

Schildberg-Hörisch, Hannah; Schwarz, Marco A.; Trieu, Chi; Willrodt, Jana

Working Paper

Perceived Fairness and Consequences of Affirmative Action Policies

CESifo Working Paper, No. 10198

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Schildberg-Hörisch, Hannah; Schwarz, Marco A.; Trieu, Chi; Willrodt, Jana (2022) : Perceived Fairness and Consequences of Affirmative Action Policies, CESifo Working Paper, No. 10198, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/271842>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Perceived Fairness and Consequences of Affirmative Action Policies

Hannah Schildberg-Hörisch, Marco A. Schwarz, Chi Trieu, Jana Willrodt

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Perceived Fairness and Consequences of Affirmative Action Policies

Abstract

Debates about affirmative action often revolve around fairness. Accordingly, we document substantial heterogeneity in the fairness perception of various affirmative action policies. But do these differences translate into different consequences? In a laboratory experiment, we study three different quota rules in tournaments that favor individuals whose performance is low, either due to discrimination, low productivity, or choice of a short working time. Affirmative action favoring discriminated individuals is perceived as fairest, followed by that targeting individuals with a short working time, while favoring low productivity individuals is not perceived as fairer than an absence of affirmative action. Higher fairness perceptions coincide with a higher willingness to compete and less retaliation against winners, underlining that fairness perceptions matter for the consequences of affirmative action. No policy harms overall productivity or post-competition teamwork, but affirmative action may reduce the average output of tournament winners.

JEL-Codes: C910, D020, D630.

Keywords: affirmative action, fairness ideals, experiment, tournament, real effort.

*Hannah Schildberg-Hörisch**
Düsseldorf Institute for Competition Economics
(DICE), Düsseldorf / Germany
schildberg-hoerisch@dice.hhu.de

Marco A. Schwarz
Düsseldorf Institute for Competition
Economics (DICE), Düsseldorf / Germany
schwarz@dice.hhu.de

Chi Trieu
Düsseldorf Institute for Competition Economics
(DICE), Düsseldorf / Germany
trieu@dice.hhu.de

Jana Willrodt
Düsseldorf Institute for Competition
Economics (DICE), Düsseldorf / Germany
willrodt@dice.hhu.de

*corresponding author

We thank Björn Bartling, Lorenz Götte, Paul Heidhues, Steffen Huck, Muriel Niederle, Hans-Theo Normann, Jan Potters, Gerhard Rieger, Bettina Rockenbach, Daniel Schunk, Marie-Claire Villeval, Roberto Weber, and two anonymous reviewers for useful comments and suggestions. Additionally, we thank participants of ESA 2018 in Berlin, EWEBE 2018 in Tilburg, MBEPS 2018 in Maastricht, the Birthday Symposium for Monika Schnitzer and Klaus Schmidt 2022 in Munich, and seminars in Aachen, Cologne, Frankfurt, Göttingen, at Howard University in Washington, D.C., at MPI for Research on Collective Goods in Bonn, and at DICE for helpful discussions. Funding by the Deutsche Forschungsgemeinschaft (German Research Foundation) through 235577387/GRK1974 is gratefully acknowledged. Our experiment complies with the code of conduct for experiments at the DICE Lab and Bonn-EconLab. According to rules of Heinrich-Heine-University Düsseldorf, study-specific IRB approval was not required for this study because the experiment did not create any harm or distress beyond participants' everyday experiences and because we comply with German laws, particular legislation pertaining to physical integrity, privacy rights, and data protection. Declaration of interest: none.

1 Introduction

Affirmative action is defined as a policy that promotes the opportunities of defined minority groups within a society to give them equal access to that of the privileged majority population (Affirmative Action, Harvard Law School Blog, archived on 18 May 2015). Ever since its introduction, affirmative action has been subject to heated debates (see, e.g., Fish, 2000; Fullinwider, 2011). A common argument professed by its opponents is that under affirmative action decisions such as university admission, hiring, or promotion choices are not purely based on merit. This goes against the ideal of a “fair” policy that should select or reward the best performers. Proponents of affirmative action on the other hand may argue that a truly fair policy should take disadvantages (e.g., due to family background, poverty, race, and gender) into account when evaluating performance to compensate for a lack of opportunities due to discrimination, historical injustice, or the “accident of birth”.

The underlying conflict in this debate is the collision of different fairness ideals originating from different theories of distributive justice (see Konow (2003) for a comprehensive overview). Each professes a distinct view on which factors of their performance individuals should be held accountable for. The three stylized factors that determine an individual’s performance are working time, productivity, and discrimination (Cappelen et al., 2010).

Building on previous work (see, e.g., Cappelen et al., 2010), we focus on four different fairness ideals. From a *libertarian* point of view, individuals should be held accountable for all three factors, making affirmative action unnecessary. *Meritocracy*, however, posits that only personal factors, i.e., working time and productivity, should be considered when assessing someone’s performance, justifying affirmative action that compensates for discrimination. According to *choice egalitarianism*, people should be held accountable only for factors within their control. This means affirmative action policies should offset discrimination and differences in productivity, but not self-chosen working time. Finally, *strict egalitarianism* strives for complete equality, thus calling for policies that counterbalance differences in all three factors.

Despite a growing amount of empirical evidence on the consequences of affirmative action, the perceived fairness of different affirmative action policies and the implications of those fairness perceptions for the effectiveness of affirmative action policies have barely been investigated. In this paper, we provide evidence from a laboratory experiment that implements several affirmative action policies in the

form of quota rules in a tournament setting, and explicitly discuss the association of a broad set of outcomes of those policies with their fairness perception. As a further novelty, the affirmative action policies address all three determinants of performance in a unified framework. One policy favors individuals disadvantaged due to discrimination. In our design, being discriminated against means that output is downgraded. Who is discriminated against is randomly assigned and remains unchanged throughout the experiment. This resembles stable causes of discrimination such as sex, skin color, or race. Similarly, Balafoutas et al. (2016), Calsamiglia et al. (2013), Fallucchi and Quercia (2018), Petters and Schroeder (2020), and Ip et al. (2020) analyze affirmative action policies that compensate for differences in a randomly assigned, exogenously given characteristic. Another affirmative action policy favors those disadvantaged by low productivity (a personal, but out-of-control factor) in the experimental task in which practicing does not improve performance. A third policy favors individuals whose performance is lower because they have chosen a shorter working time (an in-control factor). To our knowledge, we are the first to investigate affirmative action policies that compensate for differences in working time and add to the scarce evidence on affirmative action favoring individuals with low productivity (Ip et al., 2020).

Similar policies exist outside the laboratory: the Council Directive 97/81/EC of the European Union states that part-time employees may not be treated less favorably than full-time employees (see Council of European Union, 1997). This includes their equal access to promotions (although their overall performance in terms of output is typically lower). Moreover, in several countries, among them Germany, time taken off work due to parental leave (a form of voluntarily reducing working time) is counted as working time towards automatic pay raises for employees in the public sector. Several policies are designed to compensate for worse performance due to low productivity. For example, students with dyslexia or physical restraints such as typist's cramps or poor eyesight can get some extra time in exams (see, e.g., Disability Rights Commission, 2007). While we use quotas as a form to implement affirmative action policies in the laboratory, the above compensation policies are similar in spirit as they favor the outcomes of specific groups to compensate for some initial disadvantage.

An advantage of our approach to study affirmative action based on the separate determinants of performance (productivity, working time, and discrimination) is that fairness judgments regarding these affirmative action policies can be unequivocally attributed to holding different fairness ideals since these policies specifically

state the reason for a favorable treatment of the respective target group. Criteria for affirmative action commonly used in laboratory and field experiments such as gender or ethnicity, on the other hand, may invoke different perceptions of the reasons for the affirmative action policy. Take, for example, the most widely studied example of a gender quota: if a woman's performance is not among the best, some might perceive this as being the result of discrimination, while others may attribute her performance to low productivity, or a choice of working part-time. Usually, we cannot observe which of the three perceptions (or a mixture thereof) is invoked. However, this seems crucial for enhancing our understanding of the fairness perception of and reaction to quotas.

We address three related research questions. Do the consequences of affirmative action policies depend on which of the three factors (discrimination, low productivity, or short working time) they compensate for? Are the three affirmative action policies perceived as differently fair? And do the fairness perceptions of these policies predict their consequences? Answering these questions is key to understanding the approval and the implications of different affirmative action policies.

To address these questions, we elicit individual fairness perceptions for all affirmative action policies and consider a broad set of outcomes. First, we look at the immediate consequences of affirmative action within tournaments that are based on performance in a real effort task, namely willingness to compete and efficiency. We then explore potential spillover effects on post-competition outcomes when affirmative action is no longer in place, focusing on teamwork and spiteful behavior targeting those who benefited from affirmative action ("retaliation"). To guide our analysis, we introduce a theoretical framework. In particular, we allow subjects to have different fairness types (libertarian, meritocratic, choice egalitarian, and strict egalitarian) that are reflected in a potential fairness utility loss term in their utility function in addition to monetary payments.

We find that none of the affirmative action policies we investigate is considered less fair than no affirmative action, documenting wide acceptance for affirmative action from a normative fairness point of view. Affirmative action targeting discriminated individuals is perceived as fairest, followed by a policy in favor of individuals choosing a short working time, while affirmative action targeting individuals with low productivity and no affirmative action are perceived as equally and least fair. Moreover, self-serving bias is a significant predictor of fairness perceptions: individuals that potentially benefit from a specific affirmative action policy consider it as fairer than non-favored ones.

Importantly, none of the three affirmative action policies under study harms overall willingness to compete or productivity and only the affirmative action policy that favors subjects with short working time reduces the overall output of tournament winners. While our results on comparisons between the various policies are causal, our findings related to fairness perceptions are of correlational nature. The heterogeneity in fairness perceptions is associated with more specific consequences of these policies. Non-favored subjects are more willing to enter a tournament with a quota rule if they *personally* perceive the quota rule as fair. Additionally, favored subjects are more likely to compete under a quota rule instead of working under a piece rate payment the fairer the quota rule is *generally* perceived to be. Regarding potential spillover effects, we observe no differences in overall post-competition teamwork or retaliation between the various policies. However, winners of a tournament who were favored by the quotas are retaliated against by non-favored individuals who perceive a quota rule as less fair. In sum, the fairness perception of an affirmative action policy seems to shape its consequences. This is an important insight for the successful implementation of such policies if they are politically desired. For example, providing evidence on existing discrimination against a favored group may help to increase the acceptance of an affirmative action policy and, in turn, positively impact its consequences.

Our paper contributes to a growing literature researching the effectiveness of affirmative action in various domains (see, e.g., Balafoutas and Sutter (2012), Baltrunaite et al. (2014), Beaman et al. (2009), Beaurain and Masclet (2016), De Paola et al. (2010), and Niederle et al. (2013) for studies on gender quotas in labor market or political settings; Bagde et al. (2016), Banerjee et al. (2018, 2020), Bertrand et al. (2010), and Jensenius (2015) for results on affirmative action relating to caste membership; or Schotter and Weigelt (1992) and Calsamiglia et al. (2013) for affirmative action based on an exogenously assigned disadvantage in a lab setting). Many studies find overall positive effects. For example, there is evidence that such policies increase women's willingness to enter competitions without discouraging men (Balafoutas and Sutter, 2012; Balafoutas et al., 2016; Niederle et al., 2013; Ibanez and Riener, 2018). Estevan et al. (2019) show that awarding disadvantaged applicants bonus points in university entrance exams effectively raises university admission of that group without harmful effects on study efforts. Regarding post-competition outcomes, Kölle (2017) finds that gender quotas neither harm effort provision within teams nor the willingness to work in teams.

