedback) treatment



(iii) Team (Limited feedback) treatment

(A) Distributions of subjects' relative performance between the two tasks in Part 1



(B) Subjects' voting based on relative performance between the two tasks in Part 1

Notes: *p*-values (two-sided) in panel B were calculated using two-sided Wilcoxon signed ranks tests.

5.2. Free riding in teams and workplace democracy

Let us now turn our attention to the subjects' behavior in Part 2. Pooled results from both of the two effort tasks can be used to test the performance differences across the six regimes in the experiment (Figure 2). We use subjects' earnings from effort provision as an indicator of performance since per unit returns are different for the two tasks (two Danish kroner for the addition task versus 2.5 kroner for the counting task).[13] Four interesting patterns emerge. First, the figure shows clear effects of the mode of

[13] While both a Skewness Kurtosis test and a Shapiro-Wilk's W test reject the normality of subjects' earnings in the Team (Full feedback) treatment at the 1% level, they do not reject it in the other two treatments. The failure of normality is mainly due to the presence of some high performers in the addition task (see the cumulative

remuneration. Subjects' average performances were always higher with the individual- than with team-based remuneration scheme (panel A). The differences are significant at least at the 5% level, except for the comparison between the I-No and T-No (Full) regimes. Second, parallel to the first finding, time spent watching the video was always much lower with the individual- than with the team-based remuneration scheme. The differences are often significant despite the high variation in the time data. These clearly suggest that consistent with Hypothesis 1(a), free riding in teams is a significant issue.

Result 2: *Subjects' average performance was lower with the team-based than with the individual-based remuneration scheme, consistent with Hypothesis 1(a). The subjects spent much more time watching the video with the former than with the latter scheme.*

Third, comparing democratic and non-democratic environments within each mode of remuneration, a positive effect of democracy is visible in the Individual treatment (panel A). Subjects earned on average about 10 Danish kroner more in the democratic condition than in the non-democratic one. The difference is significant at the 5% level. Parallel to this behavioral difference, we also found that subjects watched the video less in the democratic condition than in the non-democratic one. However, the difference is small and not statistically significant (panel B). This suggests that the positive effect on scores is mainly due to a democratic selection effect, whereby subjects sort according to skills, rather than differences in effort provision. We will perform detailed analyses behind the positive effect by the effort task and also by workers' task preference later.

Fourth, and by clear contrast, democratic task assignment has a much smaller effect in the two Team treatments. It has some positive impact in the Team (Limited feedback) treatment as shown in Figure 2. The difference is far from significant, however.

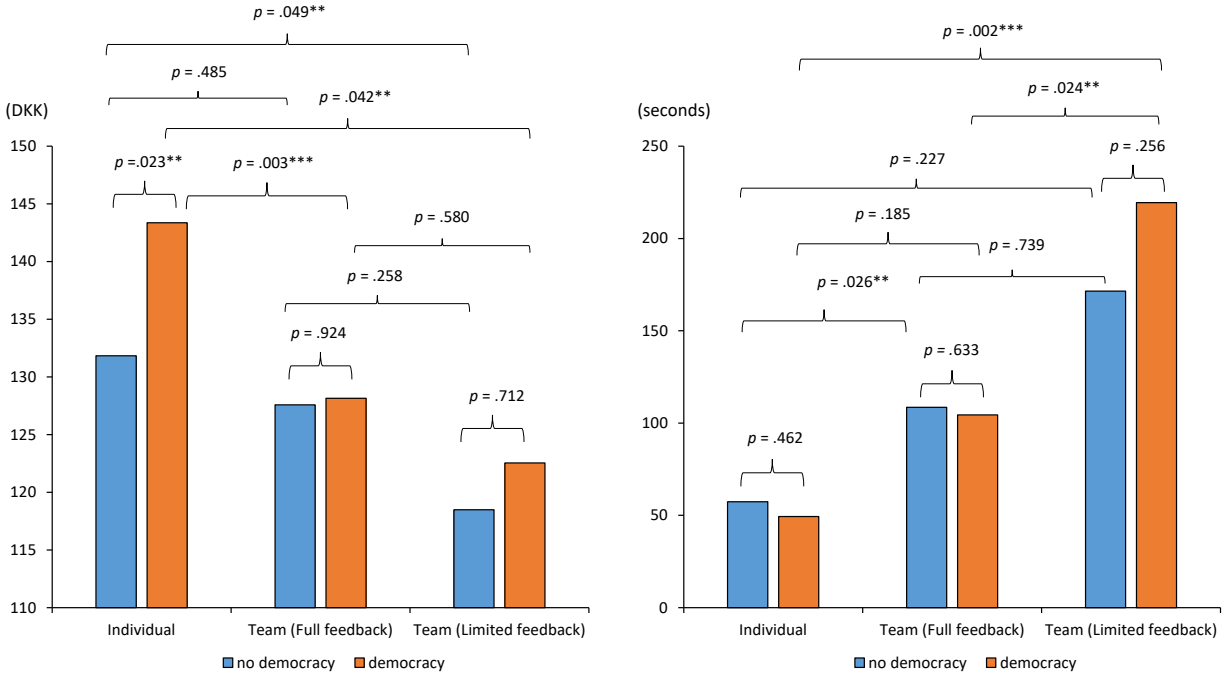
Result 3: *Consistent with Hypothesis 3, democratic task assignment improved work productivity under the individual-based remuneration scheme, but not under the team-based remuneration scheme. Subjects' video watching activities did not explain the positive effect of workplace democracy.*

We note that team members learned the vote outcomes in the Team (Full feedback) treatment, regardless of whether the voting decisions were implemented or not (i.e. in both the democratic and non-democratic teams). We discussed in Section 4 that when workers learn that others implement the task they prefer, they may expect the others to work hard, and consequently they may work hard themselves

distributions of scores in Appendix Figure C.2). In studying treatment differences, we used bootstrapped standard errors for all the regressions for consistency, considering that using a bootstrapping method generates a more conservative estimate and addresses the non-normal distribution in the Team (Full feedback) treatment.

due to interdependent preferences, such as reciprocity concerns. In both the democratic and non-democratic conditions, the subjects' performances were higher with the full feedback than with the limited feedback, and the time spent watching the video was also lower with the former than with the latter. The difference in video watching time is significant between the T-Dem (Full) and T-Dem (Limited) regimes, despite the fact that teams in these two regimes theoretically face the same situations in the main task-solving phase.[14] We emphasize, as summarized in Result 2, that free riding was a serious issue in all team treatments, including those with full information feedback. In addition, the results from the Team (Full feedback) treatments suggest that on average, unlike Hypothesis 2(b), neither motivation nor selection effects of democracy are important in the team-based remuneration scheme.

Figure 2. Average Performance (Earnings) in Part 2 by the Treatment Condition



(A) Per subject performance in Part 2

(B) Per subject video watching time

Notes: *p*-values (two-sided) for the differences between the democratic and non-democratic conditions in each treatment, and for the differences between the remuneration schemes under a given decision process (democracy, non-democracy). The *p*-values in panel A (panel B) were calculated based on subject random effect linear (tobit) regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID. The tobit model was used in panel B as 303 subjects (58.61% of the sample) did not watch the video at all. * $p < .1$; ** $p < .05$; *** $p < .01$.

[14] We speculate that the information about voting results might have been more salient in the T-Dem (Full) regime than in the T-Dem (Limited) regime because the participants in the full information treatment were told in the written instructions that this information was provided in both the democratic and non-democratic conditions, whereas in the T-Dem (Limited) regime, they were informed that it was provided only to those in the democratic condition. Hence, instructions highlighted the importance of information on voting outcomes somewhat more in the T-Dem (Full) regime than in the T-Dem (limited) regime.

Splitting the data by effort task (addition or counting) reveal further interesting patterns. Table 3 reports not only results on subjects' scores (numbers of correct responses) and video watching time, but also information on their attempts and success rates (shares of correct responses). Regardless of which task was used, team-based remuneration undermines subjects' scores and magnifies shirking (video watching) in most comparisons, consistent with Hypothesis 1(a).[15] The negative effects of team-based remuneration are mainly driven by the numbers of attempts, not by the success rates. The success rates take similar values across treatments and are not significantly different between the individual- and team-based remuneration schemes for any comparisons.[16]

A positive effect of workplace democracy is present in the addition task. The effects on the subjects' scores and attempts are significant at the 10% and 5% levels, respectively, under the individual-based remuneration scheme. Interestingly, there is also evidence of a positive effect of democracy in the Team treatments in the addition task. First, subjects' scores were significantly higher at the 10% level in the democratic than in the non-democratic condition in the Team (Full feedback) treatment. In addition, in this treatment workplace democracy significantly improved subjects' success rates and also limited their video watching weakly significantly. Second, in the Team (Limited feedback) treatment, although subjects spent more time watching the video under the democratic than under the non-democratic condition, their scores were higher (although insignificantly so) in the former than in the latter condition.

Effects of democracy are weaker and in some cases reversed in the counting task. This seems to suggest that democratic selection effects were weak for this task. As discussed earlier, performance in the addition task depends on skills as well as effort, and these skills (mathematical ability) differ across individuals. In the counting task, on the other hand, the required skills are likely to be possessed by most subjects. In the notation of the model in Section 4, c_{ik} and h_{ik} are more widely dispersed for addition than for counting. Therefore, it matters more for performance in the addition task than in the counting task whether people get to choose the task they work on or not.

Result 4: *(a) The strong effect of workplace democracy under the individual-based remuneration scheme was driven by those who selected the addition task. (b) Workplace democracy improved performance to some degree also under the team-based remuneration when the addition task was used, but not when the*

[15] The differences in the subjects' video watching time in the addition task are significant between the I-No and T-No (Full) regimes, and between the I-Dem and T-Dem (Limited) regimes at two-sided $p = .058$ and $< .001$, respectively. Also, the differences in the video watching time in the counting task are significant between the I-Dem and T-Dem (Full) (T-Dem (Limited)) regimes at two-sided $p = .077$ ($p = .060$).

[16] The tests were performed based on subject random effect linear regressions with bootstrapped standard errors clustered by session ID.

counting task was used.

Table 3. Average Performance in Part 2 by the Effort Task

	Individual			Team (Full feedback)			Team (Limited feedback)			
	(a) No democ. I-No	(b) Democracy I-Dem	H_0: (a) = (b)	(c) No democ. T-No (Full)	(d) Democracy T-Dem (Full)	H_0: (c) = (d)	(e) No democ. T-No (Limited)	(f) Democracy T-Dem (Limited)	H_0: (e) = (f)	All subjects
I. Addition task										
Score (# of correct responses)	62.4	73.0	.051*	61.9	68.7	.094*	56.0	61.1	.507	64.9
Attempts	72.4	85.0	.013**	72.8	78.2	.254	64.9	70.6	.363	75.2
Success rate (percent)	84.7	84.3	.903	82.3	87.5	.007***	86.0	84.6	.783	84.8
Time watching video (seconds)	69.5	49.9	.898	134.5	70.0	.084*	85.9	148.8	.091*	89.8
II. Counting task										
Score (# of correct responses)	55.3	56.4	.790	52.7	48.1	.102	49.7	49.1	.905	52.0
Attempts	65.7	67.3	.635	66.4	78.3[#1]	.456	63.2	58.9	.231	62.7
Success rate (percent)	83.3	82.3	.649	80.1	78.8	.912	84.7	83.4	.666	81.2
Time watching video (seconds)	46.2	48.8	.375	80.9	133.9	.239	247.6	266.5	.716	122.3

Notes: *p*-values (two-sided) for the differences between the democratic and non-democratic conditions were calculated based on subject random effect linear (tobit) regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID for the score, attempt and success rate variables (for the time watching video variable). [#1] There was an outlier who skipped a large number of questions. The record indicates that this subject attempted 1,276 questions. The average number of attempts is 59.0 if we exclude this subject.
* $p < .1$; ** $p < .05$; *** $p < .01$.

As discussed in Section 2, an advantage of our design is that it entails measurement of task preferences for all subjects. This feature enables us to perform detailed analyses by task preference, which helps us explore drivers behind free riding in teams and the effects of democratic task assignment. We first check whether the vote override procedure randomly assigned subjects into the democratic and non-democratic conditions. As shown in Appendix Table C.1, the random assignment was successful in the Individual treatment: each type's scores in Part 1 were balanced between the democratic and non-democratic conditions. In addition, it was also successful in the Team treatments, except for one case.[17] Hence, we can make a further analysis of the functioning of the free riding phenomenon and workplace democracy by conditioning on the task preferences.

[17] In the Team (Full feedback) treatment, among opponents of the counting task, high Part 1 performers were more likely to be assigned the counting task non-democratically, rather than democratically. This unequal outcome implies that estimates for the effect of democratic task assignment are likely to be conservative (i.e. biased somewhat downward) in this particular case.

Panel A of Table 4 reports subjects' average scores by task preference. Two clear patterns emerge for the addition task (panel A.I). First, Result 4a (positive effect of democracy under individual remuneration) is mainly driven by a selection effect. By design, only supporters of the addition task worked on addition in the I-Dem regime. In contrast, more than 60% of subjects assigned to the addition task in the I-No regime had voted for the counting task. The supporters of the counting task scored 17.8 percent lower than the subjects in the I-Dem regime (two-sided $p < .012$). The data also reveal that the supporters of the addition task performed higher in the democratic than in the non-democratic condition. This may be due to a motivation effect of workplace democracy as discussed in Hypothesis 2(b) – however, the difference is not significant. Second, the supporters of the addition task performed clearly better than those who voted for the counting tasks also in the two Team treatments, and the former were overrepresented in the democratic condition, relative to the non-democratic condition. This is consistent with Result 1.[18] Recall, however, that democratic task assignment on average did not have a significant effect on subjects' scores under team-based remuneration (Result 3). The reason is that democratic selection effects are stronger in the Individual than in the Team treatments, because all subjects in the democratic condition get to work on the task they desire in the former treatment, while a minority of subjects are overruled by the majority in the democratic condition of the latter treatments. This pattern is consistent with Hypothesis 3.

This picture changes drastically under the counting task (panel A.II). First, subjects' scores were indistinguishably similar between the task preferences, and also between the institutional environments (democracy or non-democracy), in the Individual treatment. Second, intriguingly, workplace democracy *negatively* affected performance among supporters of the counting task in the Team treatments. This effect was consistently seen in both the Team (Full feedback) and Team (Limited feedback) treatments, driven by their smaller number of attempts (see also Appendix Table C.2). As discussed in Section 3, the skill requirements are arguably less demanding in the counting task than in the addition task. While we assumed in Section 4 that democratic task assignment would not decrease motivations (i.e., $D \geq 0$), this result may mean that democratically selecting a less demanding task undermines cooperation by perversely influencing intrinsic motivation or beliefs.[19]

[18] The distribution of votes is significantly different between the democratic and non-democratic conditions at $p = .011$ in the Team (Full feedback) treatment, but the difference is not significant at $p = .142$ in the Team (Limited feedback) treatment, according to two-sided Fisher's exact tests.