However, there is also evidence of adverse consequences of affirmative action (Fallucchi and Quercia, 2018; Girard, 2020; Heilman et al., 1997; Leibbrandt et al., 2017; Leibbrandt and List, 2018). In particular, in Leibbrandt et al. (2017) a gender quota turns women into the target of sabotage, thereby undermining their willingness to compete. Similarly, Fallucchi and Quercia (2018) find that the threat of retaliation reduces competition entry of favored subjects. These seemingly contradictory findings bring up the question under which conditions affirmative action has adverse consequences. Answering it may provide valuable insights into how affirmative action policies that are politically desired can be implemented without causing more harm than good.

The effects of affirmative action policies can depend on whether and how they are justified. In particular, Petters and Schroeder (2020) find that, for randomly assigned quotas, favored individuals' performance is rated worse than that of non-favored individuals with a similar performance. In a gift-exchange game, Ip et al. (2020) show that quotas for female managers decrease workers' effort when subjects are primed towards perceiving a gender skill gap, but not when there is discrimination against women in the manager selection process. In an independent, representative survey with US citizens, approval for gender quotas for leadership positions is high when women are supposed to be discriminated against in the recruitment process, but low otherwise (regardless of whether a gender skill gap exists). These observations lay the ground for jointly studying the perceived fairness of affirmative action policies and their effects on outcomes as we do in this paper. Our approach differs from Ip et al. (2020) and Petters and Schroeder (2020) in several respects: (i) the fact that we explicitly elicit the fairness perception of affirmative action policies, (ii) the nature of the affirmative action policies under study (compensation for different working times; actual as opposed to presumed differences in productivity), and (iii) the context and outcomes we consider.

Finally, our paper moves beyond just identifying fairness views to also show their broader policy implications, focusing on affirmative action. Cappelen et al. (2020) provide a recent literature overview of the implications of fairness concerns for the design of incentive schemes, bargaining, and redistributive policies. A few other papers investigate how fairness concerns affect individuals' willingness to compete. In Buser et al. (2021), subjects may impose an unfair tournament onto others. We complement this paper as subjects in our setting decide for themselves whether to compete. Moreover, Bartling et al. (2009) and Balafoutas et al. (2012) document that individuals with preferences for equality shy away from competition.

The remainder of the paper is organized as follows. [Section 2](#) explains our experimental design, while [Section 3](#) presents a theoretical framework to guide our analysis. [Section 4](#) presents our results on the consequences of affirmative action on willingness to compete, efficiency, teamwork, and retaliation. We conclude in [Section 5](#).

2 Experimental design

Our experiment employs a combination of a within- and between-subject design, in which the four treatments are randomly assigned across subjects: one control treatment without affirmative action and three treatments with different affirmative action policies. Each affirmative action policy favors subjects with a characteristic that dampens their performance—either subjects who are discriminated against, those with low productivity, or those who have chosen a short working time.

Each subject participates in two sessions taking place in consecutive weeks: a preparatory session and the main session. The purpose of the preparatory session is to learn about all subjects' productivity and individually chosen working time to classify them into high and low productivity, long and short working time types, respectively. This determines which subjects will be favored by the respective affirmative action policies. Moreover, we use the information on subjects' type from the preparatory session to form groups in the main session that are balanced with regard to subjects' types. This is crucial to be able to attribute changes in outcomes across treatments to the different affirmative action policies instead of to possible differences in group composition. In the main session, we investigate the fairness perception of the various affirmative action policies and how they are associated with willingness to compete, efficiency, and post-competition teamwork, and retaliation. [Table 1](#) provides an overview of the experimental design.

Preparatory session	Practice round (grid task)
	Measurement of baseline productivity (grid task)
	Questionnaire
	Measurement of choice of working time (grid task)
Main session	Stage 1 Piece rate (grid task)
	Stage 2 Tournament (grid task)
	Stage 3 Choice between piece rate and tournament (grid task)
	Stage 4 Teamwork (slider task)
	Stage 5 Dictator Game
	Measurement of fairness perception

Table 1 Summary of experimental design.

2.1 The real-effort task

We apply the different affirmative action policies to performance in a tournament that is based on the grid task, a real-effort task introduced by Abeler et al. (2011). Subjects work on this task several times under different incentive schemes (see [Table 1](#)). In this task, subjects count the number of zeros in a 10-by-10 table containing 100 digits of randomly distributed zeros and ones (see [Figure 1](#)).

The grid task has several desirable attributes. First, its tediousness and pointlessness induces a positive effort cost for all subjects and minimizes experimenter demand effects (Abeler et al., 2011). Therefore, we are confident that our measure of working time (the time subjects decide to work on this task) actually captures the effort subjects are willing to spend on the task. Second, our data show substantial variation in productivity and chosen working time for this task (see [Figure A.1](#) and [Figure A.2](#) in [Appendix A](#)). Third, the grid task does not require special prior knowledge or skills, performance is easily measurable, and there is little learning possibility (Abeler et al., 2011). In line with that argument, subjects do not need a significantly different amount of time to solve a grid correctly in the first and second half of the five-minute grid task in the preparatory session (Wilcoxon signed rank test, $p = 0.429$). Moreover, as Balafoutas et al. (2016) note and our data confirm, the grid task is gender-neutral. For example, the average number of correctly solved grids in the five minute grid task in the preparatory session is 7.46 for men and 7.69 for women (Mann-Whitney U test, $p = 0.232$). Most importantly, the task allows us to clearly distinguish between the three determinants of performance that matter for perceived distributional fairness according to different fairness ideals (see, e.g., Cappelen et al., 2007).

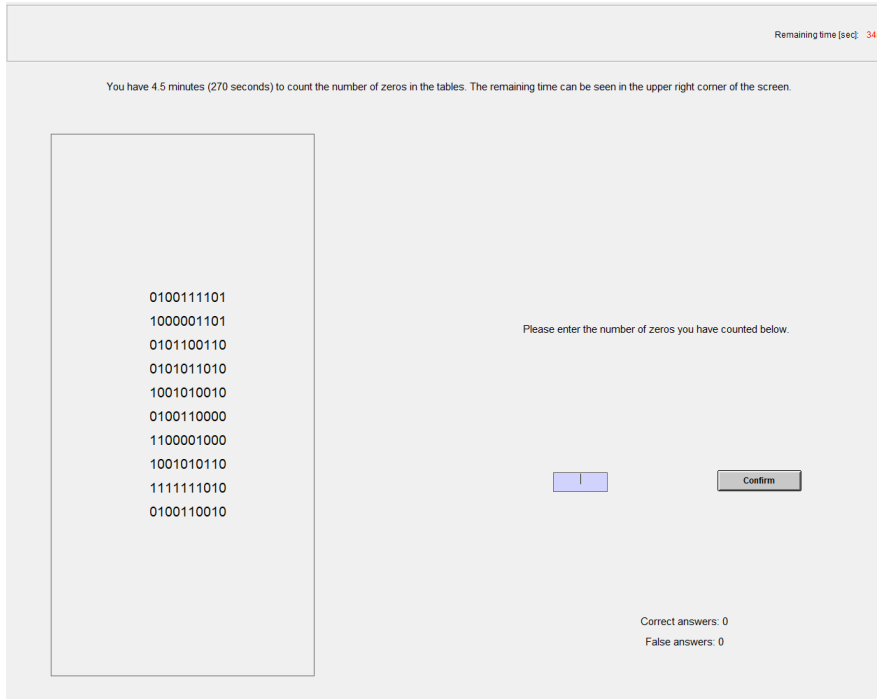


Fig. 1 Exemplary screen of grid task

2.2 Preparatory session

The preparatory session consists of four parts (see [Table 1](#)). First, subjects familiarize themselves with the grid task in a practice round of two minutes. We then measure each subject's baseline productivity, followed by a questionnaire and a choice of individual working time. The preparatory session serves the purpose to classify subjects into binary types with high and low productivity or long and short working time, respectively. A subject's type determines whether they will be among those favored by a given affirmative action policy in the main session.

Classification into productivity type. Subjects are asked to solve as many grids as they can within five minutes at a high piece-rate of 0.50 EUR for each correctly solved grid. Subjects who solve more grids correctly than the median are classified as being of the high productivity type; those below the median as low productivity type.¹ [Figure A.1](#) in [Appendix A](#) displays the distribution of the number of correctly solved grids in this stage.

¹ Our measurement of baseline productivity (number of correctly solved grids per minute) might capture both the given ability of working on the grid task and effort at the intensive margin. However, subjects tend to exert maximum effort in laboratory real effort tasks with short working period (Araujo et al., 2016; Corngnet et al., 2015; Gächter et al., 2016; Goerg et al., 2019). Therefore, our preferred interpretation of baseline productivity is that it reflects ability in the grid task, which is beyond subjects' control during the experiment.

Classification into working time type. At the end of the preparatory session, subjects work on the grid task again, now at a piece-rate of 0.10 EUR per correctly solved grid. Subjects can freely choose how long they want to work. After every grid, subjects can choose to continue or stop working by clicking on the corresponding button. If subjects choose to stop working, they finish the preparatory session and can leave the laboratory immediately. To minimize peer effects in the decision when to stop working, we implement a flexible show-up policy, meaning that subjects start the session individually and do not reach this last stage at the same time.² We truthfully communicate to subjects upfront that their chosen working time has important consequences for their working time and expected earnings in the main session in the following week. Based on whether their chosen working time is below or above the median, they are classified as being either of the short or long working time type. This determines how long they will work on the grid task in the main session and affects their expected earnings.³ The low piece-rate was deliberately chosen to make the task less attractive so subjects would choose to stop working after a reasonable time. On average they do so after 24.13 minutes. [Figure A.2 in Appendix A](#) displays the distribution of working time. In our data, the number of correctly solved grids in our measurement of baseline productivity and chosen working time are not significantly correlated (Pearson's correlation coefficient: $\rho = 0.061$, $p = 0.192$), suggesting that an individual's type in the productivity and the working time dimension is not strongly correlated.

Questionnaire. We elicit a number of control variables, including measures of risk and social preferences, cognitive ability (Raven matrices), personality (Big Five), and socio-demographics. Online [Appendix D](#) provides more details on the questionnaire.

² In the recruitment e-mail, subjects are informed that they can show up at the lab at any time within a two-hour interval.

³ In particular, the instructions state “Your working time today will determine your working time in the next session next week. Next week, you will work on a similar task and you will be given a specific amount of time to solve as many tables as possible and get paid accordingly. (...) Based on your chosen working time today, we will form two groups. One group consists of that half of the participants who choose to work for a shorter time today. This group will also be given a shorter time to work in the session next week. The other group consists of that half of the participants who choose to work for a longer time today. This group will also be given a longer time to work in the session next week. Who works shorter will, on average, solve less tables correctly and therefore earn less. The experiment is, however, shorter (it will end earlier). Who works longer will, on average, solve more tables correctly and therefore earn more. The experiment is, however, longer (it will end later).” These consequences were indeed implemented in the main session. In addition to working less, short working time subjects received their payments first and were therefore able to leave earlier. In the main session, short working time subjects earned 8.08 EUR on average, while long working time subjects earned 12.67 EUR on average.

2.3 Main session

The main session consists of five stages and a final questionnaire.⁴ In stages 1 to 3, subjects work on the grid task repeatedly, with their payoff-relevant performance being determined as follows:

Performance = Correct grids per minute × Working time × Randomly assigned multiplier.

Assignment of types. In our experiment, each subject is of one of eight (2^3) types: high or low productivity × long or short working time × discriminated or non-discriminated. Subjects are fully informed about all three dimensions of their own type before they enter the first stage of the main session and a subject's type stays constant throughout the experiment. The productivity and working time type are assigned to each subject based on the outcomes of the preparatory session as described in Section 2.2 above. While we treat productivity in this experiment as given (see also Footnote 1), subjects have more or less time to work on the grid task in the main session according to their chosen working time in the preparatory session. Subjects of the long working time type have 7.5 mins to work on the task, while those of the short working time type have only 4.5 mins. Discrimination is reflected by a randomly assigned multiplier. Half of the subjects are non-discriminated. They are assigned a high multiplier of 1.25. The other half are discriminated, being assigned a low multiplier of 0.75. The number of correctly solved grids in the total time worked is weighted with the individual multiplier. Those parameters were chosen to make the effects of each policy on the probability of winning for those favored (not favored) by it comparable in size.⁵

Stage 1: Piece rate. This stage provides a baseline measure of performance without tournament incentives. Subjects work on the grid task according to their type's working time and receive a piece-rate payment of 0.50 EUR for each correctly solved grid multiplied with the respective multiplier.

Stage 2: Tournament. In stage 2, subjects solve the grid task under tournament incentives, each competing against five other subjects. The purpose of this stage

⁴ Part of the design of the main session builds on Balafoutas et al. (2016). We thank the authors for sharing their ztree program and instructions with us.

⁵ Specifically, after observing the baseline productivity distribution in the preparatory session, we run a simple simulation that determines how likely it is on average for the low and high productivity type to win the tournament with and without affirmative action regarding productivity. We then choose parameters for working time of the short and long working time type such that the change in the average probability of winning the tournament when introducing affirmative action regarding working time is comparable. Similarly, the random multipliers are chosen such that the change in the average probability of winning the tournament when introducing affirmative action regarding discrimination is also comparable in size.

is to measure the effect of different affirmative action policies on performance. Among each group of six, the two winners of the tournament receive 1.50 EUR per correctly solved grid multiplied with their multiplier, while losers receive nothing, keeping the average payment constant compared to stage 1. If necessary, a random tie-breaking rule is applied to determine winners. Winners and rank within each group of six are not announced until the end of stage 5. How exactly the winners are determined is the main treatment variation of our experiment.

Between-subject treatments. We conduct four treatments: one control treatment without affirmative action and three different affirmative action (AA) treatments. In the control treatment without any quota rule, the two subjects with the highest performance (according to the definition above) are the winners. In the affirmative action treatments, a quota rule is added to determine the winners. If this rule is not automatically fulfilled, the subject with the second-highest performance is replaced by the highest-performing subject who fulfills the quota criterion. The following quota rules apply in the treatments:

- *Control (CTR)*: No quota rule.
- *Affirmative action w.r.t. discrimination (AAD)*: At least one subject of the discriminated type has to be among the two winners.
- *Affirmative action w.r.t. working time (AAW)*: At least one subject of the short working time type has to be among the two winners.
- *Affirmative action w.r.t. productivity (AAP)*: At least one subject of the low productivity type has to be among the two winners.

Control questions make sure that subjects understand the tournament before starting to work on the task.⁶

Stage 3: Self-selection into tournament. To elicit the willingness to compete in the tournament, subjects work on the grid task again. In this stage, they choose whether they would like to work under piece-rate incentives (exactly as in stage 1) or tournament incentives (as in stage 2). Importantly, if a subject chooses the tournament in *stage 3*, their performance will be compared to the performance of their five fellow group members in *stage 2*. This feature ensures that subjects' decision

⁶ After having read the rules of stages 2 and 3, subjects have to answer control questions correctly before they can start working on the grid task. These multiple-choice questions describe scenarios about competition within a group, provide information about each member's performance and who is favored by affirmative action, and ask about the winners. The control questions cover both cases in which affirmative action does or does not change the results of the competition.

to enter the tournament is independent of their belief about others entering the tournament (compare Niederle et al., 2013).

Belief elicitation. At the end of stage 3, subjects report their beliefs about their relative performance in stage 1, 2, and 3. Subjects are asked to guess their rank both within the whole group of six and within the group of three subjects with the same discrimination type (in treatment AAD), the same working time type (in treatment AAW), or the same productivity type (in treatment AAP). One guess is randomly chosen to be payoff-relevant. Subjects receive 1 EUR if their guess is correct.

Stage 4: Teamwork. Similar to previous work (e.g., Balafoutas et al., 2016; Sutter et al., 2016), we use this stage to evaluate whether post-competition teamwork is affected by the implementation of an affirmative action policy, conditional on the type of affirmative action. We keep the group composition and treatment history from previous stages, but provide a new working environment with a new task and new payoff rules to make clear to subjects that all subjects now work for the same amount of time (5 mins), and there are neither multipliers nor affirmative action.

In the slider task (Gill and Prowse, 2012), subjects are shown a series of screens, each with 6 sliders. Each slider has a range of positions between 0 and 100. Sliders are solved by using the computer mouse to move the slider markers to the position of 50.⁷ A screen is considered “solved” if all six sliders are positioned at 50. Only then can a subject continue to the next screen.

Importantly, each correctly solved screen yields 0.60 EUR for the group as a whole, 0.10 EUR for each of its members. Since all group members benefit equally from an individual’s effort, this is a typical setup to measure teamwork and how much subjects work indicates their willingness to contribute for the benefit of the group. In order to keep the previous tournament experience with or without affirmative action salient, we introduce unequal bonuses for winners and losers of the tournament in stage 2 (as Balafoutas et al., 2016, do). Subjects receive a bonus of 5 EUR if they were the winners in stage 2, and 2 EUR otherwise.

Feedback. At the end of each of the first four stages, subjects are informed about their individual performance. Subjects never learn the performance of other subjects. However, at stage 5 they learn whether they themselves were a winner in stage 2. About the five other subjects, they learn whether they were a winner in stage 2 or loser and favored or not.

⁷ To make sure that subjects use only the computer mouse to solve the task, the left and right arrow keys of the keyboard are disabled.

Stage 5: Dictator game. Subjects play one dictator game with each of their five group members. The only thing they know about the other group members is whether they were winners in stage 2 and whether they were favored by affirmative action. All five dictator games are displayed on the same screen. For each game, subjects are endowed with 5 EUR, and can decide how much to give away in 0.1 EUR increments. We use this setup to learn whether previously non-favored subjects (in particular losers) retaliate against previously favored winners after the tournament phase.

Fairness perception. After stage 5, subjects are asked how fair they perceive the different quota rule to be. They first rate the rule in their own treatment on a 7-point Likert scale and then those of the other three treatments. The fairness questions describe the policies neutrally and do not mention the term “affirmative action” or “discrimination” (see Online [Appendix B](#) for the exact wording).

2.4 Procedures

The experiment was conducted at the DICE Lab at the University of Düsseldorf in April 2018 and the BonnEconLab at the University of Bonn in August 2018 using the software zTree (Fischbacher, 2007). On average, the preparatory sessions lasted about 60 minutes, the main sessions about 90 minutes. Subjects were recruited via ORSEE (Greiner, 2015) and Hroot (Bock et al., 2014) from the subject pools of the respective labs, both of which include students of various disciplines. In the recruitment email, subjects are informed that the experiment consists of two mandatory sessions and that payments will only be realized at the end of the second session. Only 7 out of 463 subjects who participated in the preparatory session did not show up in the main session, implying an attrition rate of 1.5%. To be able to match the data of both sessions while ensuring anonymity, we asked subjects to generate an ID (that is never connected to their name) at the beginning of the preparatory session, and to re-enter it in the main session.

In order to ensure that tournament outcomes such as efficiency or beliefs about winning the tournament are not driven by an unbalanced composition of types within groups or treatments, each group of six consists of three subjects with high and three with low productivity; three with long and three with short working time; as well as three with the high and three with the low multiplier.

At the beginning of stage 2, subjects are informed about this rule for group composition, but not about the specific type of each other group member. Group composition remains the same in all following stages. To achieve such a balanced

group composition, 18% of subjects could not participate in the main session, since they could not be matched to a balanced group within their specific main session. These subjects receive their earnings from the preparatory session plus a 4 EUR show-up fee for the main session. We observe virtually no differences between this group and subjects who participated in both sessions (Mann-Whitney U tests, $p = 0.949$ regarding baseline productivity, $p = 0.967$ regarding working time). In total, the final number of subjects in each treatment is 108, 84, 90, and 90 for CTR, AAD, AAW, and AAP, respectively. Gender composition does not vary significantly between treatments (59% females on average; Kruskal-Wallis test: $p = 0.582$).

The different treatments were run in different sessions. Subjects were randomly assigned to a treatment: they signed up for a specific main session; only after that we randomized in which session to run which treatment.

On average, subjects earned 26 EUR for both sessions. The payoff of the preparatory session is the sum of payoffs from the measurements of baseline productivity and choice of working time in the grid task, a risk choice list, and a fixed payment of 2 EUR for completing the questionnaire. The payoff of the main session consists of a 4 EUR show-up fee and the earnings from one randomly chosen stage. Subjects know that one of the five parts of the main session will be randomly chosen to be paid. Instructions are distributed stage-by-stage.

3 Theoretical framework

In this section, we introduce a theoretical framework to derive predictions for the main session of our experiment. We will refer to it throughout the paper to help guide our data analysis.

Let $b_i \sim G(b_i)$ be subject i 's baseline productivity, i.e., how many grids a subject can solve in one minute under full effort. $G(\cdot)$ is symmetric, atomless, and has full support on $[0, B]$. Also let $t_i \in \{4.5, 7.5\}$ be a subject's working time, $e_i \in [0, 1]$ a subject's effort, $c(e_i) \geq 0$ the associated effort cost (with $c'(e_i) > 0$ and $c(0) = 0$), and $m_i \in \{0.75, 1.25\}$ a subject's multiplier. A subject's performance is a function of all of those factors as well as a random variable $\varepsilon_i \sim U[-E, E]$, where $E > 0$ is small such

that performance is non-negative:⁸

$$p_i(b_i, t_i, e_i, m_i, \varepsilon_i) = b_i t_i e_i m_i + \varepsilon_i.$$

A subject's payoff π_i further depends on a state variable S that can take values pr , cl , and cw . Under a piece rate, they receive a monetary payment of $\pi_i(b_i, t_i, e_i, m_i, \varepsilon_i, pr) = 0.5(b_i t_i e_i m_i + \varepsilon_i)$. If subjects lose the tournament, they receive $\pi_i(b_i, t_i, e_i, m_i, \varepsilon_i, cl) = 0$, and if they win, they receive $\pi_i(b_i, t_i, e_i, m_i, \varepsilon_i, cw) = 1.5(b_i t_i e_i m_i + \varepsilon_i)$. For simplicity, we assume that subjects are risk-neutral. We focus on the case in which the effort cost is sufficiently low such that subjects choose $e_i = 1$ (see also [Footnote 1](#)).

Subjects also have fairness preferences, which enter the utility function in addition to the monetary payment. In a similar vein as Cappelen et al. (2010, 2013) and Mollerstrom et al. (2015) but applied to a tournament setting, we assume that subjects suffer a utility loss if the implemented allocation vector Y differs from the allocation vector they consider fair, $F^{k(i)}$ (with their respective entries y_j and $f_j^{k(i)}$). This loss is weakly increasing in the deviation from the fair allocation. Thus, a subject's utility function is

$$u_i(b_i, t_i, e_i, m_i, \varepsilon_i, S, \gamma_i, k(i)) = \pi_i(b_i, t_i, e_i, m_i, \varepsilon_i, S) - \gamma_i \sum_{j=1}^n (y_j - f_j^{k(i)})^2.$$

Generally, Y and $F^{k(i)}$ are context-dependent (which for brevity, we do not denote explicitly) and may include monetary payoffs as well as effort costs.⁹ However, as we assumed earlier, effort costs are constant and symmetric across individuals. Thus, we do not consider them from here on.