[19] Dal Bó, Foster and Kamei (2019) found a negative effect on performance when subjects collectively selected a fixed payment, rather than introducing piece-rate payment, in a group. However, the effect was not significant. Then why did we have a stronger negative effect unlike Dal Bó, Foster and Kamei (2019)? There is a stark difference in the design between our paper and Dal Bó, Foster and Kamei (2019). While Dal Bó, Foster and Kamei

Panel B summarizes subjects' time spent watching the video, showing roughly similar patterns to Table 3 and Table 4.A. The data for video watching time are quite noisy and some results are affected by high outliers. We do not obtain any significant results because of the diminished sample sizes in comparing between the sub-categories within each effort task.

An important result emerging from Table 4 is that free riding in teams is largely driven by those who worked on a task they did not prefer. These workers preferred to work on the other task on which they had relative advantage (Result 1). Thus, this free riding result is a surprise, since theoretically, the most productive workers should free ride more than the least productive workers as discussed in Section 4 (Hypothesis 1(b)). Specifically, when supporters of the counting task were non-democratically assigned to the addition task, they scored 7.3% and 21.0% less in the Team (Full feedback) and Team (Limited feedback) treatments, respectively, relative to the Individual treatment. Likewise, when supporters of the addition task were non-democratically assigned the counting task, they scored 13.8% and 25.0% less in the Team (Full feedback) and Team (Limited feedback) treatments, respectively, relative to the Individual treatment. By contrast, free riding in teams is not visible among those who worked on their desired task. For example, supporters of the addition task answered on average 67.3, 69.6 and 66.3 addition questions correctly in the I-No, T-No (Full) and T-No (Limited) regimes, respectively. This result is remarkable because subjects could choose between only two tasks, both of which most subjects presumably find boring to perform for 30 minutes. Even in this setting, those who were assigned to their most desired task show little evidence of free riding in teams under the team-based remuneration scheme. A reason behind the no free riding result may be that workers who were assigned to their most desired task mistakenly believed that their peers have the same preferences and technologies as them. Under this condition, theoretically, it is possible that workers free ride less when they exhibit strong aversion to advantageous inequality (see Equation (7) in Section 4).

Similarly strong free riding behaviors by those who worked on a task they did not prefer can also be seen in the democratic condition (see again Table 4). This implies that the decision process through majority voting alone may not be enough to legitimate the mismatching and workers want more control.

(2019) studied a setup where one's play did *not* affect the payoff of the other group members, subjects' payoffs in the present paper *did* depend on the others' performances under the team-based remuneration. This difference in the design setup suggests that the strong negative effects in the Team treatments could have been driven by the nature of strategic interactions among team members. For example, workers' selecting of a less demanding task may self-justify their shirking and free riding behavior in teams. Another potential way to rationalize this result is to invoke workers' inequality aversion. Workers may believe that their team members enjoy working when the team adopted the counting task by voting. If this is the case, workers may not feel bad about free riding on others' effort provision.

Regression analyses confirm these results (Table 5). Columns (2) and (4) include only data from subjects who worked on the task they did *not* prefer.[20] The estimates of these columns show strong, negative and significant effects of team-based remuneration in the Team (Limited feedback) treatment. The effects of being in the Team (Full feedback) treatment are also negative, although coefficient estimates are somewhat smaller than for the Team (Limited feedback) treatment. By clear contrast, we do not find negative effects of team-based remuneration among those who preferred to work on addition (counting) when the addition (counting) task was assigned (see columns (1) and (3)).[21] We conclude that free riding in teams is largely driven by those who work on tasks they do not like. We note that the regression analyses also confirm the conclusion that workplace democracy significantly undermines work performance in teams in the counting task among those who preferred to work on counting (column (3)).

Result 5: *Unlike Hypothesis 1(b), free riding in teams is driven by those who work on an undesired task.*

We note that we also repeated the exercises of Table 5 but using time spent watching the video as the dependent variable instead of score and also using a tobit regression considering the nature of data (Appendix Table C.3). For the addition task, the same pattern as in Table 5 emerges: the effect of team-based remuneration is much stronger for those who do not work on the task they voted for than for those who did. For the counting task, results are more mixed. Parallel to the results discussed above, democracy magnified video watching activities among those who preferred to work on counting when the counting task was assigned in the Team (Limited feedback) treatment.

[20] As all subjects worked on the task they preferred in the I-Dem regime, the interaction term between the Individual treatment dummy and the Democracy dummy drops out of the model in these columns.

[21] The coefficient estimates for the Team (Full feedback) dummy and Team (Limited feedback) dummy show effects of team-based remuneration in the non-democratic conditions, in which task assignment is exogenous and hence the distributions of types are balanced across the treatments.

Table 4. Average Performance in Part 2 by Task Preference

(A) Scores (number of correct responses)

	Individual					Team (Full feedback)					Team (Limited feedback)				
	No democracy I-No		Democracy I-Dem		*p* (2-sided) for H_0:	No democracy T-No (Full)		Democracy T-Dem (Full)		*p* (2-sided) for H_0:	No democracy T-No (Limited)		Democracy T-Dem (Limited)		*p* (2-sided) for H_0:
	fraction	(a) score	fraction	(b) score	(a) = (b)	fraction	(a) score	fraction	(b) score	(a) = (b)	fraction	(a) score	fraction	(b) score	(a) = (b)
I. Addition task															
(i) Supporter of addition task	33.6%	67.3	100.0%	73.0	0.451	44.4%	69.6	70.4%	74.8	0.319	45.8%	66.3	70.8%	65.1	0.878
(ii) Supporter of counting task	66.4%	60.0		n.a.	n.a.	55.6%	55.6	29.6%	54.0	0.789	54.2%	47.4	29.2%	51.3	0.789
Number of subjects	43		45			54		54			24		24		
II. Counting task															
(i) Supporter of addition task	44.7%	55.2		n.a.	n.a.	49.0%	47.6	23.8%	42.1	0.331	44.4%	41.4	25.0%	43.0	0.858
(ii) Supporter of counting task	55.3%	55.4	100.0%	56.4	0.839	51.0%	57.6	76.2%	50.0	0.038**	55.6%	56.3	75.0%	51.2	0.045**
Number of subjects	47		49			51		63			27		36		

(B) Video watching time (seconds)

	Individual					Team (Full feedback)					Team (Limited feedback)				
	No democracy I-No		Democracy I-Dem		*p* (2-sided) for H_0:	No democracy T-No (Full)		Democracy T-Dem (Full)		*p* (2-sided) for H_0:	No democracy T-No (Limited)		Democracy T-Dem (Limited)		*p* (2-sided) for H_0:
	fraction	(a) time	fraction	(b) time	(a) = (b)	fraction	(a) time	fraction	(b) time	(a) = (b)	fraction	(a) time	fraction	(b) time	(a) = (b)
I. Addition task															
(i) Supporter of addition task	33.6%	153.9	100.0%	49.9	0.211	44.4%	103.5	70.4%	51.2	0.418	45.8%	99.4	70.8%	152.5	0.302
(ii) Supporter of counting task	66.4%	28.7		n.a.	n.a.	55.6%	159.4	29.6%	114.7	0.386	54.2%	74.4	29.2%	139.8	0.138
Number of subjects	43		45			54		54			24		24		
II. Counting task															
(i) Supporter of addition task	44.7%	74.8		n.a.	n.a.	49.0%	91.6	23.8%	179.1	0.243	44.4%	412.3	25.0%	319.5	0. 749
(ii) Supporter of counting task	55.3%	23.1	100.0%	48.8	0.212	51.0%	70.7	76.2%	119.7	0.285	55.6%	115.9	75.0%	248.9	0.133
Number of subjects	47		49			51		63			27		36		

Note: *p*-values (two-sided) were calculated based on subject random effect linear (tobit) regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID in panel A (panel B). * $p < .1$; ** $p < .05$; *** $p < .01$.

Table 5. Free Riding in Teams by Task Preference

Dependent variable: Scores (number of correct responses)

	I. Addition Task		II. Counting task	
	Assigned task = chosen task	Assigned task ≠ chosen task	Assigned task = chosen task	Assigned task ≠ chosen task
Independent variable:	(1)	(2)	(3)	(4)
a. Team (Full feedback) dummy	2.339 (8.181)	-4.367 (4.951)	2.154 (4.383)	-7.630 (5.003)
b. Team (Limited feedback) dummy	-1.013 (8.532)	-12.615*** (2.847)	.844 (3.131)	-13.773*** (4.004)
[Democracy effect:]				
Individual treatment dummy × Democracy dummy	5.670 (7.116)	---	.985 (4.296)	---
Variable a × Democracy dummy	5.217 (5.587)	-1.633 (6.902)	-7.598* (4.359)	-5.427 (4.501)
Variable b × Democracy dummy	-1.155 (3.541)	3.901 (15.727)	-5.119** (1.931)	1.583 (9.580)
Constant	67.286*** (7.033)	60.000*** (2.413)	55.423*** (2.990)	55.190*** (2.177)
Prob > Wald χ^2	.596	.035	.001	.001
# of observations	149	95	191	82

Notes: Subject random effect linear regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID. Numbers in parenthesis are standard errors. The reference group is observations in the I-No regime. * $p < .1$; ** $p < .05$; *** $p < .01$.

5.3. The Dynamics of effort provision

As the duration of the effort task in Part 2 was a full 30 minutes (considerably longer than for the typical task-solving phase in real effort experiments), we are well placed to study the dynamics of free riding and effects of workplace democracy. Figure 3 shows average performance over time in each task. Apart from a learning effect during the first minute, and a "sprint to the finish line" toward the end, performance levels are reasonably stable over time. However, a close look at the dynamics not only confirms our key findings reported in Section 5.2, but it also uncovers intriguing new patterns.

First, the dynamics provide a strong support to Hypothesis 2: workplace democracy improved productivity on the addition task in all three treatments. The dynamics, nevertheless, differ by the treatment (Figure 3.A). On the one hand, large positive effects of democracy persisted throughout the main task-solving phase in the Individual treatment. In the two team treatments, on the other hand, positive effects of workplace democracy were observed in the second half of the task-solving phase, but not in the first half. A regression analysis confirms these patterns, showing that (a) workplace democracy was a strong predictor, significant at the 1% level, of performance in the addition task in the Individual treatment (see column (1) of Appendix Table C.4), and that (b) performance increased significantly [weakly significantly] more over time in the democratic than in the non-democratic condition in the Team (Full feedback) [Team (Limited feedback)] treatment (see columns (2) and (3) of Appendix Table C.4).

Second, there was no evidence for positive effects of workplace democracy when the counting task was used, whether in early or in late parts of the task-solving phase (Figure 3, columns (4)-(6) of Appendix Table C.4). This is consistent with the observations made in Section 5.2.

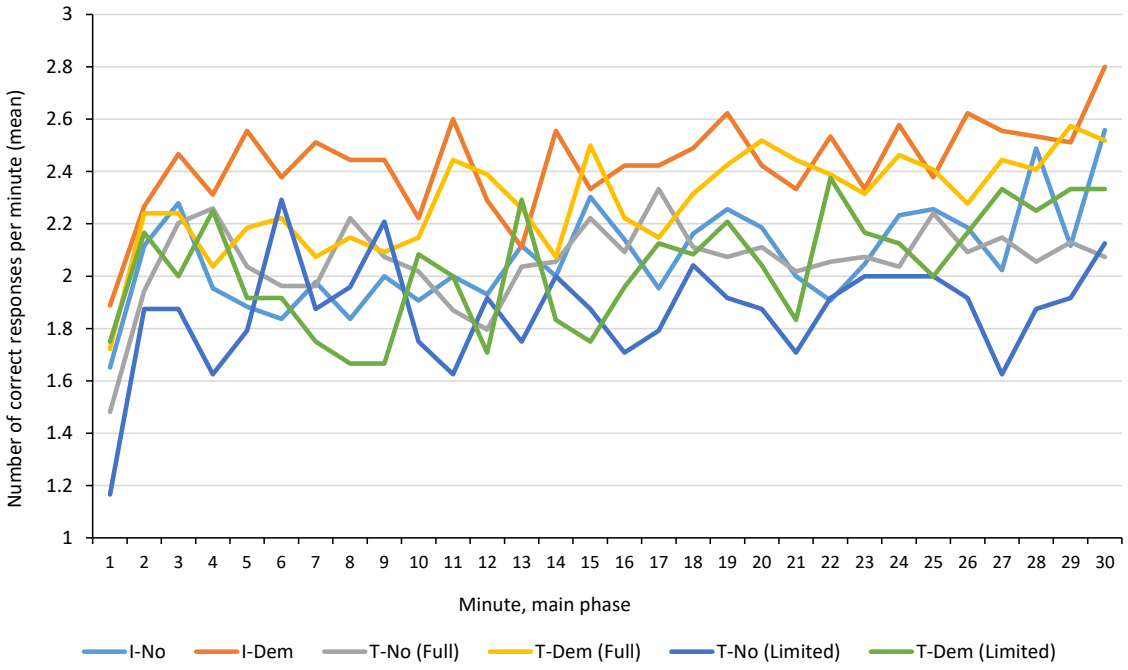
Result 6: *(a) The strongly positive effect of workplace democracy in the addition task lasted across the entire 30-minute task-solving phase under the individual-based remuneration. (b) Performance in the addition task increased significantly over time in the Team treatments when it was assigned democratically rather than non-democratically.*

Figure 4 shows dynamics for free riding in teams, depending on whether people worked on their desired task or not. We depicted the dynamics separately for each decision process (democracy or non-democracy). The data in the non-democratic condition are better suited to study effects of task preferences on workers' effort provision since there are no effects of democracy, including democratic selection effects, involved. Consistent with the results presented in Table 5 above, the figure shows no free riding among those who worked on their most desired tasks throughout the 30-minutes task-solving phase (panel A.a(i), panel B.a(i)). In contrast, free riding by those who work on undesired tasks is

significant. Under the addition task, the performance difference between the Individual and the Team treatments gradually increased over time because performance improved in the Individual treatment while it remained stable in the Team treatments (panel A.a(ii)). Under the counting task, a large degree of free riding was seen among those who preferred to work on addition throughout the entire task-solving phase (panel B.a(ii)).[22] These results strengthen Result 5.

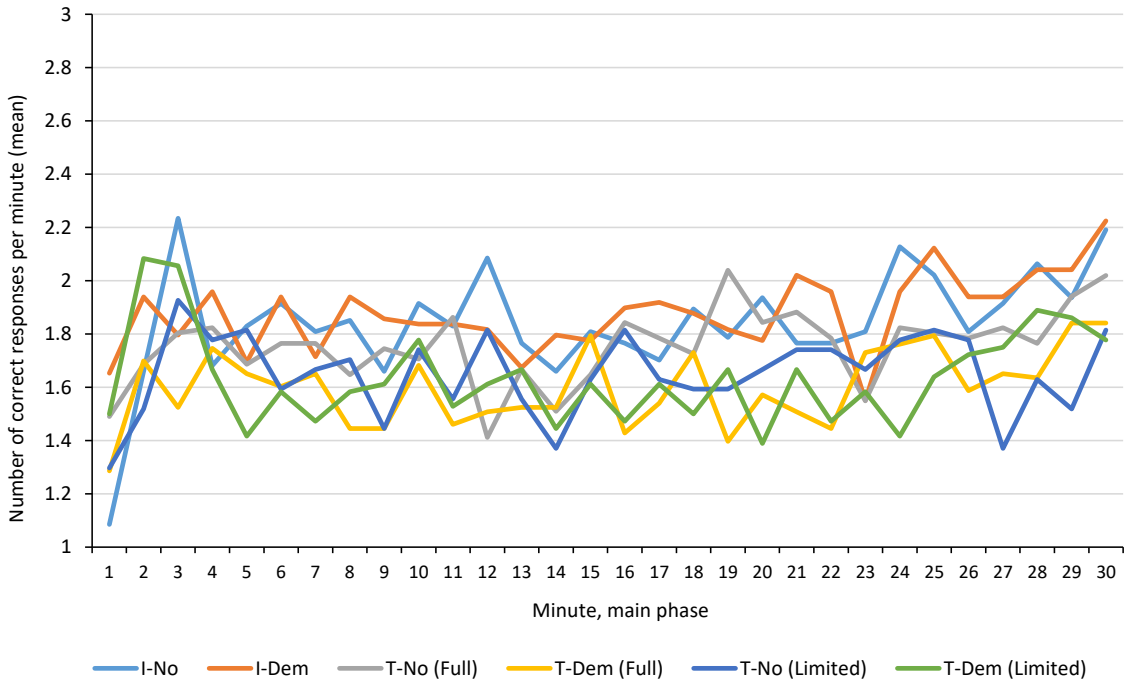
Panels A.b(ii) and B.b(ii) of Figure 4 confirms that the strong free riding phenomenon in teams in the democracy condition was largely driven by those who worked on undesired tasks. In the addition task, there is some evidence of free riding in teams even among those who worked on the task they desired and were in the democratic condition (panel A.b(i)). However, this effect is only present in the T-Dem (Limited) regime and gradually diminishes over time.

Figure 3. Performance (Score) Over Time



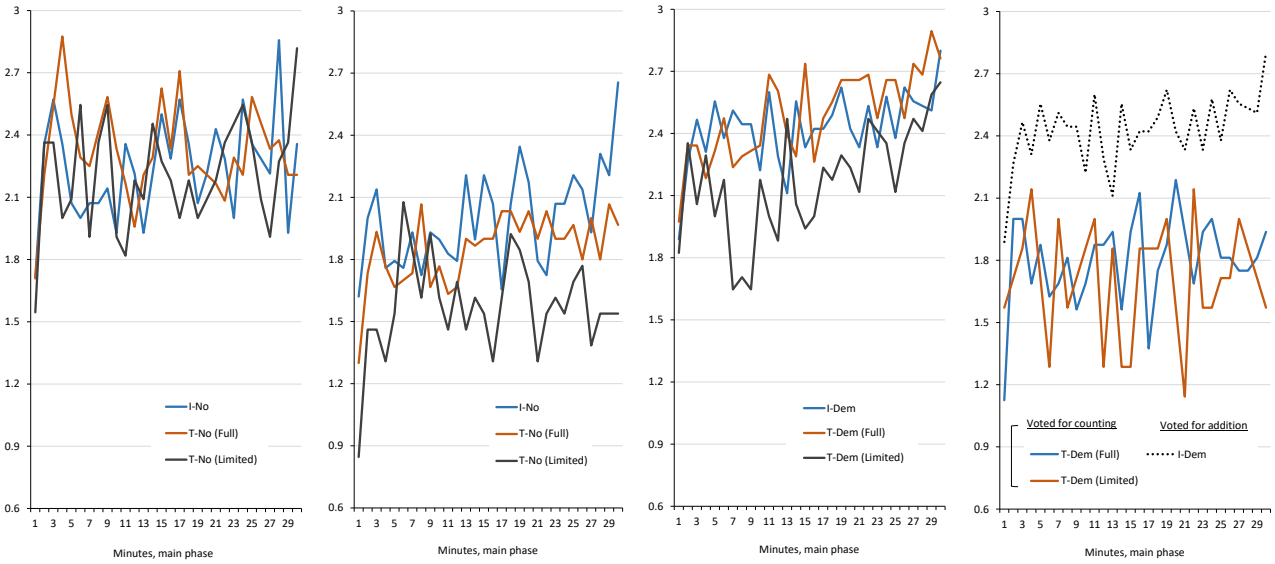
(A) Addition Task

[22] A regression analysis, shown in Appendix Table C.5 (column (2) in panel A, column (2) in panel B), confirms the result for the addition task, but not for the counting task due to high variations in subjects' scores.



(B) Counting task

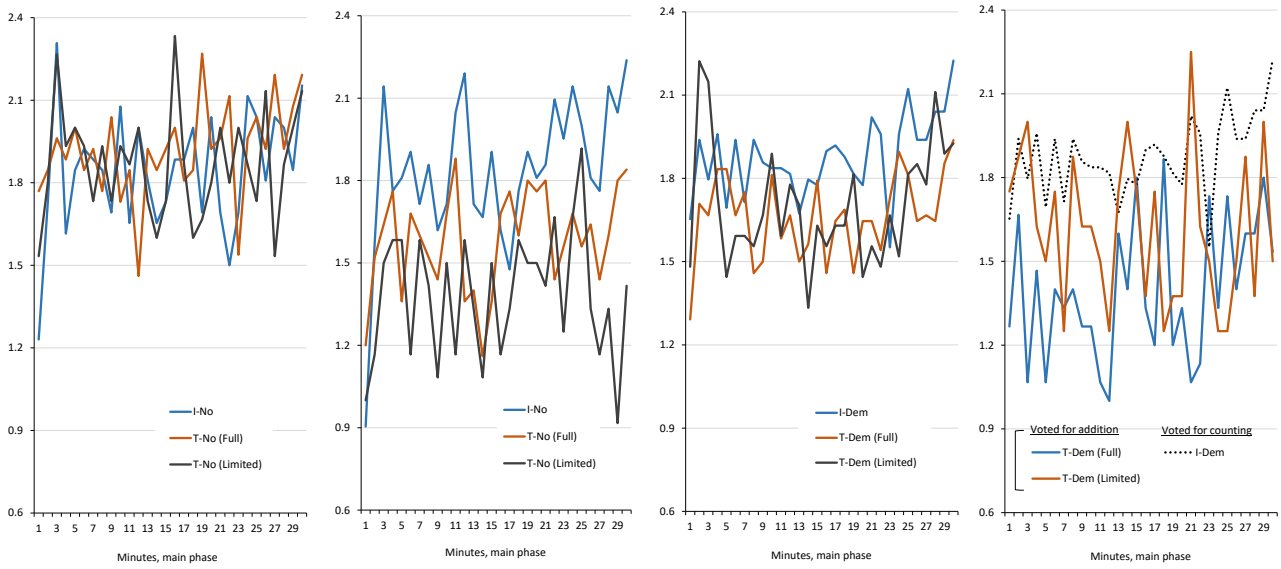
Figure 4. Dynamics of Performance by Task Preference



(i) Voted for addition task (ii) Voted for counting task (i) Voted for addition task (ii) Voted for counting task

(a) The non-democratic condition (b) The democratic condition

A. Scores in the addition task



(i) Voted for counting task (ii) Voted for addition task (i) Voted for counting task (ii) Voted for addition task

(a) The non-democratic condition (b) The democratic condition

B. Scores in the counting task

6. Conclusion

Using data from a real effort experiment, this paper investigated the relationship between workers' task preferences, free riding in teams and workplace democracy. Results showed strong evidence of free riding. People exert less effort when remuneration is based on revenue sharing (average performance in a team) than when it is based solely on their own performance (piece-rate payment). An important, new result is a strong interaction between task assignment and free riding. Free riding is much more prevalent among people working on a task they did not desire than among those who got the task they wanted. This result contrasts with the prediction from a standard model of utility from income and disutility of effort. This model predicts that the most productive workers (i.e. predominantly those who work on the task they desire) free ride the most, in absolute terms. The empirical result suggests that factors other than income maximization and inequity aversion are important for effort choice. For example, a worker may identify less strongly with her team when the team works on a task she does not like, and therefore she may free ride strongly.

The literature has almost neglected sorting among different tasks, presumably due to the unavailability of field data to study this causal relationship. However, mismatching between workers' task preferences and assignment could easily happen in real workplaces, since task preferences are often not

fully observed by employers. Our results underscore the importance of devising efficient task assignment mechanisms within firms.

We further studied workplace democracy (defined as worker influence on task assignment) as a possible remedy against the matching problem and against free riding in teams. Results showed that workplace democracy mitigates the matching problem and therefore increases productivity. Democracy allows more workers to work on the task they like best, which in turn increases work performance. This sorting effect is stronger when workers choose tasks individually than when they do so by majority voting in teams. The reason is that a minority of team members do not get to work on the task they desire, even when tasks are chosen democratically. An implication is that firms may benefit from allowing workers some flexibility in terms of shifting between teams working on different tasks. On the other hand, this may give rise to adverse selection issues, since workers may select teams based on *other* team members' skills, rather than their own. If the team working on task A has very productive members, it might be attractive for a worker to join the team, even if he is not skilled at task A himself. How to form efficient teams in an environment of heterogeneous tasks is an interesting avenue for future research.

We also investigated whether workplace democracy, in addition to facilitating productivity-enhancing sorting, also increases performance through information or motivation effects. In contrast to findings in the literature on social dilemmas (e.g. Sutter, Haigner and Kocher 2010, Dal Bó, Foster and Putterman 2010), we only found limited evidence of such effects.

Lastly, we remark that our results have implications for research on organizations and incentives in workplace using real effort or field experiments. At the beginning of the paper, we discussed how workers' free riding in teams remained empirically unsettled in real effort experiments. The findings presented in our paper suggest that the answer to a given research question, including possible free riding in teams, may depend on researchers' selection of effort tasks and on the distribution of workers' task preferences (which are usually unobserved by the researchers). Workers' behavioral responses to effort tasks and incentives may differ according to their preferences towards tasks assigned in a given study.

References

Abeler, J., Falk, A., Goette, L., Huffman, D. (2009). Reference Points and Effort Provision. *American Economic Review*, 101(2), 470-492.

Alchian, A., Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *American Economic Review*, 62(5), 777-95.

Araujo, F., Carbone, E., Conell-Price, L., Dunietz, M., Jaroszewicz, A., Landsman, R., Lamé, D., Vesterlund, L., Wang, S., Wilson, A. (2016). The slider task: an example of restricted inference on incentive effects. *Journal of Economic Science Association*, 2(1), 1-12.

Barro, J., Beaulieu, N. (2003). Selection and Improvement: Physical Responses to Financial Incentives. National Bureau of Economic Research working paper 10017.

Bergstrom, T., Blume, L., Varian, H. (1986). On the Private Provision of Public Goods. *Journal of Public Economics*, 29, 25-49.

Blume, A., Ortmann, A. (2007). The effects of costless pre-play communication: Experimental evidence from games with pareto-ranked equilibria. *Journal of Economic Theory*, 132(1), 274-290.

Boning, B., Ichniowski, C., Shaw, K. (2007). Opportunity Counts: Teams and the Effectiveness of Production Incentives. *Journal of Labor Economics*, 25(4), 613-650.

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

Cadsby, B., Song, F., Tapon, F. (2007). Sorting and Incentive Effects of Pay for Performance: An Experimental Investigation. *Academy of Management Journal*, 50(2), 387-405.

Charness, G. (2000). Self-Serving Cheap Talk: A Test of Aumann's Conjecture. *Games and Economic Behavior*, 33, 177-194.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14, 47-83.

Cooper, R., DeJong, D., Forsythe, R., Ross, T. (1992). Communication in coordination games. *Quarterly Journal of Economics*, 107, 739-771.

Corgnet, B., Hernan-Gonzalez, R.,Rassenti, S. (2011). Real effort, real leusire and real-time supervision: Incentives and peer pressure in virtual organizations. Working Paper, Chapman University.

Dal Bó, P., Foster, A., Kamei, K. (2019). The democracy effect: A weights-based identification strategy. National Bureau of Economic Research working paper 25724.

Dal Bó, P., Foster, A., Putterman, L. (2010). Institutions and Behavior: Experimental Evidence on the Effects of Democracy. *American Economic Review*, 100(5), 2205-2229.

Deci, E. L., Ryan, R.M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Deci, E. L., Ryan, R.M. (2000). The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227-268.

Dohmen, T., Falk, A. (2011). Performance Pay and Multi-Dimensional Sorting: Productivity,

Preferences and Gender. *American Economic Review*, 101(2), 556-590.

Dow, Gregory. (2003). *Governing the Firm - Workers' Control in Theory and Practice*. Cambridge University Press.

Duffy, J., Feltovich, N. (2002). Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talk. *Games and Economic Behavior*, 39, 1-27.

Duffy J., Feltovich, N. (2006). Words, deeds, and lies: Strategic behavior in games with multiple signals. *Review of Economic Studies*, 73, 669-688.

Eckarz, K., Kirchkamp, O., Schunk, D. (2012). How do incentive affect creativity? Jena Economic Research Papers 2012-2068

Erev, I., Bornstein, G., Galili, R. (1993). Constructive intergroup competition as a solution to the free rider problem: A field experiment. *Journal of Experimental Social Psychology*, 29, 463-478.

Eriksson, T., Teyssier, S., Villeval, M.C. (2009). Self-selection and efficiency of tournaments. *Economic Inquiry*, 47(3), 530-548.

Eriksson, T., Villeval, M.C. (2018). Performance-pay, sorting and social motivation. *Journal of Economic Behavior & Organization*, 69, 412-421.

Fehr, E., Schmidt, K. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114(3), 817-868.

Fehr, E., Schmidt, K. (2010). On Inequity Aversion: A Reply to Binmore and Shaked. *Journal of Economic Behavior & Organization*, 73, 101-108.

Falk, A., Huffman, D., Mierendorff, K. (2006). Incentive properties and political acceptability of workfare: Evidence from real effort experiments. Working Paper.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.

Gächter, S., Huang, L., Sefton, M. (2016). Combining "real effort" with induced effort costs: The ball-catching task. *Experimental Economics*, 19, 687-712.

Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association*, 1(1), 114-125.

Hamilton, B., J. Nickerson, Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, 111(3), 465-497.

Hansen, D. G. (1997). Worker Performance and Group Incentives: A Case Study. *Industrial and Labor Relations Review*, 51(1), 37-49.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, 13, 324–340.

Ichniowski, C., Kochan, T. A., Levine, D., Olson, C., Strauss, G. (1996). What Works at Work: Overview and Assessment. *Industrial Relations*, 35(3), 299-333.

Ichniowski, C., K. Shaw, Prennushi, G. (1997). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *American Economic Review*, 87(3), 291-313.

Jensen, T., Markussen, T. (2019). Group size, signaling and the effect of democracy. mimeo, University of Copenhagen.

Kamei, K. (2016). Democracy and resilient pro-social behavioral change: An experimental study. *Social Choice and Welfare*, 47, 359-378.

Kamei, K. (2019). Cooperation and Endogenous Repetition in an Infinitely Repeated Social Dilemma: Experimental Evidence. *International Journal of Game Theory*, 48(3), 797-834.

Kruse, D. L. (1992). Profit Sharing and Productivity: Microeconomic Evidence from the United States, *Economic Journal*, 102(410), 24-36.

Lazear, E. (2000). Performance Pay and Productivity. *American Economic Review*, 90(5), 1346-1361.

Ledyard, J. (1995). Public Goods: A Survey of Experimental Research, in *Handbook of experimental economics*, ed. by J. H. Kagel and A. E. Roth. Princeton: Princeton University Press, 111-94.

Leuven, E., Oosterbeek, H., Sonnemans, J., van der Klaauw, B. (2011). Incentives versus Sorting in Tournaments: Evidence from a Field Experiment. *Journal of Labor Economics*, 29(3), 637-658.

Lezzi, E., Fleming, P., Zizzo, J. (2015). Does It Matter Which Effort Task You Use? A Comparison of Four Effort Tasks When Agents Compete for a Prize. Working Paper, Newcastle University.

Markussen, T., Putterman, L., Tyran, J.-R. (2014). Self-Organization for Collective Action: An Experimental Study of Voting on Sanction Regimes. *Review of Economic Studies*, 81(1), 301-324.

Melizzo, P., Carpenter, J., Matthews, P.H. (2014). Workplace democracy in the lab. *Industrial Relations Journal*, 45(4), 313-328.

Niederle, M., Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *Quarterly Journal of Economics*, 122(3), 10-67.

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, 37(1), 7-63.

Román, F. (2003). An analysis of changes to a team-based incentive plan and its effects on productivity, product quality, and absenteeism. *Accounting, Organizations and Society*, 34, 589-618.

Samuelson, P. (1954). The Pure Theory of Public Expenditure. *Review of Economics and Statistics*, 36(4), 387-389.

Sloof, R., van Praag, M. (2010). The effect of noise in a performance measure on work motivation: A real effort laboratory experiment. *Labour Economics*, 17, 751-765.

Stiglitz, J. (1975). Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy. *Bell Journal of Economics*, 6(2), 552-579.

Sutter, M., Haigner, S. Kocher, M. (2010). Choosing the Stick or the Carrot? – Endogenous Institutional Choice in Social Dilemma Situations. *Review of Economic Studies*, 77, 1540-1566.

Tyran, Jean-Robert Feld, L. (2006). Achieving Compliance when Legal Sanctions are Non-deterrent. *Scandinavian Journal of Economics*, 108(1), 1-22.

Van Dijk, F., Sonnemans, J., Van Winden, F. (2001). Incentive systems in a real effort experiment. *European Economic Review*, 45, 187-214.

Electronic Supplementary Material for

Kamei and Markussen (2020):

"Free Riding and Workplace Democracy

– Heterogeneous Task Preferences and Productivity Sorting"[#]

This Appendix includes theoretical analyses, the instructions, and additional figures and tables that supplement Kamei and Markussen (2020).

[#] Kamei: Durham University, Email: kenju.kamei@gmail.com; kenju.kamei@durham.ac.uk. Markussen: University of Copenhagen, Email: Thomas.Markussen@econ.ku.dk.

Appendix A: Inequity-averse Subjects' Optimal Effort Provision under the Team-based Remuneration Scheme

Suppose that subjects have inequity-averse preferences as discussed in the paper. In this part of the Appendix, we show that $e_1 = e_2 = e_3 = e^{***} \in \left[\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}, \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}\right]$ in equilibrium if $(\alpha, \beta) \neq (0, 0)$.

Step 1: $e_1 = e_2 = e_3$ in equilibrium.

Without loss of generality, suppose that $e_1 \leq e_2 \leq e_3$. In this case, member 1 receives the highest payoff, while member 3 receives the lowest payoff in the team. The optimal effort level of member 1 (e_1^{***}) is $\frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}$ because $\frac{\partial U_{1k}}{\partial e_1} = \frac{c_k}{3} - (1-\beta)[2(h_k - \delta D)e_1 + 1]$.

With a similar calculation, we can also find that the optimal effort level of member 3 (e_3^{***}) is given by $\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}$ because $\frac{\partial U_{ik}}{\partial e_i} = \frac{c_k}{3} - (1+\alpha)[2(h_k - \delta D)e_i + 1]$.

These calculations suggest that $e_3^{***} = \frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)} < \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)} = e_1^{***}$ unless α and β are both equal to 0. This contradicts our initial assumption of $e_1 \leq e_3$.

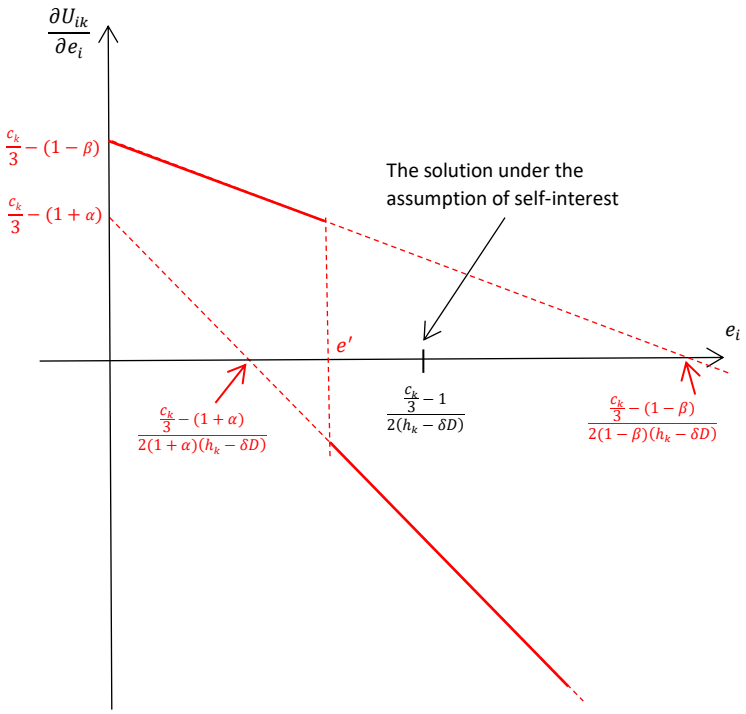
Step 2: $e_1^{***} = e_2^{***} = e_3^{***} = e^{***} \in \left[\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}, \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}\right]$.

Consider member *i*'s decision to work, given that the other two team members' effort levels are e'. *i*'s optimal response is determined by the following condition:

$$\frac{\partial U_{ik}}{\partial e_i} = \frac{c_k}{3} - (1-\beta)[2(h_k - \delta D)e_i + 1] \text{ when } e_i < e'. \quad \text{(A1)}$$

$$\frac{\partial U_{ik}}{\partial e_i} = \frac{c_k}{3} - (1+\alpha)[2(h_k - \delta D)e_i + 1] \text{ when } e_i \geq e'. \quad \text{(A2)}$$

As shown in the graph below, $e_i = e'$ can be an optimal effort provision if $e' \in \left[\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}, \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}\right]$. This suggests that the team members face a coordination problem whose set of equilibrium is $e_1 = e_2 = e_3 = e^{***} \in \left[\frac{\frac{c_k}{3}-(1+\alpha)}{2(1+\alpha)(h_k-\delta D)}, \frac{\frac{c_k}{3}-(1-\beta)}{2(1-\beta)(h_k-\delta D)}\right]$.


$\frac{\partial U_{ik}}{\partial e_i}$
$\frac{c_k}{3}-(1-\beta)$
$\frac{c_k}{3}-(1+\alpha)$
The solution under the assumption of self-interest
e_i
e'
$\frac{c_k}{3}-(1+\alpha)$ / $2(1+\alpha)(h_k-\delta D)$
$\frac{c_k}{3}-1$ / $2(h_k-\delta D)$
$\frac{c_k}{3}-(1-\beta)$ / $2(1-\beta)(h_k-\delta D)$

Appendix B: Experimental Instructions

(1) The Individual treatment

Welcome

You are now taking part in a decision-making experiment. You will be able to earn money depending on your decisions and also on the decisions of other participants in some tasks. Please read the following instructions carefully.

During the experiment you are not allowed to communicate with other participants. If you have questions, please raise your hand. One of us will come to answer your question.

This experiment has two parts. In Part 1 you engage in two tasks, each of which lasts three minutes. In Part 2, you engage in one of the two tasks based on your choice.

Part 1

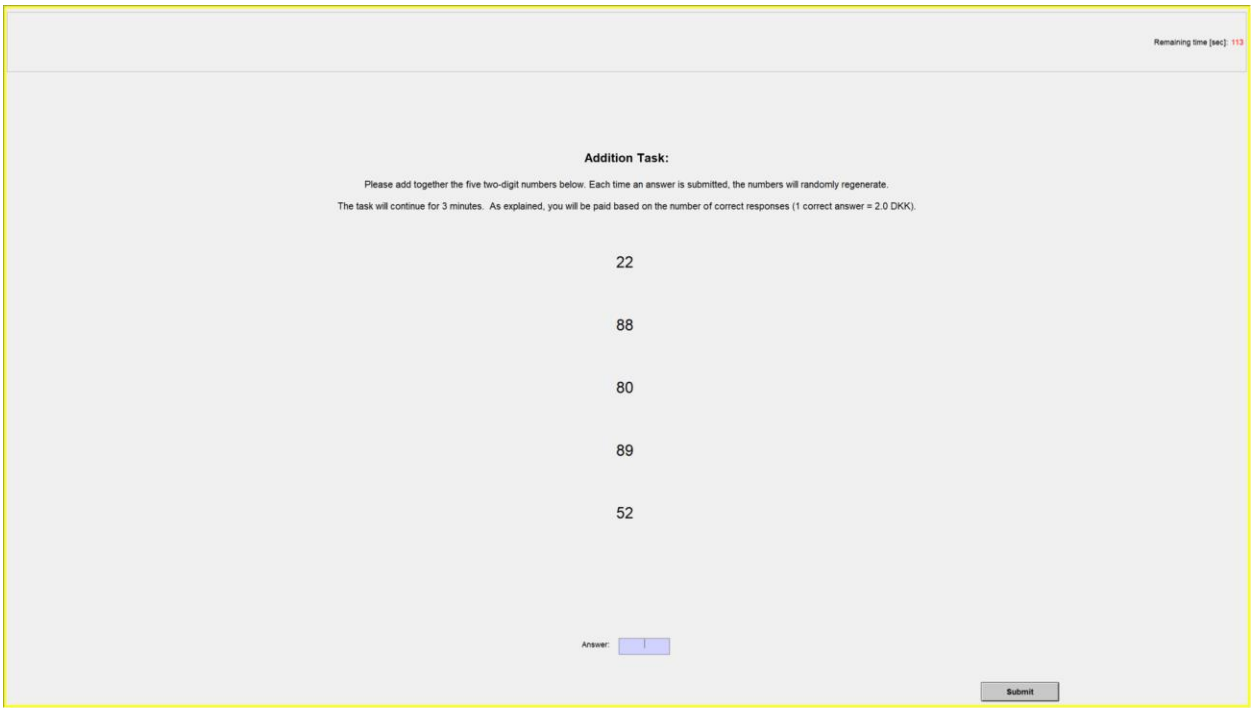
In Part 1, you will undertake two tasks: (a) **an addition task** for three minutes, and (b) **a counting zeros task** for three minutes. The order of tasks (which task you will undertake first) is randomly selected for each participant. We will now explain each task.

(a) The addition task:

You will be asked to answer addition questions. In each task, you will add five two-digit numbers. The more questions you answer correctly, the more money you earn.

Once you add the five numbers and fill the blank with the sum, please click the "Submit" button. Then, you will move on to the next question and five two-digit numbers will be randomly generated. We note that if you do not want to answer a question, you can skip it by simply pressing the Submit button without writing an answer. You will not be given feedback as to whether your calculation is correct or not for each question. However, at the end of the 3 minutes, you will be informed of the number of questions you answered correctly. You will be paid based only on your own performance. You will receive 2 Danish Kroner per correct answer. Any questions?

Screen Image:



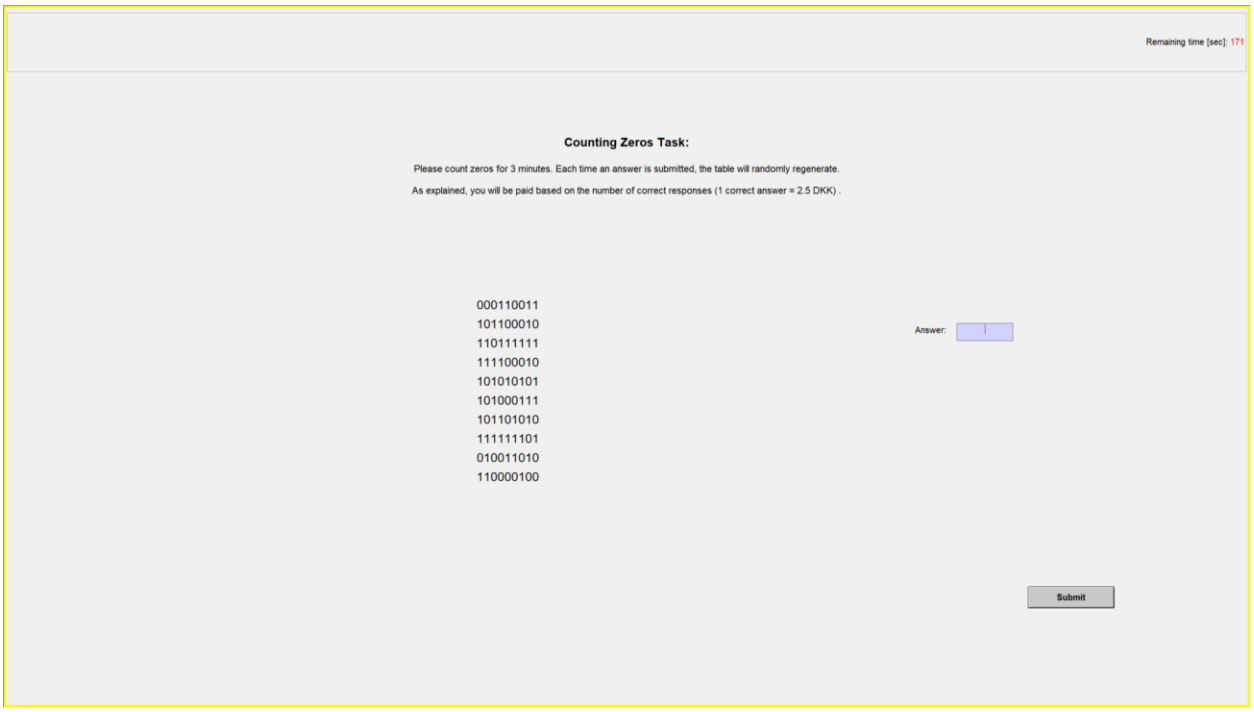
Note: Numbers shown in the screen shot are for illustration only.