Subjects differ in the weight they attach to fairness considerations, indicated by a fairness sensitivity parameter $\gamma_i \geq 0$. We assume that fairness concerns are not dominant, i.e., the direct monetary utility gains outweigh any potentially increasing losses from fairness concerns. That means holding other subjects' payments constant, a subject has a higher utility u_i if their own monetary payment increases.

Moreover, subjects differ in their fairness type $k \in \{\text{lib, mer, ceg, seg}\}$ that determines which allocation $F^{k(i)}$ they consider fairest, taking the institutional envi-

⁸ For simplicity, we take a subject's number of solved grids per minute during the first part of the preparatory session at face value to group subjects into productivity types and call it baseline productivity, assuming that subjects exert maximal effort. Thus, ε represents a combination of factors outside of our model that influence performance in the preparatory and main session of our experiment, such as randomly occurring distractions.

⁹ Y and $F^{k(i)}$ may depend on the various factors that determine all subjects' performance (t, b, m, ε), whether they enter the competition, and whether they win.

ronment as given. While we use the same overarching concept of libertarianism, meritocratism, and choice and strict egalitarianism, respectively, as, e.g., Cappelen et al. (2010), a specific fairness type's most preferred allocation does not coincide one-to-one with the previous literature because we consider different institutional environments, i.e., tournaments or a piece-rate payments instead of dictator games. Libertarians' fairest allocation F^{lib} is the one without any kind of affirmative action. For a tournament, this coincides with the outcome of treatment CTR, while, for the piece rate, this coincides with a piece rate payment based on performance. Meritocrats believe that the fairest allocation, F^{mer} , is the one that corrects for different m_i and realizations of ε_i , or stated differently, the one implemented if all subjects had the average multiplier $m_i = 1$ and the average realization of $\varepsilon_i = 0$. Choice egalitarians think that out-of-personal-control factors, i.e., discrimination and differences in baseline productivity, should not be accounted for. Thus, F^{ceg} is the allocation that is implemented if, in addition to the average m_i and the average realization of ε_i , all subjects had the average baseline productivity $b_i = B/2$. Finally, strict egalitarians' fairest allocation F^{seg} also imposes the average working time $t_i = 6$ for all subjects, resulting in a symmetric, random determination of winners in a tournament setting or a flat payment in a piece rate setting.

In a tournament environment, subjects know that they are put in groups of six such that in each group, half of the subjects have a high baseline productivity \bar{b} ($b_i \geq B/2$), half of the subjects have a high multiplier \bar{m} ($m_i = 1.25$), and half of the subjects have a long working time \bar{t} ($t_i = 7.5$); analogously for \underline{b} , \underline{m} , and \underline{t} . Subjects know their own baseline productivity (and whether it is high or low), whether they have a high or low multiplier, and whether they have a long or short working time. About other subjects, they can only make inferences given what they know about themselves. We assume that those features are independently distributed within the given constraints on group composition.

4 Hypotheses and results

We start by presenting evidence on the fairness perception of the various affirmative action policies. We then analyze the consequences of affirmative action for willingness to compete in the tournament and efficiency, before we examine its impact on post-competition teamwork and retaliation. Moreover, we investigate the predictive power of individual-level fairness perceptions for behavior. For each of these outcomes, we use the framework from [Section 3](#) to derive hypotheses and guide our analysis.

4.1 Fairness perceptions of affirmative action

To be able to make statements about the fairness ranking of the affirmative action policies in our treatments based on the concepts of libertarianism, meritocratism, choice egalitarianism, and strict egalitarianism, it is necessary that subjects indeed perceive productivity as a personal but fixed, i.e., not self-chosen factor, working time as personal but self-chosen, and the multiplier as a neither personal nor self-chosen factor. As the multiplier is randomly assigned to subjects the last aspect can be taken for granted. In the instructions, we explicitly remind subjects that, as pre-announced, working time in the main session is based on the self-chosen working time of the preparatory session. Therefore, we are confident that subjects actually view it as self-chosen. Regarding productivity, we neither find evidence of learning effects nor of diminishing productivity over time (as we discuss when introducing the grid task).

As people differ in the fairness ideals they adhere to (e.g., Konow, 2003; Cappelen et al., 2010), we anticipate substantial heterogeneity between subjects in the fairness perception of the various affirmative action policies. This is also what we find, see [Figure A.3](#).

Hypotheses on average fairness perceptions. In order to derive more specific hypotheses on how subjects rank the fairness of the various affirmative action policies, we consider the part of the utility function that describes fairness preferences, $Prob(cw)E_{cw}[-\gamma_i \sum_{j=1}^n (y_j - f_j^{k(i)})^2] + Prob(cl)E_{cl}[-\gamma_i \sum_{j=1}^n (y_j - f_j^{k(i)})^2]$, i.e., subject i 's expected fairness utility loss from the competition, weighted by the respective probabilities. In the following, we discuss the fairness evaluations of the changes introduced through different affirmative action rules in the competition compared to competition with no such rule.

Libertarians dislike all affirmative action rules as the implemented allocation will tend to differ from F^{lib} (it strictly differs as soon as the affirmative action rule is binding). In particular, libertarians dislike a rule the more, the more it affects subjects' probability of winning the competition. All meritocrats, choice egalitarians, and strict egalitarians consider AAD as fairer than CTR as the allocations it implements are closer to their preferred allocation F^k than allocations in the absence of any affirmative action rule in CTR. For analogous reasons, choice and strict egalitarians judge AAP as fairer than CTR, whereas meritocrats—as libertarians—

conceive it less fair.¹⁰ Finally, strict egalitarians perceive AAW as fairer than CTR, whereas meritocrats, choice egalitarians, and libertarians consider it less fair.

As we can neither estimate the exact share of each type in our sample, nor make precise statements about the relative strength of each preference for each fairness ideal, we can only make limited statements on the aggregate ranking. However, as long as there are not mostly libertarians and there are some choice egalitarians, the following should hold:¹¹

Hypothesis 1. *On average, subjects consider the affirmative action rule that favors subjects with a low multiplier as fairer than all other affirmative action rules in this experiment and than the absence of an affirmative action rule. Moreover, they consider the affirmative action rule that favors subjects with low productivity as fairer than the affirmative action rule that favors subjects with a short working time.*

Results on average fairness perceptions. As hypothesized, [Figure 2](#) illustrates that AAD is indeed perceived as significantly fairer than CTR, AAP, and AAW. Its average fairness rating is 4.70 on a Likert scale from 1 to 7 and 48% of subjects rate it at or above 5. [Figure 2](#) also provides results of Wilcoxon signed rank tests for all pairwise policy comparisons.¹²

While the relative ranking of AAP compared to AAW contrasts [Hypothesis 1](#), it is in line with the results of Cappelen et al. (2010) who find that a substantial majority holds others responsible for their productivity, even in situations where productivity should be considered beyond individual control. Both our results and those of Cappelen et al. (2010) suggest that, in contrast to choice egalitarians, many people accept individual merits and fairness perceptions reflect more than the distinction between factors that are beyond and factors that are within individual control.

In line with our predictions, AAW (average score 4.22) is considered significantly less fair than AAD. It is also judged as significantly fairer than no affirmative action on average.

¹⁰ Generally, the effects of AAP on fairness perceptions should be smaller than the effects of other rules for the following reason: the favored winner will likely be someone with a relatively high baseline productivity among subjects with a low baseline productivity and the subject they potentially replace will likely be someone with a relatively low baseline productivity among subjects with a high baseline productivity; by contrast, the other factors (working time and multiplier) are binary.

¹¹ We find it reasonable to assume that libertarians are not in the majority for two reasons. First, in our data, only about 11% of subjects rank CTR as fairer than all three other treatments (as libertarians would). Second, Cappelen et al. (2010) estimate the share of libertarians to be 26% among Norwegian students of economics and business administration and about 45% among the respective alumni with labor market experience. These percentages are upper bounds if people who tend to hold libertarians fairness ideals are more likely to study business administration or economics.

¹² We provide power calculations for these and all following tests that refer to our main results in [Section A.6 of Appendix A](#).

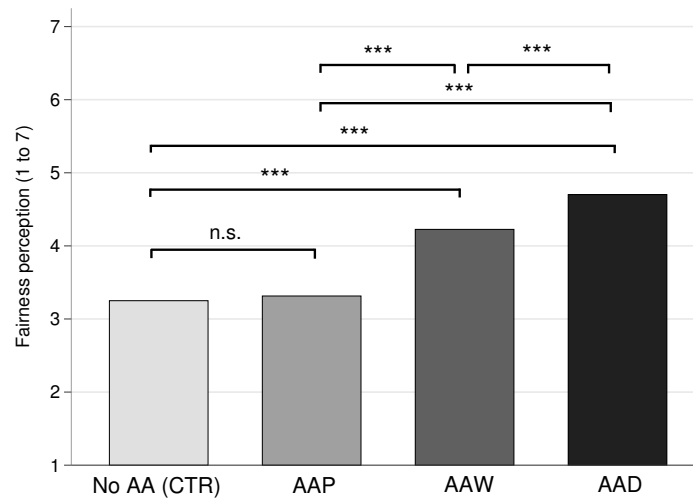


Fig. 2 Fairness perception

Notes: Perceived fairness of different policies elicited on a 7-point Likert scale, answering the item "How fair do you perceive the rules of the competition to be?". Higher numbers indicate that a policy is perceived as fairer. No AA (CTR): No affirmative action. AAP: Affirmative action w.r.t. productivity. AAW: Affirmative action w.r.t. working time. AAD: Affirmative action w.r.t. discrimination. The brackets and stars above each bar show results of Wilcoxon signed rank tests. *** indicates $p < 0.001$. $N = 372$.

Compatible with the fact that two fairness types rank AAP as fairer and two types rank it as less fair than CTR, [Figure 2](#) documents that its average fairness rating is roughly equal to no affirmative action; the average score is 3.31 for AAP and 3.25 for CTR.

All results regarding the fairness ranking in [Figure 2](#) remain robust in a regression framework; see [Table A.1](#) where we regress individual fairness perceptions on binary variables indicating which treatment the fairness perception refers to, using the control treatment as baseline.

Result 1. *There is broad support for all three affirmative action policies. None of them is ranked as less fair than the absence of affirmative action (CTR). On average, the affirmative action policy that compensates for discrimination (AAD) is considered as fairest. Affirmative action favoring individuals with short working time (AAW) is perceived as significantly fairer than affirmative action based on low productivity (AAP).*

Further determinants of fairness perceptions. We continue by exploring further possible determinants of subjects' fairness perceptions: do they exhibit a self-serving bias? Are they affected by having experienced a specific policy? In our context, self-serving fairness beliefs would imply that low types (i.e., low productivity, short working time, and discriminated subjects, respectively) who have been or

would have been favored by a specific affirmative action policy consider this policy as fairer than the respective high types. This is what we find for all three affirmative action policies under study, see Panel 1 of Table 2. Those who (would) have been favored by an affirmative action policy rank it as between 0.48 and 1.07 points fairer on a 7-point Likert scale.

	(1)	(2)	(3)
Panel 1:	AAP	AAW	AAD
High type	-1.070*** (0.180)	-0.806*** (0.195)	-0.479*** (0.186)
Constant	3.849*** (0.134)	4.629*** (0.133)	4.941*** (0.137)
Observations	372	372	372
Panel 2:	AAP	AAW	AAD
Low type in AA	0.660* (0.365)	-0.999*** (0.381)	-1.242*** (0.380)
High type in CTR	-0.949*** (0.334)	-1.350*** (0.379)	-1.280*** (0.381)
High type in AA	-0.535 (0.347)	-1.177*** (0.400)	-1.242*** (0.367)
Stage 2 winner	0.602** (0.258)	0.263 (0.313)	0.312 (0.313)
Constant	3.348*** (0.262)	4.800*** (0.252)	5.304*** (0.226)
Observations	198	198	192

OLS regressions, standard errors in parentheses, clustered at subject level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The dependent variable is perceived fairness of different policies elicited on a 7-point Likert scale, answering the item "How fair do you perceive the rules of the competition to be?". Higher numbers indicate that a policy is perceived as fairer. In Panel 1, the baseline category is Low type; in Panel 2, the baseline category is Low type in CTR. Wald tests on the equality of coefficients of High type in CTR and High type in AA in Panel 2 do not indicate significant differences (all p-values larger than 0.188).

Table 2 Fairness perception

By contrast, fairness judgments regarding affirmative action do not seem to be systematically shaped by having experienced a specific policy. To explore the role of personal experience, we compare how fair a policy is rated by a type actually affected by it in an affirmative action treatment to how it is rated by the same type in CTR. Panel 2 of Table 2 displays these results for each of the affirmative action

treatments. For high types, there is no evidence that experiencing an affirmative action policy as opposed to judging it by description only affects fairness perceptions. None of the three Wald tests for equality of coefficients of “High type in CTR” and “High type in AA” indicates a significant difference and differences are small. Favored low types types in AAW and AAD consider affirmative action as significantly less fair than their control group counterparts, while the respective coefficient is positive, but not significantly different from zero at the conventional 5% level in AAP. Finally, Panel 2 also includes a binary indicator whether a subject won the forced stage 2 tournament that does not turn out to be significant in AAW and AAD, but only in AAP in predicting subjects’ fairness judgments.

4.2 Willingness to compete

Hypotheses on willingness to compete. According to our framework, subjects maximize their expected payment minus their expected utility loss from fairness considerations when deciding whether to enter the competition in stage 3. They should enter if and only if their expected utility from the tournament (which they win with $Prob(cw)$ and lose with $Prob(cl)$) is weakly larger than the expected utility from the piece rate:

$$Prob(cw)\mathbb{E}_{cw}[1.5(b_i t_i m_i + \varepsilon_i) - \gamma_i(1.5(b_i t_i m_i + \varepsilon_i) - f_i^{k(i)})^2] + Prob(cl)\mathbb{E}_{cl}[-\gamma_i(0 - f_i^{k(i)})^2] \geq 0.5(b_i t_i m_i) - \gamma_i\mathbb{E}_{pr}[(0.5(b_i t_i m_i + \varepsilon_i) - f_i^{k(i)})^2].$$

Note that other subjects’ payments are independent of subject i ’s decision at this stage, as i is competing against other subjects’ performances from stage 2.

Beliefs enter the inequality through the probability of winning and the value of ε . As a benchmark, beliefs are assumed to be correct. Even if beliefs are incorrect, however, subjects’ willingness to enter increases in the belief of winning. It seems reasonable to assume that subjects’ beliefs of winning increase in baseline productivity, working time, multiplier, the presence of a rule that favors them, and the absence of a rule that does not favor them. We then expect the following to be true:

Hypothesis 2. *Subjects are more likely to enter the competition when they are favored by affirmative action.*

Next, we turn to how fairness concerns influence willingness to compete. Holding the chance of winning constant, the fairer subjects consider an affirmative ac-

tion rule, the more likely they are to enter the competition as the affirmative action rule only influences the tournament allocation but not the piece rate allocation. That means the left hand side of the inequality above increases due to a smaller fairness utility loss for a constant right hand side. As we designed the experiment such that being favored increases the average probability of winning by the same amount across all affirmative action rules, we can test the following:

Hypothesis 3. *The fairer subjects consider an affirmative action rule, the more likely they are to enter the competition.*

This holds for both favored and non-favored subjects. As there are no favored and non-favored subjects in the absence of an affirmative action rule, a direct comparison of average willingness to compete in the affirmative action treatments with the control treatment is not possible. Whereas the fraction of competing subjects across different affirmative action rules should increase in the perceived fairness of those rules, it is not clear how the control treatment ranks: libertarians consider the tournament allocation in the control treatment fair and those in all affirmative action treatments unfair, while they always perceive the piece rate as fair. All other types find the tournament allocation always less fair than the piece rate allocation—unless subjects expect to be very likely to lose when they compete, in which case they should not compete anyway. This is because winners receive three times the piece rate, which increases the fairness utility loss if the piece rate is already perceived as unfair (which it is by every type but libertarians as it is based on subjects' unadjusted performance). An affirmative action rule may reduce this gap. Thus, the share of competing subjects in the control treatment relative to the affirmative action treatments weakly increases in the share of libertarians among subjects.

Results: aggregate level. As discussed above, there is exists no clear-cut prediction how the average willingness to compete in CTR compares to that in the affirmative action treatments. Combining [Hypothesis 3](#) with the empirically observed, average fairness rating of the affirmative action treatments leads us to expect that average willingness to compete should be lower in AAP than AAW than AAD. This is what we find: 45% of subjects are willing to compete in the absence of affirmative action (CTR) compared to 39% in AAP, 44% in AAW, and 51% in AAD. However, the overall proportion of subjects selecting the tournament differs only modestly across treatments and pairwise comparisons with CTR are not significant (Fisher's exact tests yield $p = 0.388$ for AAP, $p = 1.000$ for AAW, and $p = 0.468$ for AAD). Likewise, a

Kruskal-Wallis test does not detect any significant differences across all four treatments ($p = 0.446$).

Results: determinants of willingness to compete at the individual level. Table 3 reports results of OLS regressions on the determinants of willingness to compete. The dependent variable is an indicator variable that equals 1 if a subject chooses to enter the tournament and 0 if they choose the piece rate payment in stage 3. As expected, Table 3 shows that the belief regarding own rank in the tournament is indeed predictive for individual willingness to compete. If subjects' belief about their own rank increases by 1 (on a scale from 1 to 6, with 1 as the best rank), their willingness to enter the tournament decreases by about 10 percentage points in the specification reported in column (1). Table A.2 replicates specifications (1)-(3) of Table 3 except for estimating the coefficient of each value of belief separately. Results closely resemble those in Table 3. Moreover, results are qualitatively the same and quantitatively very similar if we use beliefs regarding own rank in stage 2 instead of stage 3.

While our theoretical framework assumes risk neutrality for simplicity, a clear expectation is that risk averse subjects are less willing to enter the tournament. We observe that for a 1 point higher willingness to take risks (on an 11-point Likert scale), willingness to compete raises by 3 percentage points. Also, in line with the literature (see, e.g., Almås et al., 2016; Buser et al., 2014; Croson and Gneezy, 2009; Gneezy et al., 2003; Leonard et al., 2009; Niederle and Vesterlund, 2007, 2011; Sutter and Glätzle-Rützler, 2015), women tend to be less likely to compete than men, even controlling for risk aversion and beliefs.¹³ Sign and often also size of these coefficients are stable across specifications in Table 3.

In order to learn about the effect of being favored, we add a binary indicator variable favored (1 if favored, 0 if non-favored) in column (2). This requires dropping observations from CTR. In line with Hypothesis 2, we observe a strong overall encouragement effect of being favored: on average, favored subjects are 21 percentage points more likely to enter the tournament than non-favored subjects when considering all affirmative action treatments jointly. Columns (4)-(6) analyze the effect of being favored or non-favored for each affirmative action policy separately, comparing the willingness to compete of a given type in an affirmative action treatment to that of the same type in CTR. The encouragement effect of being favored is most striking for AAD, in which favored subjects' tournament participation increases by 28 percentage points compared to CTR. Subjects with short working time

¹³ For a conceptual contribution on this topic refer to Van Veldhuizen (2022).

	(1)	(2)	(3)	(4)	(5)	(6)
	All	AA	AA	AAP	AAW	AAD
Belief on rank	-0.102*** (0.017)	-0.116*** (0.021)	-0.120*** (0.020)	-0.133*** (0.023)	-0.121*** (0.024)	-0.106*** (0.026)
Risk attitude	0.026** (0.011)	0.029** (0.013)	0.029** (0.012)	0.012 (0.015)	0.046*** (0.014)	0.009 (0.015)
Female	-0.093* (0.050)	-0.088 (0.058)	-0.087 (0.057)	-0.137** (0.068)	-0.011 (0.068)	-0.115 (0.072)
Fairness perception	0.019 (0.013)	0.017 (0.016)	0.048** (0.022)	0.025 (0.020)	-0.005 (0.019)	-0.001 (0.021)
Favored		0.209*** (0.062)	0.460*** (0.134)			
Favored × Fairness			-0.064** (0.030)			
Low type in AA				0.020 (0.097)	0.247** (0.095)	0.282*** (0.104)
High type in CTR				0.132 (0.088)	0.236*** (0.090)	0.152 (0.092)
High type in AA				-0.031 (0.092)	0.011 (0.095)	-0.038 (0.104)
Constant	0.635*** (0.097)	0.555*** (0.113)	0.454*** (0.122)	0.759*** (0.142)	0.507*** (0.147)	0.739*** (0.147)
Observations	372	264	264	198	198	192

OLS regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The binary dependent variable is willingness to compete in stage 3 (1 if tournament, 0 if piece rate). Base categories are non-favored subjects in (2) and (3); Low type in CTR in (4), (5), (6). Wald tests on the equality of coefficients between High type in CTR and High type in AA indicate significant differences for (4), (5), and (6); $p=0.075$, $p=0.015$, $p=0.060$, respectively. *Belief on rank* is a categorical variable reflecting beliefs regarding own rank in stage 3 (between 1 and 6); *Risk attitude* is the answer to the general risk question elicited on an 11-point scale on which higher numbers indicate a higher willingness to take risks; *Female* is an indicator variable for gender (1 if female, 0 if male); *Fairness perception* reflects fairness rating of own treatment, elicited on a 7-point scale, on which higher numbers indicate higher perceived fairness; *Favored* is an indicator variable (1 if favored, 0 otherwise). Due to the inclusion of the interaction effect in (3), we estimate OLS instead of a non-linear model (Ai and Norton, 2003). The sample in column (1) includes subjects from all four treatments; in columns (2) and (3) subjects from all three affirmative action (AA) treatments; in (4), (5), and (6) subjects from CTR and the respective AA treatment.