(b) The counting zeros task:

You will be asked to count the number of zeros appearing on your screen. The more counting tasks you solve correctly, the more money you earn. Each question has a table which consists of 90 randomly ordered 0s and 1s as in the following screen image. Your task is to count the number of 0s in the table.

Once you count the number of 0s and fill the blank with the number, please click the "Submit" button. Then, you will move on to the next question and a new table with 0s and 1s will be randomly generated. We note that if you do not want to answer a question, you can skip it by simply pressing the Submit button without writing an answer. As in the addition task, you will not be given feedback as to whether your answer is correct or not for each question. However, at the end of the 3 minutes, you will be informed of the number of questions you answered correctly. You will be paid based only on your own performance. Specifically, you will receive 2.5 Danish Kroner per correct answer. Any questions? Once you finish undertaking the two tasks, you will move on to Part 2.

Screen Image:

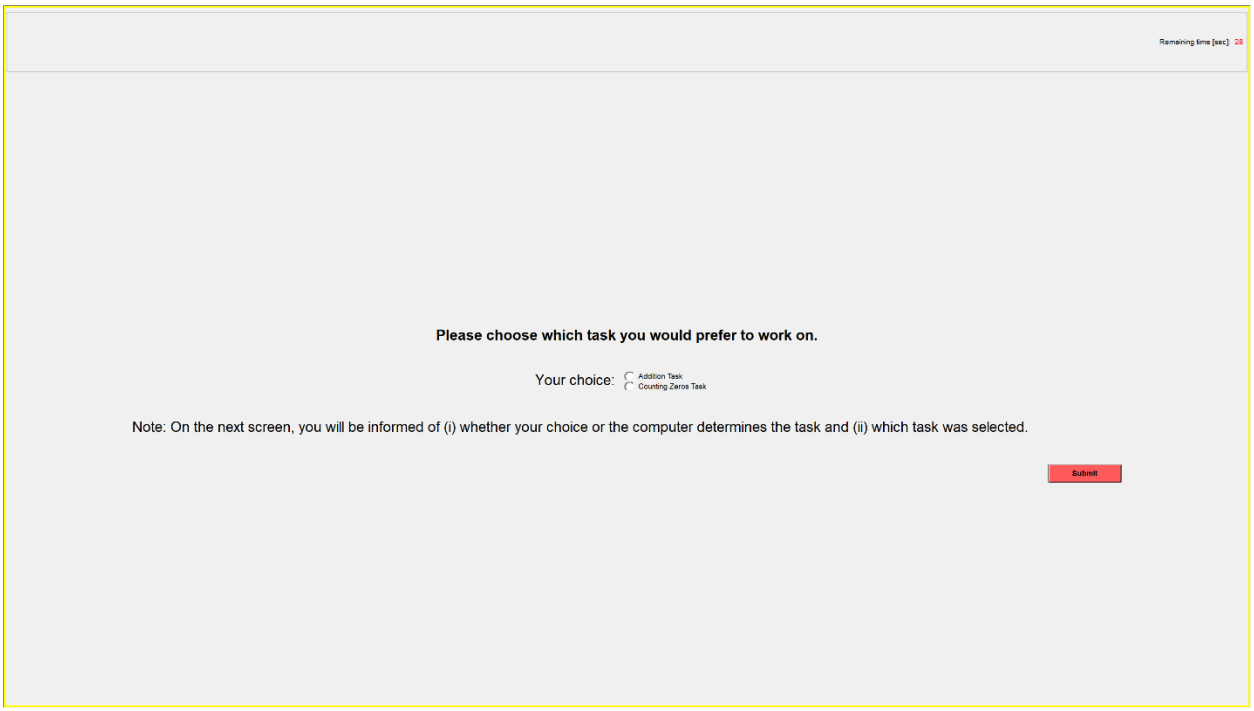


Note: Numbers shown in the screen shot are for illustration only.

Part 2

In this part, you will perform one of the two tasks (the addition or counting zeros task) for a longer duration – namely 30 minutes. We will now explain how Part 2 proceeds.

Step 1: You decide which task you want to undertake in this part.



Step 2: After you have made your choice, the computer **randomly (i.e., with a probability of 50%) decides whether or not to consider your choice**.

- If the computer decides to consider you choice, **you will use the task you selected in step 1**.
- If the computer decides not to consider your choice, it will **randomly choose which task to be used** regardless of what you chose.

At the end of step 2, you will be informed whether or not your choice was considered and which task was selected.

Step 3: You will undertake the selected task **for 30 minutes**. You will receive 2 Danish Kroner per correct answer if you undertake the addition task. You will receive 2.5 Danish Kroner per correct answer if you undertake the counting zeros task.

In the work stage (Step 3), you can take a rest and watch an entertaining video clip (see the screen image below) whenever you like. To hear the audio track of the video, please use the headphone on your desk. There is a button named "Watch Video." When you click this button, you will move on to another screen

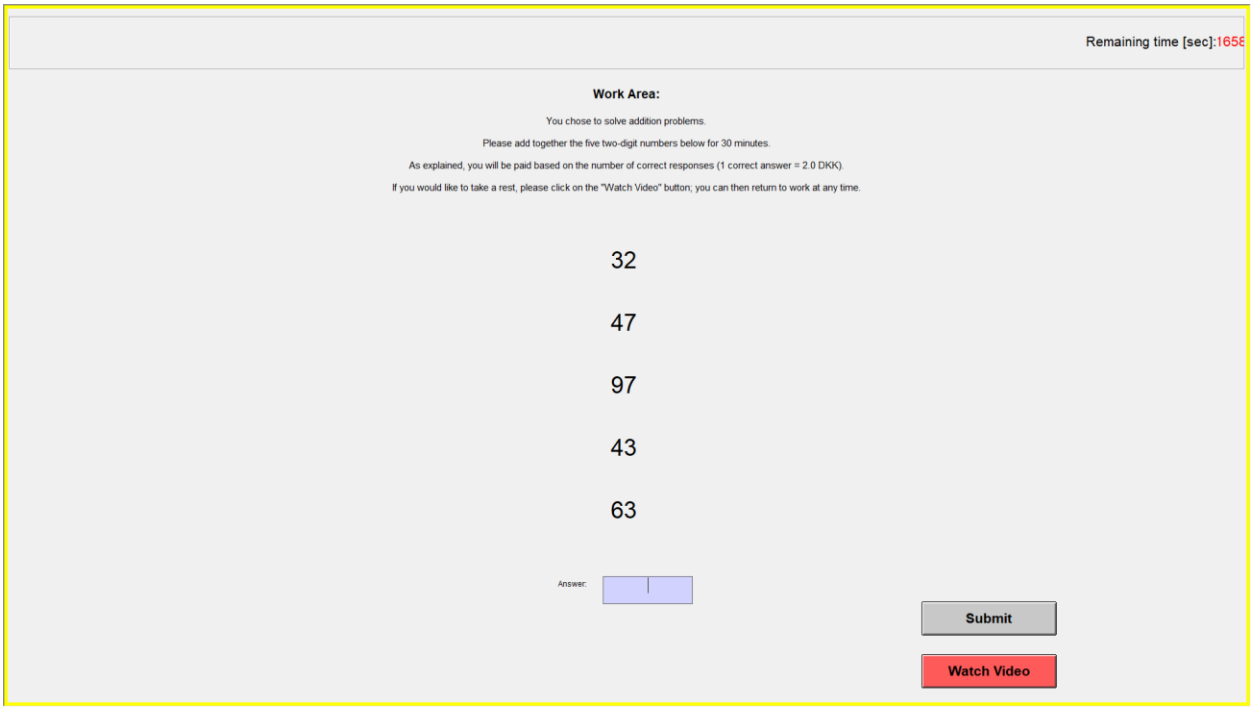
that contains a video clip. You can pause the video and return to the work site by clicking the "Pause and return to work" button freely. You will earn a small amount of money by watching the video clip also instead of working on the selected task. In particular, you will receive 1 Danish Krone per minute of watching the video clip. Note that you cannot watch the video while working on a task.

Please remain silent at all times during the experiment.

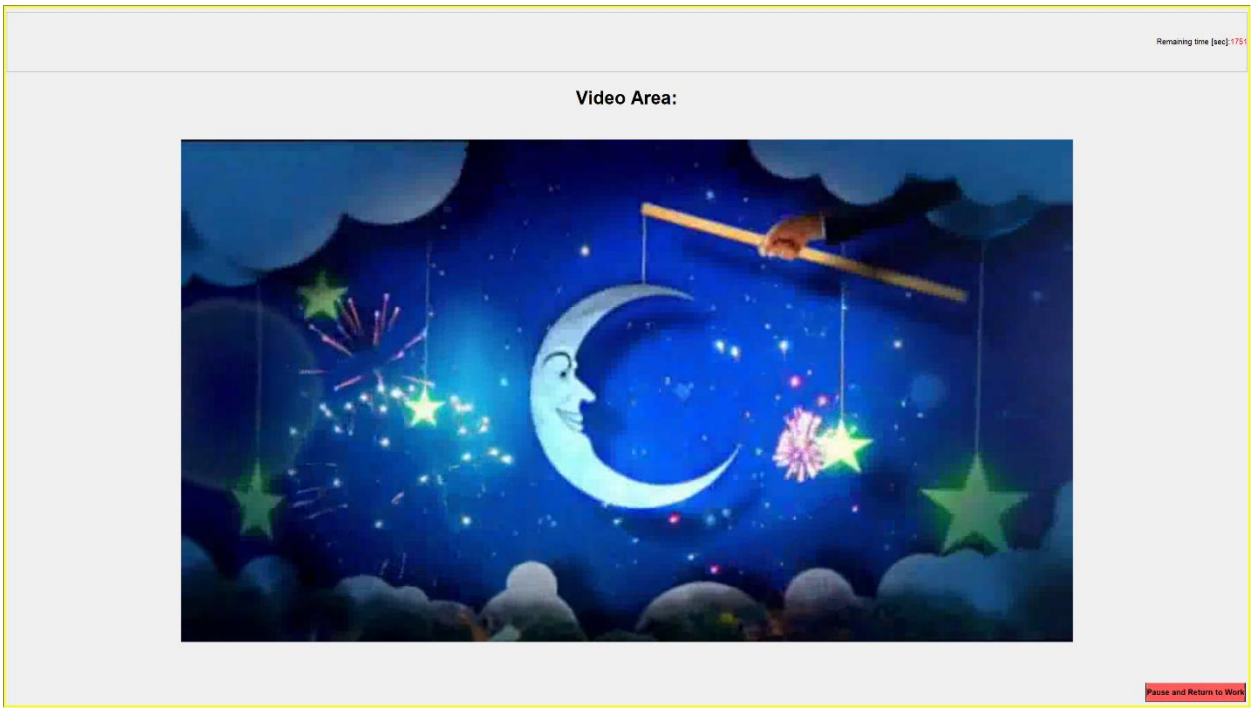
Any questions?

Screen Image of the addition task in Part 2:

[1. Work site:]

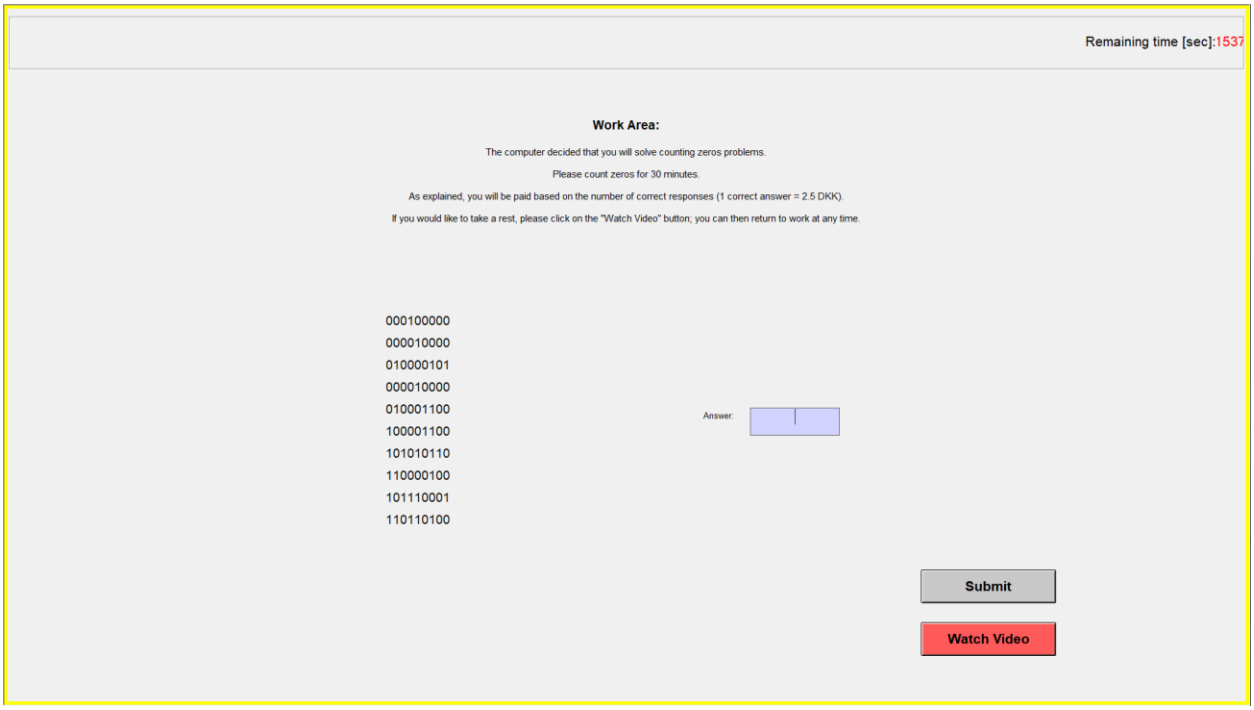


[2. Video site:]

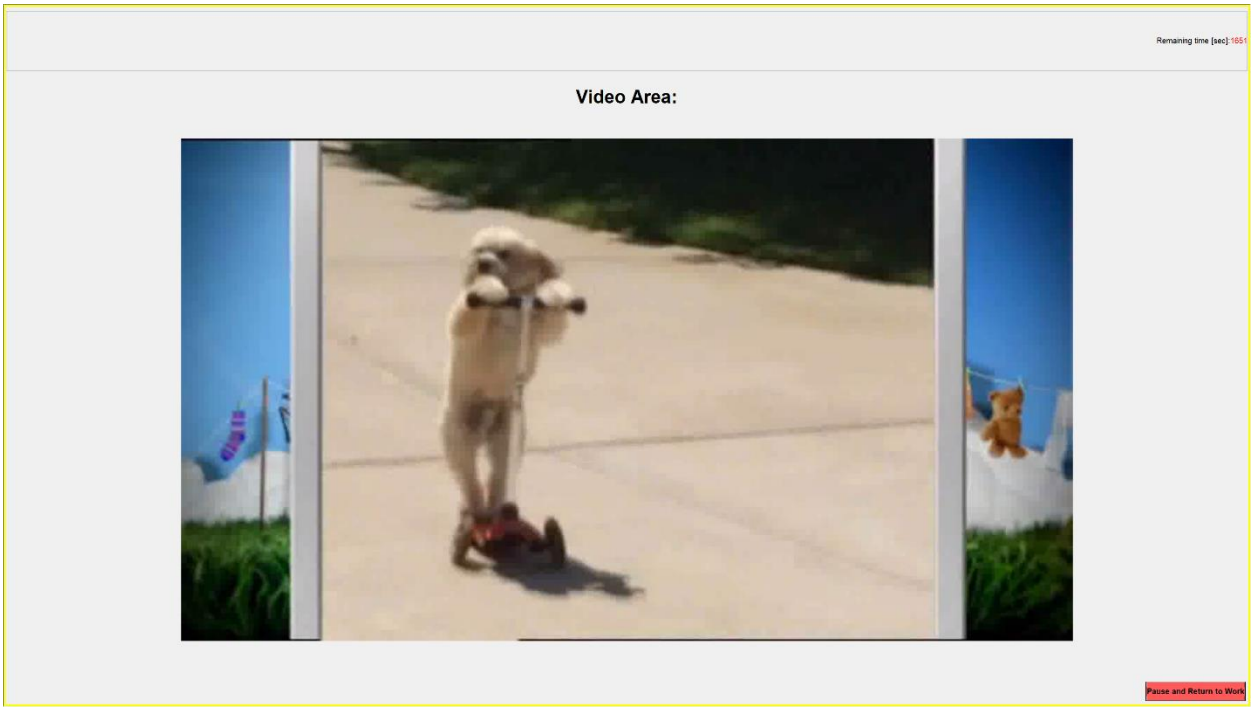


Screen Image of the counting zeros task in Part 2:

[1. Work area:]



[2. Video area:]



(2) The Team (Full feedback) treatment

Welcome

You are now taking part in a decision-making experiment. You will be able to earn money depending on your decisions and also on the decisions of other participants in some tasks. Please read the following instructions carefully.

During the experiment you are not allowed to communicate with other participants. If you have questions, please raise your hand. One of us will come to answer your question.

This experiment has two parts. In Part 1 you engage in two tasks, each of which lasts three minutes. Part 2 consists of a voting stage and a work stage.

Part 1

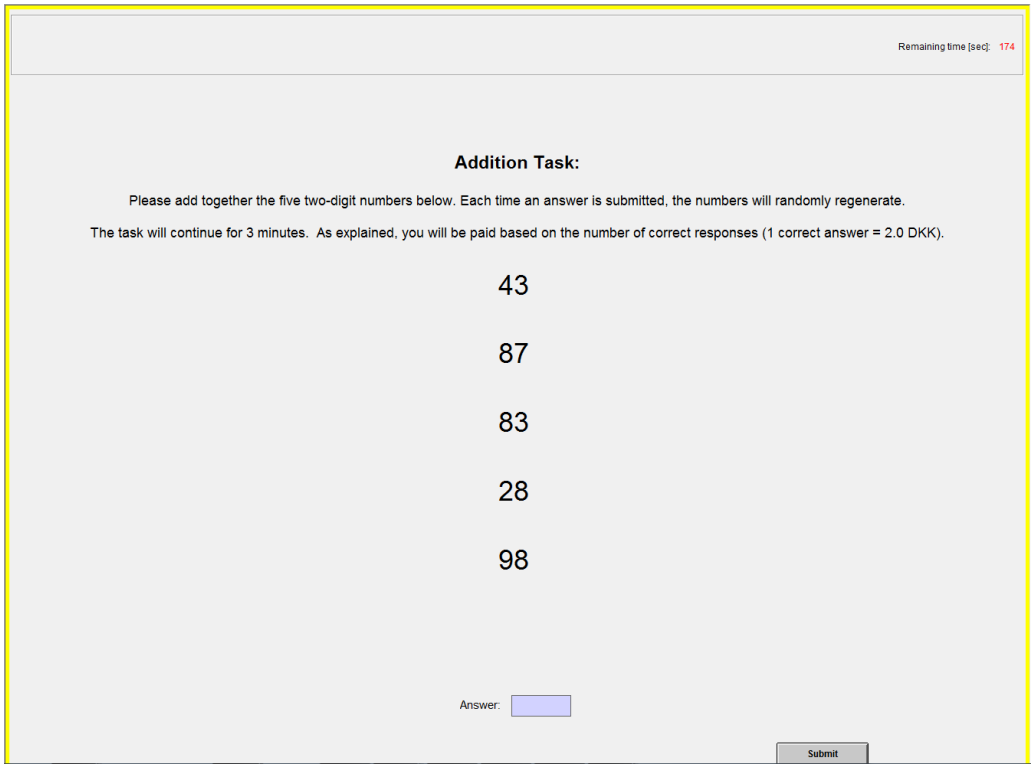
In Part 1, you will undertake two tasks: (a) **an addition task** for three minutes, and (b) **a counting zeros task** for three minutes. The order of tasks (which task you will undertake first) is randomly selected for each participant. We will now explain each task.

(a) The addition task:

You will be asked to answer addition questions. In each task, you will add five two-digit numbers. The more questions you answer correctly, the more money you earn.

Once you add the five numbers and fill the blank with the sum, please click the "Submit" button. Then, you will move on to the next question and five two-digit numbers will be randomly generated. We note that if you do not want to answer a question, you can skip it by simply pressing the Submit button without writing an answer. You will not be given feedback as to whether your calculation is correct or not for each question. However, at the end of the 3 minutes, you will be informed of the number of questions you answered correctly. You will be paid based only on your own performance. You will receive 2 Danish Kroner per correct answer. Any questions?

Screen Image:



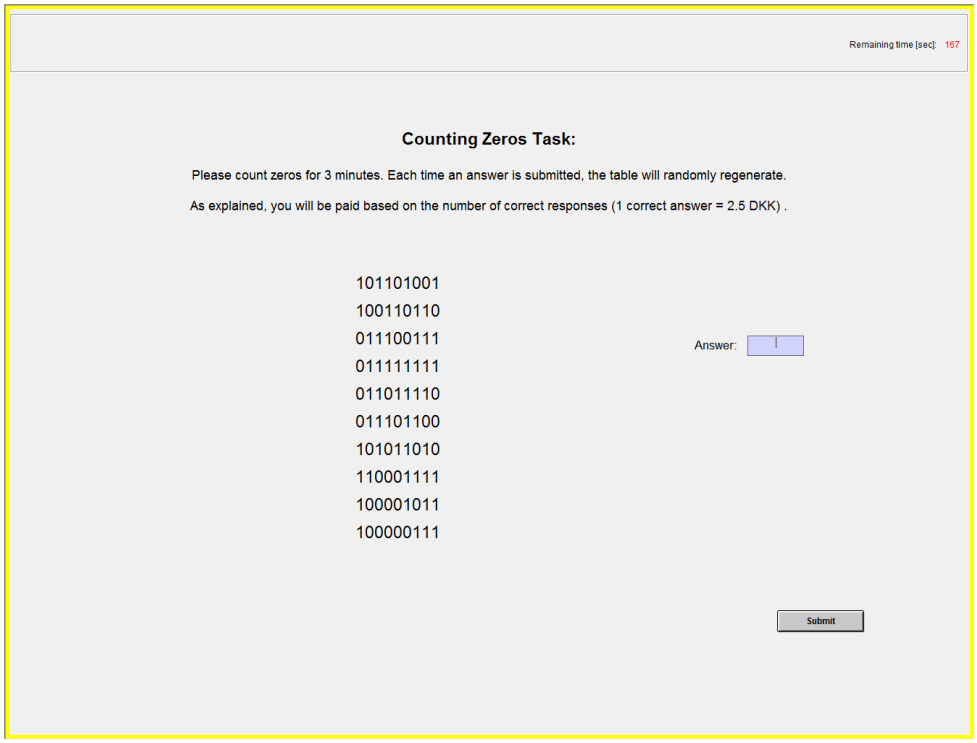
Note: Numbers shown in the screen shot are for illustration only.

(b) The counting zeros task:

You will be asked to count the number of zeros appearing on your screen. The more counting tasks you solve correctly, the more money you earn. Each question has a table which consists of 90 randomly ordered 0s and 1s as in the following screen image. Your task is to count the number of 0s in the table.

Once you count the number of 0s and fill the blank with the number, please click the "Submit" button. Then, you will move on to the next question and a new table with 0s and 1s will be randomly generated. We note that if you do not want to answer a question, you can skip it by simply pressing the Submit button without writing an answer. As in the addition task, you will not be given feedback as to whether your answer is correct or not for each question. However, at the end of the 3 minutes, you will be informed of the number of questions you answered correctly. You will be paid based only on your own performance. Specifically, you will receive 2.5 Danish Kroner per correct answer. Any questions? Once you finish undertaking the two tasks, you will move on to Part 2.

Screen Image:



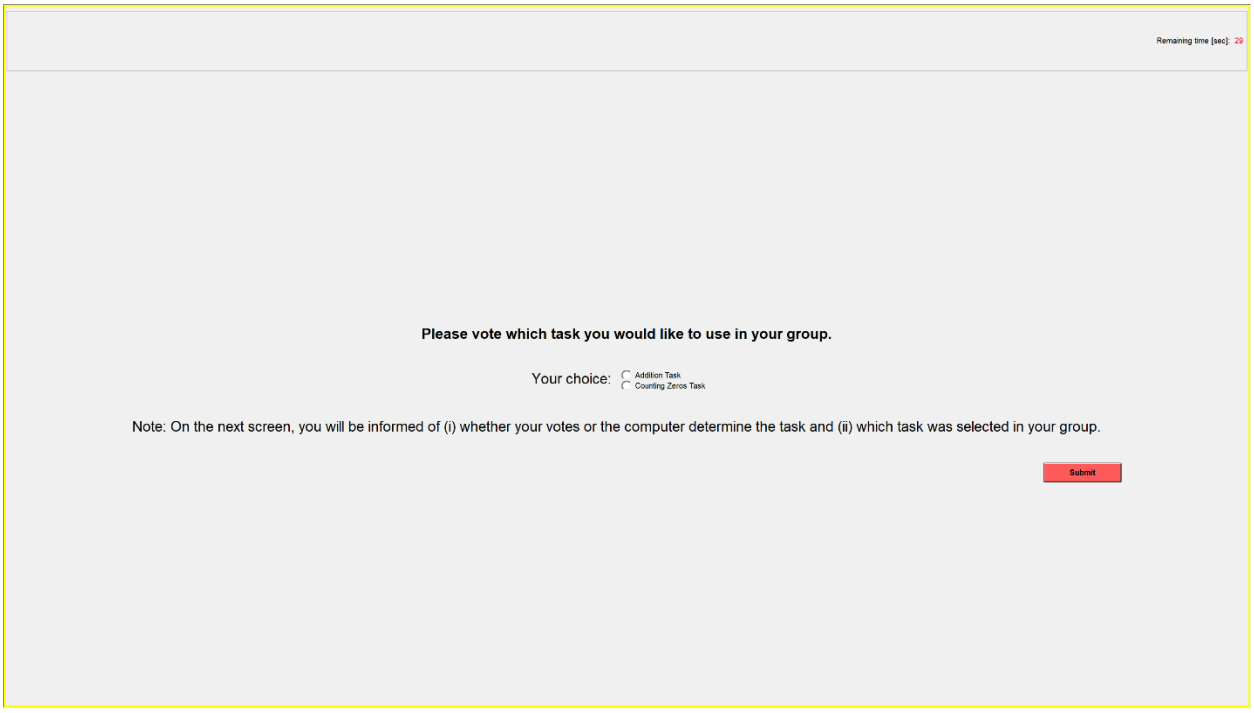
Note: Numbers shown in the screen shot are for illustration only.

Part 2

At the onset of Part 2, each participant is randomly assigned to **a group of three**. Group assignment is completely random. Neither your performance in Part 1 nor others' performance in that part affects the group assignment process. In this part, you will perform one of the two tasks (the addition or counting zeros task) for a longer duration – namely 30 minutes. After the 30 minutes, the total number of correctly answered questions (and accordingly the resulting total earnings) are calculated for each group. Each group member receives **one-third of their group's total earnings**. This means that for each correct response in the counting zeroes task, you earn 2.5/3 = 0.83 Danish Krone, and each of the other two members of your group also earn 0.83 Krone. For each correct response in the addition task, you earn 2/3 = 0.67 Danish Krone, and each of the other two members of your group also earn 0.67 Krone.

The procedure to select a task in each group consists of two steps (Step 1 and Step 2 below). We will now explain how Part 2 proceeds.

Step 1: Each member, including you, votes on which task they want to use in their group.



Step 2: After all individuals complete the voting decision, the computer **randomly (i.e., with a probability of 50%) decides whether or not their votes are considered in each group**.

- If the computer decides to consider their votes in a group, **whichever task receives a majority of votes (2 or 3 votes)** will be used in that group.

- If the computer decides not to consider a group's votes, it will **randomly choose which task to be used in that group** regardless of how the individuals in the group voted.

At the end of step 2, you will be informed whether or not the votes were considered in your group and which task was selected.

You will be informed of which task received a majority of votes, even if the computer randomly decides not to consider your votes.

Step 3: You will undertake the selected task **for 30 minutes**. Your group will receive 2 Danish Kroner per correct answer if you undertake the addition task. Your group will receive 2.5 Danish Kroner per correct answer if you undertake the counting zeros task.