Table 3 Willingness to compete

are 25 percentage points more likely to compete in AAW than in CTR. By contrast, tournament entry of low productivity subjects in AAP hardly increases compared to CTR (2 percentage points). These encouragement effects of affirmative action on favored subjects closely mirror the average fairness rating of the affirmative action treatments in [Figure 2](#). Favored subjects seem to be more willing to compete under affirmative action policies with higher average fairness ratings. Again in line with [Hypothesis 2](#), we observe at least marginally significant discouragement effects on non-favored subjects (10 percentage points in AAP, $p = 0.075$; 23 percentage points in AAW, $p = 0.015$, and 11 percentage points in AAD, $p = 0.060$).¹⁴

Overall, results in [Table 3](#) suggest that an individual's own fairness perception is not a significant predictor of willingness to compete when controlling for beliefs, risk attitudes, and gender. The relevant coefficient is small and not significantly different from 0 in any of the specifications in columns (1), (2), and (4) through (6). A simple correlation between fairness perception and willingness to compete is, however, positive and significant, but not large (Spearman's $\rho = 0.131$, $p = 0.012$, $N = 372$, i.e., pooling data from all four treatments). However, these findings mask an interesting heterogeneity. In column (3) of [Table 3](#), we investigate whether the role of fairness perceptions differs for favored and non-favored subjects. Results in column (3) imply that an individual's own fairness perception is a significant predictor of non-favored subjects' willingness to compete: if their fairness perception of the affirmative action treatment increases by 1 (on a 7-point Likert scale), they are 5 percentage points more likely to enter the tournament. For favored subjects, however, the individual fairness perception of an affirmative action policy does not affect their willingness to compete. A test of joint significance of Fairness perception and Fairness perception \times Favored yields $p = 0.488$. Both results are confirmed in the alternative specification in column (3) of [Table A.2](#). Likewise, the simple correlation is substantially stronger for non-favored (Spearman's $\rho = 0.253$, $p = 0.003$, $N = 132$) than for favored subjects (Spearman's $\rho = 0.076$, $p = 0.388$, $N = 132$). A possible interpretation of these two findings is that γ , the weight subjects attach to fairness considerations, may be endogenous to being favored and higher for non-favored subjects. Favored subjects may—consciously or unconsciously—choose to downsize their γ to stronger benefit monetarily from their favored status when deciding to enter the competition. This result has a somewhat similar flavor as that of

¹⁴ By contrast, studies on gender quotas typically do not find that men are less likely to compete if a quota for women is in place (Balafoutas and Sutter, 2012; Niederle et al., 2013; Balafoutas et al., 2016).

Buser et al. (2021) who find that concerns for the chance of winning trump concerns for fairness for most people.

Result 2. *Subjects are more likely to enter the competition when they are favored by affirmative action.*

Result 3. *For all subjects jointly, individual fairness perceptions and willingness to compete are positively, but only weakly correlated. Non-favored subjects are more likely to compete, the fairer they consider an affirmative action policy, while this is not the case for favored subjects.*

Overall, a higher fairness perception of affirmative action seems to increase the share of competing subjects in two ways: a higher *individual* fairness perception predicts non-favored subjects' willingness to compete, while favored subjects seem to be more willing to compete under affirmative action policies with a higher *average* fairness perception. While encouraging favored individuals to enter competitions is a key aim of affirmative action, raising non-favored subjects' willingness to compete is also desirable since competition tends to increase efficiency (see, e.g., Balafoutas et al., 2016, Niederle and Vesterlund, 2007).

Our results thus point at an important aspect in designing and communicating affirmative action policies: their fairness perception. For an affirmative action policy to encourage favored subjects to compete without discouraging non-favored subjects, it is vital that the policy is perceived as fair. From a policy perspective, providing a convincing rationale for the implementation of quota rules to ensure that they are largely perceived as fair seems key to make them a success. An example is providing evidence on discrimination against women when introducing a gender quota (Ip et al., 2020).

4.3 Efficiency

A prominent worry of opponents of affirmative action is that it harms efficiency by not selecting or rewarding the "best". A 50% quota that may replace the second-highest performer as a winner by someone else may seem damaging to efficiency at first glance. But at closer inspection this may only be the case under rather narrow circumstances.

To provide a comprehensive analysis, we offer four different perspectives on efficiency. First, we discuss two measures of efficiency: (i) productivity defined as number of correctly solved grids per minutes in stage 3, not considering differences in performance that arise due to discrimination or working time, and (ii)

output defined as the overall number of correctly solved grids in stage 3 that reflects both productivity and working time, but not the multiplier that we use to introduce discrimination. Moreover, we report results for both measures of efficiency for all subjects and for the winners of the competition only. In our view, it depends on the specific context under consideration which perspective is the most relevant one. Focusing on winners only is appropriate from, e.g., an employer's perspective in a promotion or selection context. Focusing on all individuals informs us about the consequences of various affirmative action policies for the population as a whole, taking into account individuals who decided against or in favor of entering the competition, and, within the latter group, those who lost or won. Output may be the relevant measure of efficiency when a principal like an employer cannot observe an individual's actual time worked and condition rewards on it, productivity if that is feasible.

Hypotheses on efficiency. As our theoretical framework assumes that fairness concerns are not dominant and effort costs are sufficiently small, subjects should exert maximal effort (see Section 3). Consequently, we do not expect average productivity and output to differ across treatments when analyzing all subjects jointly, i.e., irrespective of whether they entered the tournament or not.

Hypothesis 4. *Considering all subjects jointly, average productivity and output in the affirmative action treatments do not change compared to the control treatment.*

Tournament winners' efficiency reflects two separate aspects: how well winners perform and who the winners are. The latter is determined both by subjects' entry decisions and winner selection, which in turn are both affected by the tournament rules (a quota or its absence). We have already analyzed effects on willingness to compete in Section 4.2. Generally, if more subjects compete, winners' ε increases in expectation, leading to slightly higher productivity and output. Encouraging subjects with a low baseline productivity or working time to enter, tends to make them more likely to win, in particular those subjects with a high multiplier, reducing winners' productivity and/or output. As those two effects oppose each other, are small to begin with, and require that the eventual winners would not have competed without an affirmative action rule, we expect the overall effect due to selection into the competition to be limited.¹⁵

The following arguments refer to winner selection for a given composition of tournament entrants. Effects of any affirmative action policy on winners' average

¹⁵ Moreover, our results on willingness to compete show that the differences between treatments are not large, further limiting the potential impact of this channel.

productivity will be small. In general, at most one out of two winners will be replaced by a quota rule. In case of actual replacement of a winner, productivity will mainly be affected if a high baseline productivity subject replaces a low baseline productivity subject and vice versa; the probability for those events is relatively low, though. In AAW and AAD, it tends to require that all three non-favored subjects have high baseline productivity (such that the rule replaces a non-favored high with a favored low baseline productivity subject) or at most one non-favored subject has high baseline productivity (such that the rule likely replaces a non-favored low with a favored high baseline productivity one). Moreover, it requires that the favored winner would not already have won in the absence of a rule. Even though effects under AAP are more direct, they tend to be small as well because those subjects with a relatively high baseline productivity among those classified as low baseline productivity subjects tend to replace subjects with a relatively low baseline productivity among those classified as high baseline productivity subjects.

Regarding winners' average output, the effect of AAP will be limited for the same reason. AAW has a direct, negative, and larger effect on output because a short working time subject is likely to replace a long working time one. AAD's effect on output should be small as it is indirect. If anything, the winners' average output will weakly go up for AAD as a subject with a high multiplier that has the second highest performance (and not the highest one) is likely to have a short working time and/or low baseline productivity and will likely be replaced with a subject that has a long working time and/or higher baseline productivity.

Hypothesis 5. *Winners' average productivity does not change much for any affirmative action rule. Winners' average output does not change much for the affirmative action rules that favor subjects with low productivity or a low multiplier either. Winners' average output decreases for the affirmative action rule that favors subjects with a short working time.*

Results on efficiency. Table 4 presents results on efficiency. In columns (1) and (2), output is the dependent variable that is regressed on dummy variables for each affirmative action treatment, using CTR as a baseline, and individual fairness perception. Columns (3) and (4) refer to productivity. Columns (1) and (3) use observations on winners only, while columns (2) and (4) refer to all subjects' output and productivity.

	(1)	(2)	(3)	(4)
	Output		Productivity	
	Winners	All	Winners	All
AAP	-0.107 (1.687)	-0.350 (0.784)	0.065 (0.244)	0.002 (0.098)
AAW	-3.608** (1.613)	-0.741 (0.790)	-0.123 (0.234)	-0.109 (0.099)
AAD	-0.845 (1.584)	-0.755 (0.818)	-0.036 (0.229)	-0.074 (0.102)
Fairness perception	-0.301 (0.297)	0.071 (0.156)	-0.020 (0.043)	-0.008 (0.019)
Constant	18.582*** (1.347)	12.148*** (0.672)	2.534*** (0.043)	2.039*** (0.084)
Observations	68	372	68	372

OLS regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. For columns (1) and (2), the dependent variable is the number of correctly solved grids in stage 3. For columns (3) and (4), the dependent variable is the number of correctly solved grids per minute in stage 3. The base category is CTR. *Fairness perception* reflects the fairness rating of own treatment, elicited on a 7-point scale on which higher numbers indicate higher perceived fairness.

Table 4 Efficiency in stage 3

In line with [Hypothesis 4](#), we do not observe large or significant differences in output and productivity in the affirmative action treatments compared to the control treatment when considering all subjects jointly in columns (2) and (4).

Result 4. *Considering all subjects jointly, average productivity and output are not affected by introducing affirmative action policies.*

Treatment effects on winners' output and productivity in columns (1) and (3) of [Table 4](#) are well in line with [Hypothesis 5](#). On average, winners in AAW solve 3.6 fewer grids than in the absence of affirmative action ($p = 0.029$) because AAW often replaces a winner with long working time by a short working time one. Importantly, this finding demonstrates that affirmative action favoring individuals with short working time can be detrimental to efficiency by decreasing winners' output. Although there is a tendency of lower output in AAP and AAD, coefficients are small and not significantly different from 0. Similarly, results in column (3) of [Table 4](#) document that winners' productivity does not differ significantly across treatments. Our findings extend those of Balafoutas and Sutter (2012) and Niederle et al.

(2013) that a 50% quota favoring women does not harm efficiency to 50% quotas for discriminated individuals and those with low productivity.

Result 5. *Considering winners only, average productivity and output are not affected by introducing affirmative action policies, with one exception. Winner's output is significantly lower under affirmative action that favors subjects with short working time.*

Finally, [Table 4](#) provides thorough null results on the relation between individual fairness perception and efficiency. This is as expected given that the different quotas do not lead to a large variation in subjects' willingness to compete in our data. In principle, if higher fairness perceptions increased the number of competing subjects, one could expect small positive effects on winners' efficiency as this should lead to slightly more efficient winners on average. However, the fraction of competing subjects does not change much across treatments.

4.4 Post-competition teamwork and retaliation

Teamwork. One might worry that the previous exposure to affirmative action policies may harm post-competition teamwork. We measure teamwork by performance in the slider task (number of correctly solved slider screens) in stage 4, in which a higher individual performance yields equal benefits for all group members. Subjects now act in a new environment with the same monetary incentives for all and without a policy that favors particular individuals.