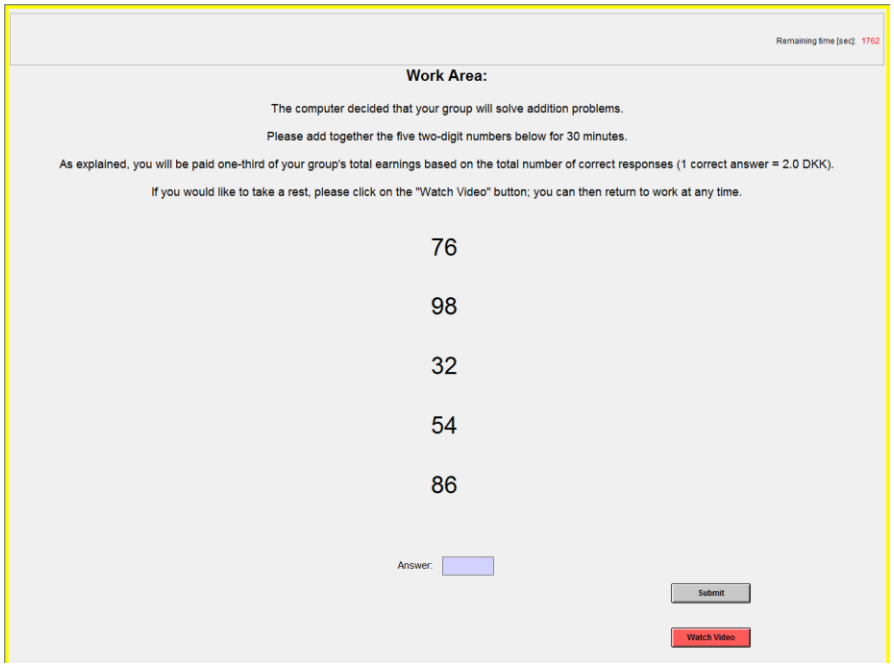
In the work stage (Step 3), you can take a rest and watch an entertaining video clip (see the screen image below) whenever you like. To hear the audio track of the video, please use the headphone on your desk. There is a button named “Watch Video.” When you click this button, you will move on to another screen that contains a video clip. You can pause the video and return to the work site by clicking the “Pause and return to work” button freely. You will earn a small amount of money by watching the video clip also instead of working on the selected task. In particular, you will receive 1 Danish Krone per minute of watching the video clip. Note that you cannot watch the video while working on a task. Note also that the earnings from watching the video clip will be added to your own earnings (will not be shared with the other members of your group).

Please remain silent at all times during the experiment.

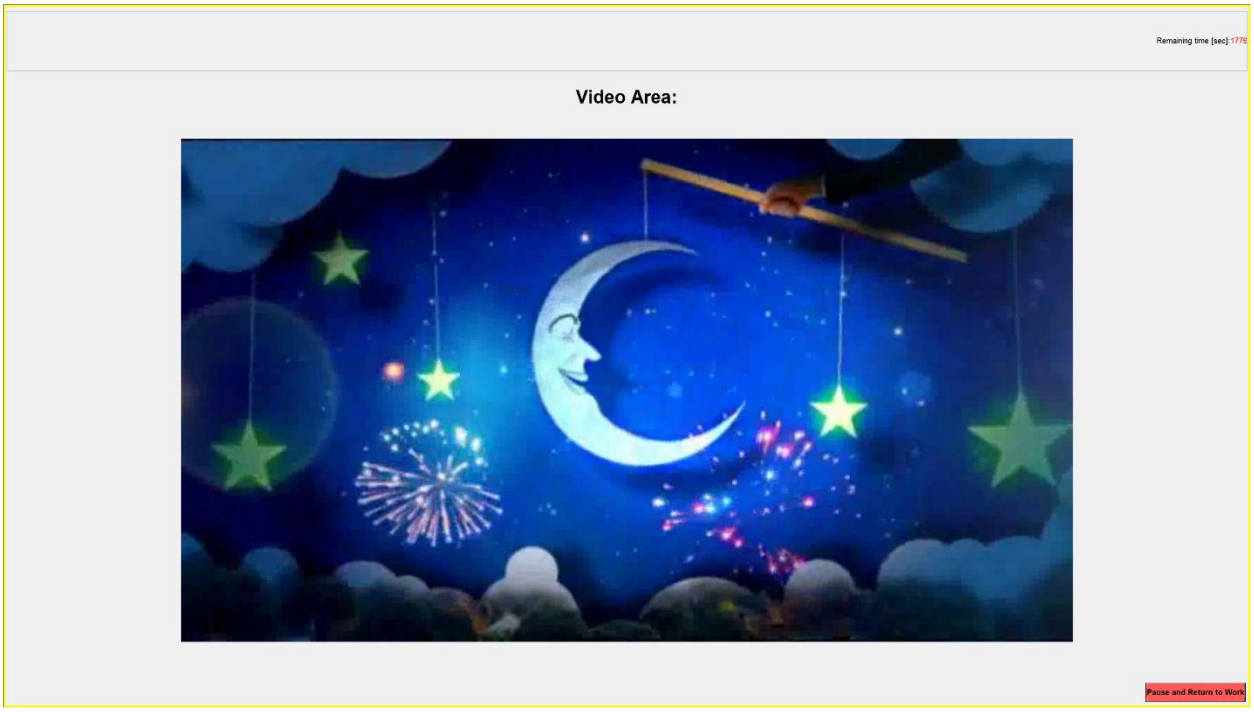
Any questions?

Screen Image of the addition task in Part 2:

[1. Work site:]

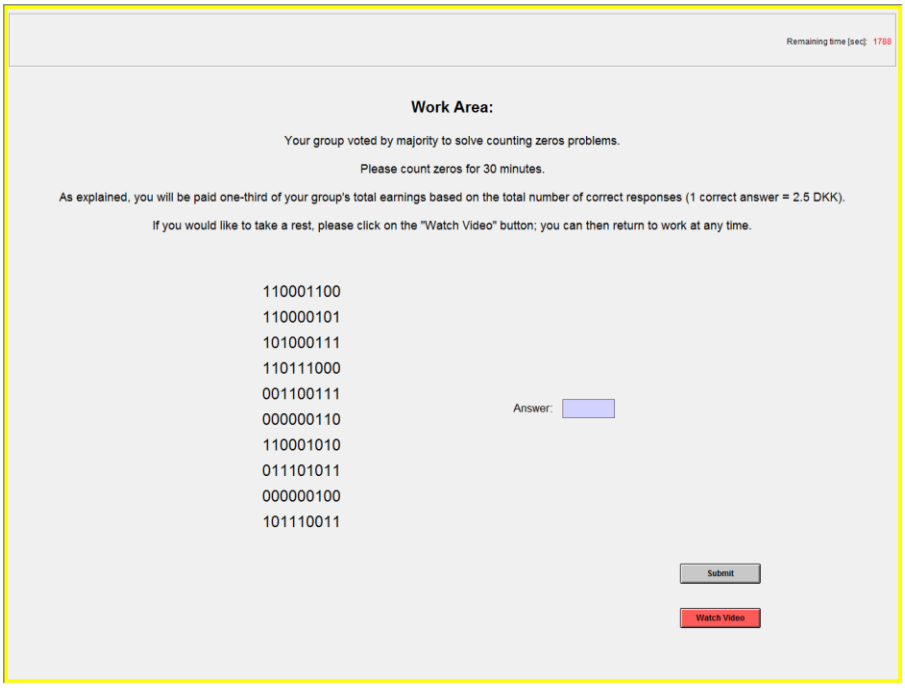


[2. Video site:]

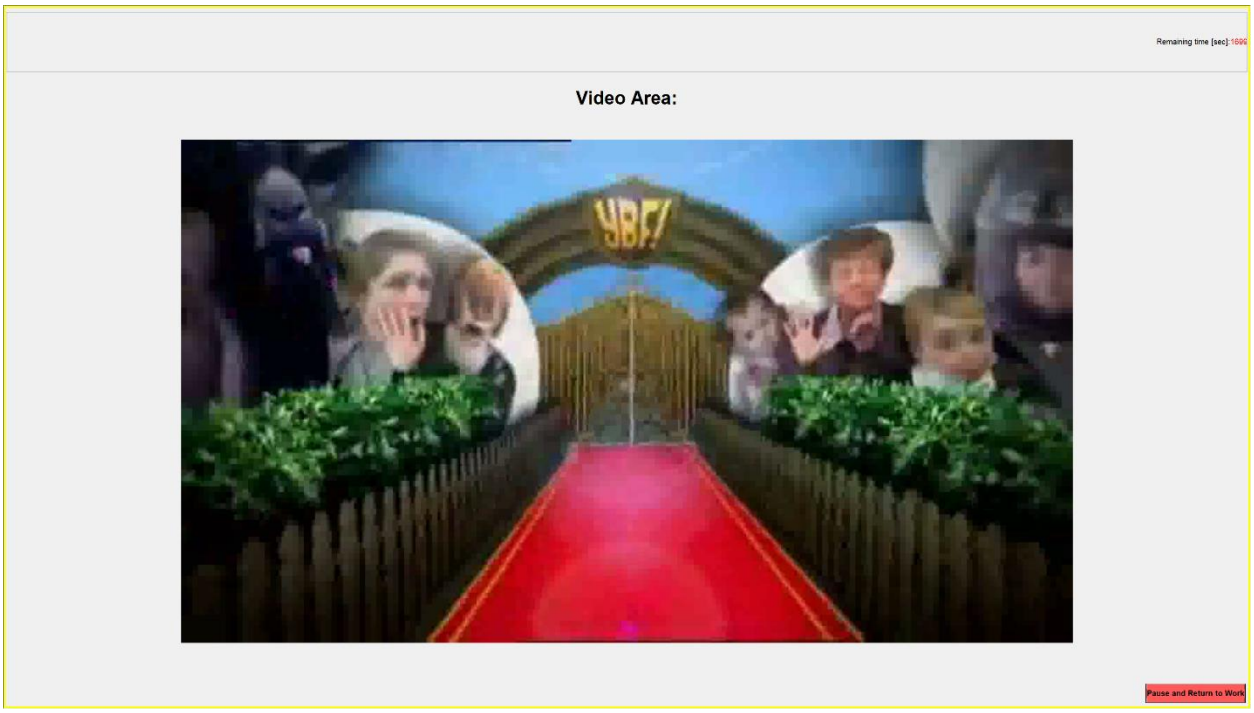


Screen Image of the counting zeros task in Part 2:

[1. Work area:]

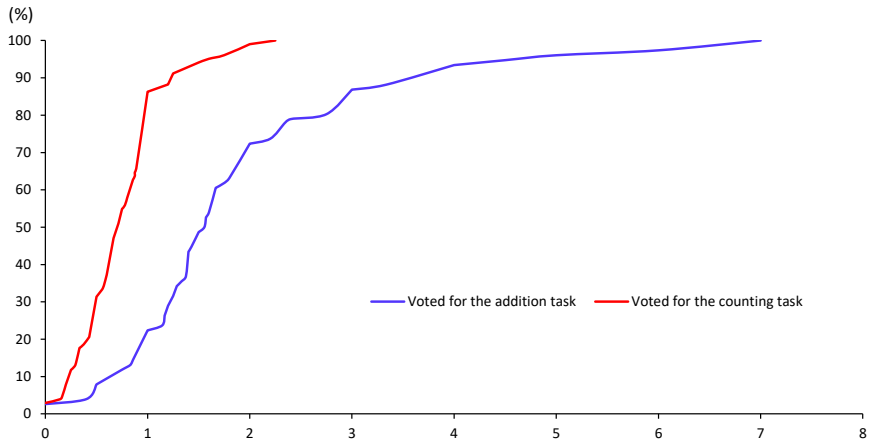


[2. Video area:]

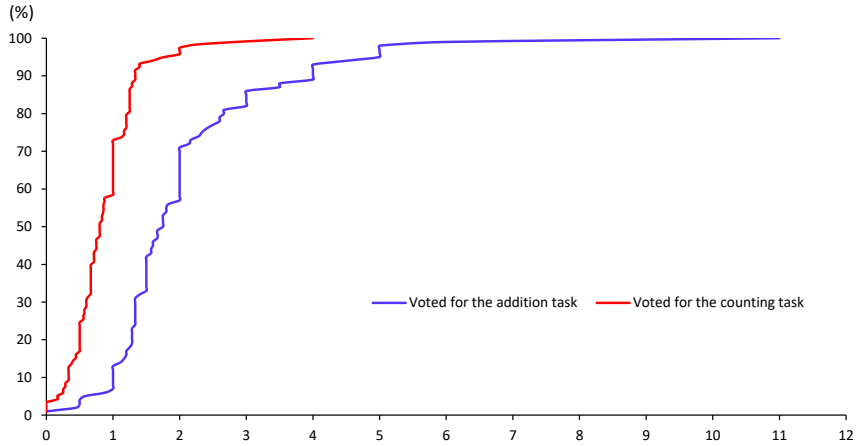

Remaining time [sec]: 1899
Video Area:
УВГ!
Pause and Return to Work

Appendix C: Additional Figures and Tables

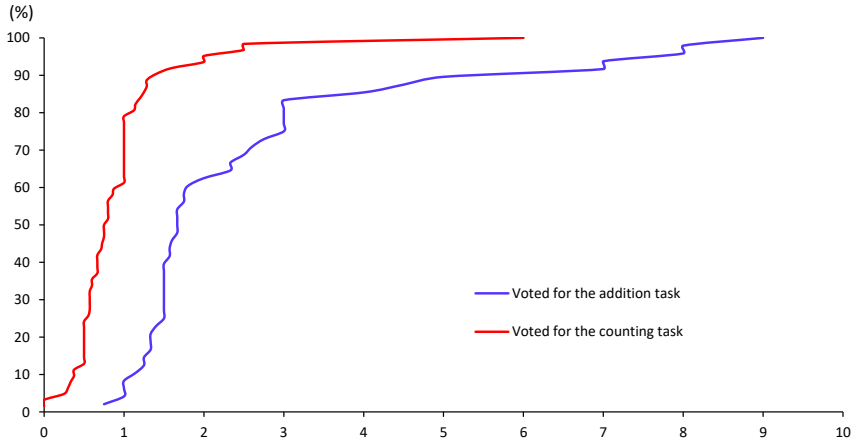
Figure C.1. Cumulative Distributions of Subjects' Relative Performance in the Two Tasks in Part 1