Hypotheses on teamwork. At the teamwork stage, subjects maximize their payment $\sum_{i=1}^6 \tilde{b}_i \tilde{e}_i$ (where $\tilde{b}_i \geq 0$ and $\tilde{e}_i \in [0, 1]$ are a subject's baseline productivity and effort in the slider task), minus an additional effort cost (with $\tilde{c}'(\tilde{e}_i) > 0$ and $\tilde{c}(0) = 0$), plus a bonus payment of 5 EUR or 2 EUR (depending on whether they won or lost in stage 2, respectively), minus their expected utility loss from fairness considerations (where the allocation vector Y and the fairness ideal $F^{k(i)}$ not only take the payments from the teamwork stage into account but also the ones from the tournament in stage 2). Exerted effort will not be affected by the different treatments as the payoffs from stage 2 cannot be changed anymore and the same pure shifts of absolute payment levels for all team members do not affect the fairness term of subjects' utility function. Note that libertarians and meritocrats may find the mode of compensation in the teamwork task unfair but they do so to the same extent after experiencing different affirmative action policies, and we are interested

	(1)	(2)	(3)	(4)
	All	AAP	AAW	AAD
AAP	-0.145 (0.262)			
AAW	-0.140 (0.263)			
AAD	0.111 (0.273)			
Low type in AA		-0.391 (0.366)	-0.128 (0.409)	0.315 (0.384)
High type in CTR		0.371 (0.333)	0.230 (0.378)	0.065 (0.343)
High type in AA		0.542 (0.350)	0.150 (0.405)	-0.042 (0.384)
Fairness perception	0.021 (0.052)	-0.019 (0.073)	-0.013 (0.077)	0.026 (0.073)
Constant	7.145*** (0.224)	7.069*** (0.309)	7.124*** (0.326)	7.100*** (0.306)
Observations	372	198	198	192

OLS regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The dependent variable is performance (number of correctly solved slider screens) in stage 4. Base categories are CTR in (1) and Low type in CTR in (2), (3), and (4). Wald tests on the equality of coefficients between High type in CTR and High type in AA are not significant for (2), (3), and (4) (all p-values larger than 0.624).

Table 5 Post-competition teamwork: performance in the slider task

in comparisons between treatments. In sum, we do not expect to find any spillovers of previously experienced affirmative action on post-competition teamwork.

Hypothesis 6. *Post-competition teamwork is affected neither by previously experienced affirmative action rules nor by fairness perceptions.*

Results on teamwork. In line with [Hypothesis 6](#), post-competition teamwork does not differ significantly between treatments (Kruskal-Wallis test, $p = 0.440$; see also column (1) of [Table 5](#)). Moreover, post-competition teamwork of any given subject type is not affected by whether a subject was formerly favored or not (see columns (2) to (4) of [Table 5](#)). In line with the expectations discussed above, we do not find the perceived fairness of the formerly experienced tournament rules to affect teamwork in the new environment; see [Table 5](#).

Result 6. *Post-competition teamwork is affected neither by previously experienced affirmative action rules nor by fairness concerns.*

In sum, worries that the previous exposure to affirmative action policies may harm post-competition teamwork seem unwarranted; see Balafoutas et al. (2016), Sutter et al. (2016), and Kölle (2017) for related findings in different affirmative action contexts. This is due to the fact that being less cooperative in teamwork cannot reduce utility losses due to fairness concerns, but comes at the cost of lower earnings. We therefore proceed by studying whether exposure to affirmative action policies may decrease transfers in dictator games where, for example, less altruistic behavior towards previous winners may increase an individual's payoff and, at the same time, may decrease utility losses from fairness considerations.

Retaliation. In contrast to the slider task, in which an individual's behavior affects all other group members equally and individual incentives are aligned with group incentives as long as effort costs are small enough, decisions in stage 5's dictator games allow subjects to treat each individual group member in a more or less favorable way. Subjects can condition their transfer on whether someone was favored or non-favored by the respective affirmative action policy in the stage 2 tournament and a winner or loser. In particular, we can explore whether there is evidence of a further possible, adverse post-competition consequence of affirmative action policies: retaliation towards favored winners. In our setup, we observe retaliation if previously non-favored subjects (in particular losers) transfer lower amounts to previously favored than non-favored winners.

Hypotheses on retaliation. In the dictator game, subjects maximize their payment minus their expected utility loss from fairness considerations, where the allocation vector Y and the fairness ideal $F^{k(i)}$ not only take the payments from the dictator games into account but also the ones from the tournament in stage 2. Clearly, subjects (and in particular non-favored subjects) are less willing to give a positive amount of money to a subject that has already earned more in the tournament than what they consider fair and vice versa. Thus, ceteris paribus, libertarians should give less to favored winners who possibly benefited from any affirmative action rule. Meritocrats and choice egalitarians should do so under the rule that favors subjects with a short working time, and meritocrats under the rule that favors subjects with low baseline productivity.

Hypothesis 7. *The fairer non-favored subjects consider an affirmative action rule, the weakly more they give to favored winners.*

In other words, retaliation should be weakly more pronounced, the unfairer a non-favored individual perceives a quota rule to be. This also means that the un-

fairer a treatment is considered on average, the weakly less non-favored subjects give to favored winners.

Results on retaliation.¹⁶ We start by analyzing the most pointed situation, in which retaliation seems most likely to occur: in [Figure 3](#), we investigate whether—on average and disregarding the role of individual fairness perceptions—non-favored losers give less to favored winners than to non-favored winners in the affirmative action treatments. Comparing transfers to favored and non-favored winners within subject not only controls for the dictator’s identity, but also holds the recipient’s expected earnings other than from the dictator game rather constant. On average, we find no evidence for retaliation under any of the three affirmative action policies (Wilcoxon signed rank tests yield $p = 0.317$ for AAP, $p = 0.906$ for AAW, and $p = 0.198$ for AAD, respectively). As a robustness check, [Figure A.4](#) in the appendix provides additional evidence on the absence of retaliation at the aggregate level based on a broader set of situations that reflect transfers from all non-favored subjects. We do not observe that transfers to favored winners increase monotonically in a treatment’s average fairness rating, i.e., from AAP to AAW to AAD in [Figure 3](#) and [Figure A.4](#), and the signs of the treatment dummies in [Table 6](#) confirm this.

¹⁶ Although we focus on analyzing retaliation, we report average transfers in the dictator games of the different treatments for completeness. Average transfers in the affirmative action treatments range between 101 Cents in AAP and 115 Cents in CTR (Kruskal-Wallis test, $p = 0.925$).

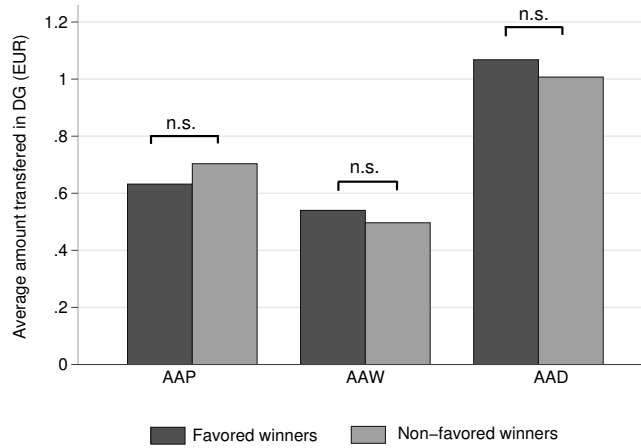


Fig. 3 Retaliation in dictator games

Notes: Average transfer (EUR) from non-favored losers to favored winners versus non-favored winners in dictator games for each affirmative action treatment. AAP: Affirmative action w.r.t productivity ($N = 28$). There are 3 more observations for transfers to favored winners in AAP that we do not report here because for these subjects there were no non-favored winners in their group which makes a within-subject comparison and signed rank tests impossible. Including these yields a slightly changed average transfer of 60 Cents to favored winners in AAP. AAW: Affirmative action w.r.t. working time ($N = 30$). AAD: Affirmative action w.r.t. discrimination ($N = 28$). The brackets and stars above each bar show results of Wilcoxon signed rank tests, * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

	(1)	(2)	(3)	(4)
From	non-favored subjects		non-favored losers	
To	fw	nfw	fw	nfw
AAW	-0.766 (0.470)	-0.557 (0.540)	-0.320 (0.544)	-0.557 (0.540)
AAD	0.303 (0.463)	0.414 (0.538)	0.662 (0.544)	0.414 (0.538)
Fairness perception	0.235** (0.105)	0.192 (0.123)	0.225* (0.125)	0.192 (0.123)
Constant	-0.661 (0.482)	-0.608 (0.534)	-0.935 (0.544)	-0.608 (0.534)
Observations	132	86	89	86

Tobit regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The dependent variable is transfer (EUR) in dictator games. The base category for all models is AAP. Abbreviations: fw – favored winners, nfw – non-favored winners. Note that columns (2) and (4) are identical since there is no group in which a non-favored winner makes a decision about a transfer to another non-favored winner. Hence, giving non-favored subjects and giving non-favored losers coincide in this case.

Table 6 Transfers in Dictator Games

While there is no evidence for retaliation at the aggregate level, subjects that perceive an affirmative action policy as fairer do transfer more to favored winners (see columns (1) and (3) of Table 6)—as predicted by Hypothesis 7. This pattern holds both when considering non-favored subjects as a whole and non-favored losers only. We find the same pattern in simple Spearman correlations between individual fairness perceptions and transfers to favored winners that do not control for treatment differences: $\rho = 0.270$ for all non-favored subjects and $\rho = 0.299$ for non-favored losers, respectively, both $p < 0.01$. Thus, the absence of evidence on retaliation at the aggregate level masks the fact that favored winners are retaliated against by individuals that perceive affirmative action as less fair, while they receive higher transfers if it is perceived as fair. In line with our expectations, fairness concerns are not a significant predictor of transfers to non-favored winners in columns (2) and (4) of Table 6.

Result 7. *The fairer non-favored subjects consider an affirmative action rule, the more they give to favored winners.*

By pointing out the association of individual fairness perceptions of affirmative action policies and post-competition interactions, our results connect previous findings on backlash against favored individuals (e.g., Fallucchi and Quercia, 2018; Leibbrandt et al., 2017) and support for affirmative action policies (Ip et al., 2020).

5 Conclusion

One defining feature of all affirmative action policies is to base hiring, promotion, or admission decisions not on observed performance alone but to complement or adjust observed performance by further criteria. For a given technology, working time, productivity, and discrimination (i.e., how much a given unit of output is valued by others) are the three key determinants of performance (Cappelen et al., 2010). In this sense, we investigate the “whole universe” of possible quota rules: quotas that favor individuals with a short self-chosen working time, with low productivity, or those facing enduring discrimination. One advantage of this stylized approach to study the implications of affirmative action policies is that the motivation to favor certain individuals is clearly stated—quite in contrast to a gender quota, for example, that some will attribute to lower skill levels of women, others to offsetting disadvantages due to part-time work, and still others to discrimination

against women. Explicitly stating the reasons for preferential treatment provides a homogeneous perception of what an affirmative action policy is about and a sound basis for eliciting fairness judgments of such policies that can be attributed to the respective criterion for such preferential treatment.

We are not aware of previous attempts to investigate affirmative action favoring individuals with short working time, although such policies exist in practice. For example, all countries of the European Union provide part-time employees equal access to promotions even if their overall performance is lower since they work shorter hours. Examples for affirmative action in favor of individuals with low productivity are instances in which individuals with dyslexia or physical restraints get extra time in exams. We provide first evidence how affirmative action favoring individuals with low productivity affects willingness to compete, post-competition retaliation, and teamwork.¹⁷

We find that quotas for discriminated individuals and those who have chosen to work shorter meet their main aim: they effectively encourage favored individuals to enter a competition. By contrast, a quota for low productivity individuals does not have such an encouragement effect. Compared to a situation without affirmative action, none of the three affirmative action policies under study harms overall productivity, post-competition teamwork, or induces significant retaliation towards the group of favored individuals as a whole. Only the affirmative action policy that favors subjects with short working time reduces the overall output of tournament winners. Thus, our results largely reinforce the rather positive findings regarding the consequences of affirmative action policies in studies on gender quotas (e.g., Balafoutas and Sutter, 2012; Balafoutas et al., 2016; Ibanez and Riener, 2018; Kölle, 2017; Niederle et al., 2013) or caste membership (Banerjee et al., 2018, 2020) and extend them to affirmative action policies targeting discriminated individuals and those who perform lower since they have chosen to work shorter hours.