(A) Individual treatment

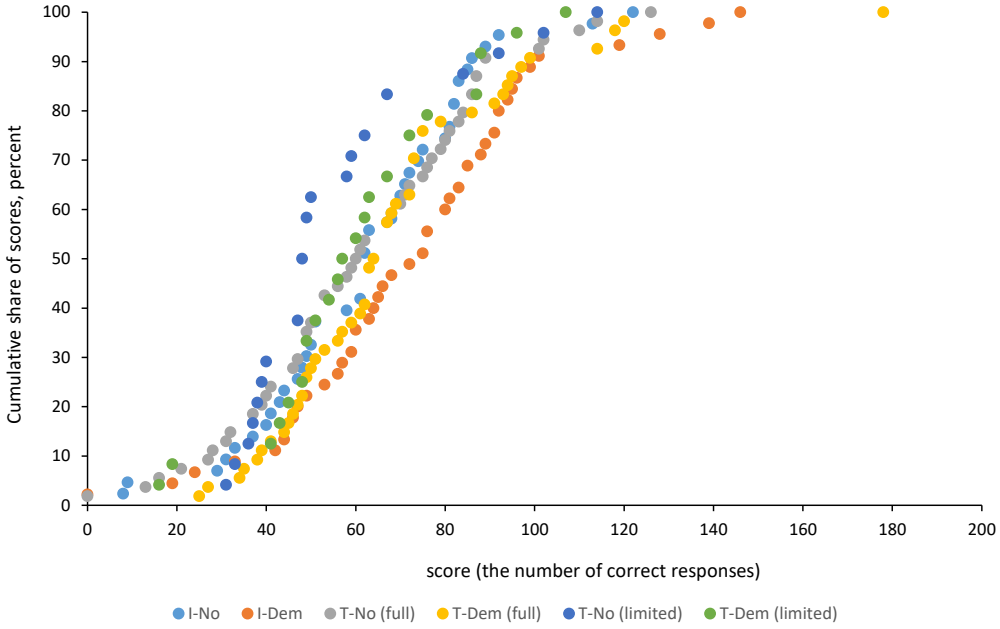


(B) Team (Full feedback) treatment

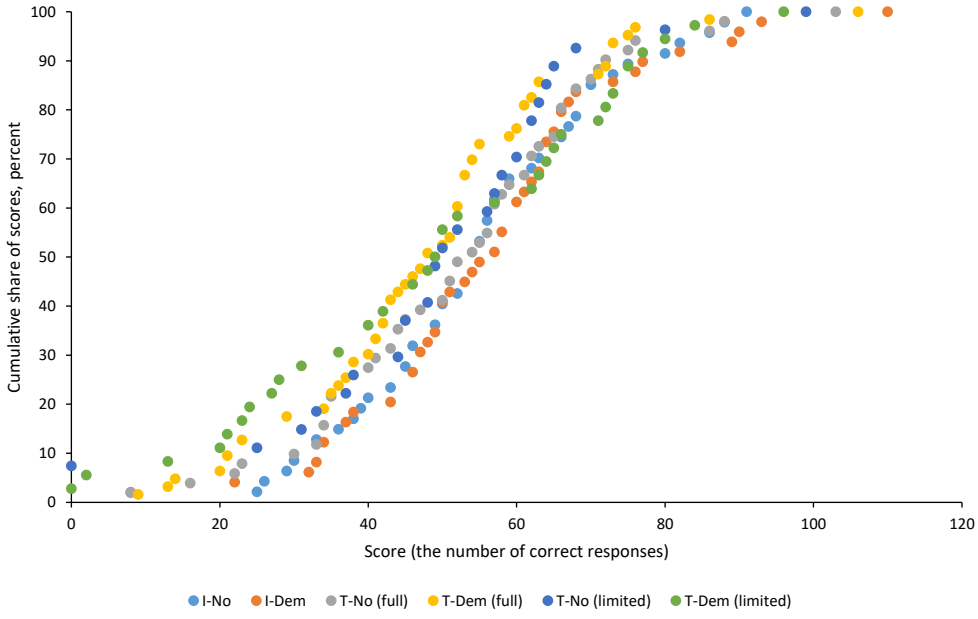


(C) Team (Limited feedback) treatment

Figure C.2. Cumulative Distributions of Subjects' Scores in Part 2



(A) The addition task



(B) The counting task

Notes: The further to the right in the diagram the dots for a particular treatment are located, the better did the subjects in the treatment perform. I-Dem [orange dots] and T-No (Limited) [dark blue dots] regimes stand out as the, respectively, most and least successful groups in the addition task.

Table C.1. Average Performance in Part 1 by the Voting

	Individual			Team (Full feedback)			Team (Limited feedback)		
	(a) No democ. I-No	(b) Democracy I-Dem	H_0: (a) = (b)	(c) No democ. T-No (Full)	(d) Democracy T-Dem (Full)	H_0: (c) = (d)	(e) No democ. T-No (Limited)	(f) Democracy T-Dem (Limited)	H_0: (e) = (f)
I. Number of correct responses in the addition task									
(i) Supporters of the addition task	5.500	6.288	.337	7.292	6.474	.128	7.182	6.294	.310
(ii) Supporters of the counting task	4.483	n.a.	n.a.	4.267	5.313	.088*	3.923	4.714	.291
II. Number of correct responses in the counting task									
(iii) Supporters of the addition task	3.667	n.a.	n.a.	4.160	2.733	.019**	4.000	5.000	.183
(iii) Supporters of the counting task	5.923	5.571	.513	5.462	4.899	.237	5.867	5.704	.709

Notes: *p*-values (two-sided) for the differences between the democratic and non-democratic conditions were calculated based on subject random effect linear regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table C.2. Average Attempts in Part 2 by Task Preference (supplementing Table 4 of the paper)

	Individual					Team (Full feedback)					Team (Limited feedback)				
	No democracy I-No		Democracy I-Dem		*p* (2-sided) for H_0:	No democracy T-No (Full)		Democracy T-Dem (Full)		*p* (2-sided) for H_0:	No democracy T-No (Limited)		Democracy T-Dem (Limited)		*p* (2-sided) for H_0:
	fraction	(a) attempt	fraction	(b) attempt	(a) = (b)	fraction	(a) attempt	fraction	(b) attempt	(a) = (b)	fraction	(a) attempt	fraction	(b) attempt	(a) = (b)
I. Addition task															
(i) Supporter of addition task	33.6%	76.6	100.0%	85.0	.308	44.4%	80.2	70.4%	85.6	.407	45.8%	75.7	70.8%	73.9	.759
(ii) Supporter of counting task	66.4%	70.3		n.a.	n.a.	55.6%	66.8	29.6%	60.9	.394	54.2%	55.7	29.2%	62.4	.586
Number of subjects	43		45			54		54			24		24		
II. Counting task															
(i) Supporter of addition task	44.7%	66.3		n.a.	n.a.	49.0%	59.6	23.8%	130.1[#]	.316	44.4%	62.4	25.0%	58.9	.638
(ii) Supporter of counting task	55.3%	65.1	100.0%	67.3	.606	51.0%	72.8	76.2%	62.1	0.000***	55.6%	63.8	75.0%	58.9	0.028**
Number of subjects	47		49			51		63			27		36		

Note: *p*-values (two-sided) were calculated based on subject random effect linear regressions with bootstrapped standard errors, based on 500 replications, clustered by session ID. * $p < .1$; ** $p < .05$; *** $p < .01$. [#] As discussed in Table 3, there was an outlier who skipped a large number of questions.

Table C.3. Free Riding in Teams by Task Preference (supplementing Table 5 of the paper)

Dependent variable: Video watching time (seconds)

	I. Addition Task		II. Counting task	
	Assigned task = chosen task	Assigned task ≠ chosen task	Assigned task = chosen task	Assigned task ≠ chosen task
Independent variable:	(1)	(2)	(3)	(4)
a. Team (Full feedback) dummy	-21.620 (173.728)	529.685** (225.718)	194.272 (162.771)	186.614 (169.135)
b. Team (Limited feedback) dummy	-59.755 (172.506)	273.023 (214.460)	285.246** (137.734)	573.955* (332.550)
[Democracy effect:]				
Individual treatment dummy × Democracy dummy	-210.107 (134.056)	---	197.857 (137.106)	---
Variable a × Democracy dummy	-170.743 (165.904)	-240.801 (424.056)	140.201 (145.791)	127.225 (138.685)
Variable b × Democracy dummy	111.818 (98.923)	318.463* (177.682)	246.300** (109.799)	-172.551 (701.924)
Constant	-10.360 (135.760)	-563.509 (205.650)	-381.833*** (110.183)	-314.106* (166.538)
Prob > Wald χ^2	.002	.065	.001	.331
# of observations	149	95	191	82
# of left-censored observations	90	62	110	41

Notes: Subject random effect tobit regressions with bootstrapped standard errors (based on 500 replications) clustered by session ID. Numbers in parenthesis are standard errors. The reference group is observations in the I-No regime. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table C.4. Dynamics for Effects of Workplace Democracy (supplementing Figure 3 of the paper)

A. Effects of workplace democracy

	I. Addition Task			II. Counting task		
Treatment:	Individual	Team (Full feedback)	Team (Limited feedback)	Individual	Team (Full feedback)	Team (Limited feedback)
Independent variable:	(1)	(2)	(3)	(4)	(5)	(6)
Democracy dummy {= 1(0) if the task was democratically (non-democratically) assigned}	.419*** (.085)	.122 (.174)	-.057 (.173)	-.037 (.138)	-.153 (.099)	-.164* (.095)
Period number = {2, 3, 4, …, 29}[#1]	.010*** (.002)	.003 (.004)	-.000 (.003)	.004 (.002)	.005 (.004)	-.003 (.003)
Democracy dummy × Period number	-.005 (.005)	.010** (.004)	.013* (.007)	.002 (.004)	-.001 (.006)	-.001 (.005)
Constant	1.879*** (.082)	1.963*** (.144)	1.849*** (.144)	1.770*** (.075)	1.615*** (.091)	1.796*** (.060)
Prob > Wald χ^2	.0000	.0000	.0063	.0000	.0648	.0000
# of observations	2,464	3,024	1,344	2,688	3,192	1,680
# of left-censored observations	178	249	133	245	447	244

B. Information effects of voting in the Team treatments[#2]

	I. Addition Task		II. Counting task	
Treatment:	No democracy	Democracy	No democracy	Democracy
Independent variable:	(7)	(8)	(9)	(10)
Team (Full feedback) dummy {= 1 for the Team (Full feedback) treatment; 0 otherwise}	.113 (.187)	.292 (.180)	-.179* (.098)	-.171 (.137)
Period number = {2, 3, 4, …, 29}[#1]	-.000 (.003)	.012 (.006)	-.003 (.004)	-.004*** (.002)
Team (Full feedback) dummy × Period number	.003 (.005)	.000 (.007)	.008 (.005)	.008** (.004)
Constant	1.849*** (.131)	1.792*** (.164)	1.798*** (.057)	1.629*** (.106)
Prob > Wald χ^2	.5355	.0000	.2819	.0125
# of observations	2,184	2,184	2,128	2,744
# of left-censored observations	227	155	244	447

Notes: Subject random effect tobit regressions with bootstrapped standard errors (based on 500 replications) clustered by session ID. Numbers in parenthesis are standard errors. The dependent variable is subject *i*'s score for a given minute. The reference group in columns (1) to (6) [columns (7) to (10)] is observations in the non-democratic condition in a given treatment [observations in the Team (Limited feedback) treatment]. Observations from the 2nd minute to 29th minute in Part 2 were used because, as explained in the paper, subjects exhibited strong learning during the first minute of the 30-minute task-solving phase in Part 2, and a sprint to the finish line toward the end in the last one minute. [#1] Period *k* indicates the one-minute span starting from the *k*-1th minute in the task-solving phase. [#2] Data from the two Team treatments were used. * $p < .1$; ** $p < .05$; *** $p < .01$.

Table C.5. Dynamics for Free Riding in Teams by Task Preference (supplementing Figure 4 of the paper)

A. Score in the addition task

	No democracy		Democracy	
Corresponding panel in Figure 4:	panel A.a(i)	panel A.a(ii)	panel A.b(i)	panel A.b(ii)
Independent variable:	(1)	(2)	(3)	(4)
Team (Full feedback) dummy {= 1 for the Team (Full feedback) treatment; 0 otherwise}	.286	-.154	-.080	-.523***
	(.243)	(.177)	(.142)	(.178)
Team (Limited feedback) dummy {= 1 for the Team (Limited feedback) treatment; 0 otherwise}	.044	-.178	-.439**	-.654*
	(.263)	(.112)	(.185)	(.384)
[Trend:]				
(a) Individual treatment dummy × Period number[#1]	.005	.012***	.005	.005
	(.005)	(.003)	(.004)	(.004)
(b) Team (Full feedback) dummy × Period number	-.006	.010**	.018***	.001
	(.004)	(.005)	(.003)	(.006)
(c) Team (Limited feedback) dummy × Period number	.001	-.001	.018***	-.002
	(.004)	(.005)	(.004)	(.010)
Constant	2.125***	1.761***	2.296***	2.297***
	(.187)	(.076)	(.124)	(.122)
Prob > Wald χ^2	.5305	.0000	.0000	.0270
# of observations	1,372	2,016	2,800	1,904
# of left-censored observations	92	223	171	164

B. Score in the counting task

	No democracy		Democracy	
Corresponding panel in Figure 4:	panel B.a(i)	panel B.a(ii)	panel B.b(i)	panel B.b(ii)
Independent variable:	(1)	(2)	(3)	(4)
Team (Full feedback) dummy {= 1 for the Team (Full feedback) treatment; 0 otherwise}	-.024	-.272	-.148	-.638***
	(.152)	(.194)	(.143)	(.225)
Team (Limited feedback) dummy {= 1 for the Team (Limited feedback) treatment; 0 otherwise}	.086	-.068	-.095	-.096
	(.132)	(.179)	(.169)	(.153)
[Trend:]				
(a) Individual treatment dummy × Period number[#1]	.001	.007	.006***	.006***
	(.003)	(.006)	(.002)	(.002)
(b) Team (Full feedback) dummy × Period number	.006	.004	.001	.013**
	(.006)	(.004)	(.004)	(.006)
(c) Team (Limited feedback) dummy × Period number	-.004	-.002	-.004*	-.007
	(.005)	(.005)	(.002)	(.010)
Constant	1.811***	1.719***	1.726***	1.728***
	(.106)	(.108)	(.111)	(.117)
Prob > Wald χ^2	.8939	.1001	.0011	.0024
# of observations	1,876	1,568	3,472	2,016
# of left-censored observations	154	193	473	258

Notes: Subject random effect tobit regressions with bootstrapped standard errors (based on 500 replications) clustered by session ID. Numbers in parenthesis are standard errors. The dependent variable is subject *i*'s score for a given minute. The reference group is observations in the Individual treatment for a given comparison in period 2. Observations from the 2nd minute to 29th minutes in Part 2 were used [#1] Period number (= {2, 3, 4, …, 29}). Period *k* indicates the one-minute span starting from the *k*-1th minute in the task-solving phase. * $p < .1$; ** $p < .05$; *** $p < .01$.