Based on judgments regarding their fairness, affirmative action policies receive broad support. In particular, affirmative action policies targeting discriminated individuals or those choosing a short working time are judged as significantly fairer than no affirmative action. In times of heated debates about affirmative action, this is important news as it indicates that quotas can get broad support if they target discriminated individuals or part-time workers and are communicated as such. As fairness perceptions can affect effort provision (Charness and Kuhn, 2007; Cohn

¹⁷ Most closely related is the work of Ip et al. (2020) who prime subjects to perceive women as having lower skills than men to study whether quotas for female managers decrease workers' effort in a gift-exchange game.

et al., 2015) and physical health (Falk et al., 2017) in organizational contexts, our findings underline the importance of fostering a sense of fairness and acceptance regarding managerial decisions on affirmative action. Providing an explicit justification for affirmative action such as disadvantage or discrimination of favored groups seems to be key to achieve broad support for affirmative action policies.

Most importantly, our findings suggest that the perceived fairness of affirmative action policies predicts their consequences. Higher fairness perceptions can encourage willingness to compete and prevent retaliation against favored winners. As a whole, our results point at a so far disregarded, but vital aspect in designing and communicating affirmative action policies, namely their perceived fairness. Providing a convincing rationale for the implementation of quota rules to ensure that they are perceived as fair seems key to making them a success.

References

- Abeler, Johannes, Falk, Armin, Goette, Lorenz, and Huffman, David (2011). "Reference Points and Effort Provision." *American Economic Review*, 101(2), 470–492.
- Ai, Chunrong and Norton, Edward (2003). "Interaction Terms in Logit and Probit Models." *Economics Letters*, 80(1), 123–129.
- Almås, Ingvild, Cappelen, Alexander, Salvanes, Kjell G, Sørensen, Erik, and Tungodden, Bertil (2016). "Willingness to Compete: Family Matters." *Management Science*, 62(8), 2149–2162.
- Araujo, Felipe A., Carbone, Erin, Conell-Price, Lynn, Dunietz, Marli W., Jaroszewicz, Ania, Landsman, Rachel, Lamé, Diego, Vesterlund, Lise, Wang, Stephanie W., and Wilson, Alistair J. (2016). "The Slider Task: An Example of Restricted Inference on Incentive Effects." *Journal of the Economic Science Association*, 2(1), 1–12.
- Bagde, Surendrakumar, Epple, Dennis, and Taylor, Lowell (2016). "Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India." *American Economic Review*, 106(6), 1495–1521.
- Balafoutas, Loukas, Davis, Brent J., and Sutter, Matthias (2016). "Affirmative Action or Just Discrimination? A Study on the Endogenous Emergence of Quotas." *Journal of Economic Behavior and Organization*, 127, 87–98.
- Balafoutas, Loukas, Kerschbamer, Rudolf, and Sutter, Matthias (2012). "Distributional Preferences and Competitive Behavior." *Journal of Economic Behavior and Organization*, 83(1), 125–135. (Gender Differences in Risk Aversion and Competition).
- Balafoutas, Loukas and Sutter, Matthias (2012). "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory." *Science*, 335(6068), 579–582.
- Baltrunaite, Audinga, Bello, Piera, Casarico, Alessandra, and Profeta, Paola (2014). "Gender Quotas and the Quality of Politicians." *Journal of Public Economics*, 118, 62–74.
- Banerjee, Ritwik, Gupta, Nabanita Datta, and Villeval, Marie Claire (2018). "The Spillover Effects of Affirmative Action on Competitiveness and Unethical Behavior." *European Economic Review*, 101(C), 567–604.
- Banerjee, Ritwik, Gupta, Nabanita Datta, and Villeval, Marie Claire (2020). "Feedback Spillovers Across Tasks, Self-confidence and Competitiveness." *Games and Economic Behavior*, 123, 127–170.
- Bartling, Bjorn, Fehr, Ernst, Marechal, Michel Andre, and Schunk, Daniel (2009). "Egalitarianism and Competitiveness." *American Economic Review*, 99(2), 93–98.
- Beaman, Lori, Chattopadhyay, Raghavendra, Duflo, Esther, Pande, Rohini, and Topalova, Petia (2009). "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, 124(4), 1497–1540.
- Beaurain, Guillaume and Masclet, David (2016). "Does Affirmative Action Reduce Gender Discrimination and Enhance Efficiency? New Experimental Evidence." *European Economic Review*, 90, 350–362.

- Bertrand, Marianne, Hanna, Rema, and Mullainathan, Sendhil (2010). "Affirmative Action in Education: Evidence From Engineering College Admissions in India." *Journal of Public Economics*, 94(1), 16–29.
- Bock, Olaf, Baetge, Ingmar, and Nicklisch, Andreas (2014). "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review*, 71(C), 117–120.
- Buser, Thomas, Cappelen, Alexander W, and Tungodden, Bertil (2021). "Fairness and Willingness to Compete." *NHH Dept. of Economics Discussion Paper*, 08.
- Buser, Thomas, Niederle, Muriel, and Oosterbeek, Hessel (2014). "Gender, Competitiveness, and Career Choices." *The Quarterly Journal of Economics*, 129(3), 1409–1447.
- Calsamiglia, Caterina, Franke, Jörg, and Rey-Biel, Pedro (2013). "The Incentive Effects of Affirmative Action in a Real-effort Tournament." *Journal of Public Economics*, 98, 15–31.
- Cappelen, Alexander W., Falch, Ranveig, and Tungodden, Bertil (2020). "Fair and Unfair Income Inequality." *Handbook of Labor, Human Resources and Population Economics*, 1–25.
- Cappelen, Alexander W., Konow, James, Sørensen, Erik Ø, and Tungodden, Bertil (2013). "Just Luck: An Experimental Study of Risk-taking and Fairness." *American Economic Review*, 103(4), 1398–1413.
- Cappelen, Alexander W., Sørensen, Erik Ø, and Tungodden, Bertil (2007). "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review*, 97(3), 818–827.
- Cappelen, Alexander W., Sørensen, Erik Ø, and Tungodden, Bertil (2010). "Responsibility for What? Fairness and Individual Responsibility." *European Economic Review*, 54(3), 429–441.
- Charness, Gary and Kuhn, Peter (2007). "Does Pay Inequality Affect Worker Effort? Experimental Evidence." *Journal of Labor Economics*, 25(4), 693–723.
- Cohn, Alain, Fehr, Ernst, and Götte, Lorenz (2015). "Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment." *Management Science*, 61(8), 1741–2011.
- Corgnet, Brice, Hernán-González, Roberto, and Schniter, Eric (2015). "Why Real Leisure Really Matters: Incentive Effects on Real Effort in the Laboratory." *Experimental Economics*, 18(2), 284–301.
- Council of European Union (1997). "Council Directive 97/81/EC." (<https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:31997L0081>).
- Croson, Rachel and Gneezy, Uri (2009). "Gender Differences in Preferences." *Journal of Economic Literature*, 47(2), 448–474.
- De Paola, Maria, Scoppa, Vincenzo, and Lombardo, Rosetta (2010). "Can Gender Quotas Break Down Negative Stereotypes? Evidence From Changes in Electoral Rules." *Journal of Public Economics*, 94(5), 344–353.
- Disability Rights Commission (2007). "Code of Practice (Revised) for Providers of Post-16 Education and Related Services."
- Estevan, Fernanda, Gall, Thomas, and Morin, Louis-Philippe (2019). "Redistribution Without Distortion: Evidence From an Affirmative Action Programme at a Large Brazilian University." *The Economic Journal*, 129(619), 1182–1220.

- Falk, Armin, Becker, Anke, Dohmen, Thomas, Enke, Benjamin, Huffman, David, and Sunde, Uwe (2018). "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics*, 133(4), 1645–1692.
- Falk, Armin, Kosse, Fabian, Menrath, Ingo, Verde, Pablo E., and Siegrist, Johannes (2017). "Unfair Pay and Health." *Management Science*, 64(4), 1477–1973.
- Fallucchi, Francesco and Quercia, Simone (2018). "Affirmative Action and Retaliation in Experimental Contests." *Journal of Economic Behavior and Organization*, 156, 23–40.
- Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10(2), 171–178.
- Fish, Stanley (2000). "The Nifty Nine Arguments against Affirmative Action in Higher Education." *The Journal of Blacks in Higher Education*, 27, 79–81.
- Frederick, Shane (2005). "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 19(4), 25–42.
- Fullinwider, Robert (2011). "Affirmative Action." In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (editor).
- Gächter, Simon, Huang, Lingbo, and Sefton, Martin (2016). "Combining "Real Effort" With Induced Effort Costs: The Ball-catching Task." *Experimental Economics*, 19(4), 687–712.
- Gill, David and Prowse, Victoria (2012). "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." *American Economic Review*, 102(1), 469–503.
- Girard, Victoire (2020). "Stabbed in the Back? Mandated Political Representation and Murders." *Social Choice and Welfare*, 1–40.
- Gneezy, Uri, Niederle, Muriel, and Rustichini, Aldo (2003). "Performance in Competitive Environments: Gender Differences." *The Quarterly Journal of Economics*, 118(3), 1049–1074.
- Goerg, Sebastian J., Kube, Sebastian, and Radbruch, Jonas (2019). "The Effectiveness of Incentive Schemes in the Presence of Implicit Effort Costs." *Management Science*, 65(9), 3949–4450.
- Greiner, Ben (2015). "Subject Pool Recruitment Procedures: Organizing Experiments With ORSEE." *Journal of the Economic Science Association*, 1, 114–125.
- Heilman, Madeline E., Block, Caryn J., and Stathatos, Peter (1997). "The Affirmative Action Stigma of Incompetence: Effects of Performance Information Ambiguity." *The Academy of Management Journal*, 40(3), 603–625.
- Ibanez, Marcela and Riener, Gerhard (2018). "Sorting Through Affirmative Action: Three Field Experiments in Colombia." *Journal of Labor Economics*, 36(2), 437–478.
- Ip, Edwin, Leibbrandt, Andreas, and Vecchi, Joseph (2020). "How Do Gender Quotas Affect Workplace Relationships? Complementary Evidence from a Representative Survey and Labor Market Experiments." *Management Science*, 66(2), 503–1004.
- Jenselius, Francesca Refsum (2015). "Development from Representation? A Study of Quotas for the Scheduled Castes in India." *American Economic Journal: Applied Economics*, 7(3), 196–220.
- Kölle, Felix (2017). "Affirmative Action, Cooperation, and the Willingness to Work in Teams." *Journal of Economic Psychology*, 62, 50–62.
- Konow, James (2003). "Which is the Fairest One of All? A Positive Analysis of Justice Theories." *Journal of Economic Literature*, 41(4), 1188–1239.

- Leibbrandt, Andreas and List, John A (2018). "Do Equal Employment Opportunity Statements Backfire? Evidence From A Natural Field Experiment On Job-Entry Decisions." *NBER Working Paper*, 25035.
- Leibbrandt, Andreas, Wang, Liang Choon, and Foo, Cordelia (2017). "Gender Quotas, Competitions, and Peer Review: Experimental Evidence on the Backlash Against Women." *Management Science*, 64(8), 3501–3516.
- Leonard, Kenneth, List, John, and Gneezy, Uri (2009). "Gender Differences in Competition: Evidence From a Matrilineal and a Patriarchal Society." *Econometrica*, 77, 1637–1664.
- Mollerstrom, Johanna, Reme, Bjørn-Atle, and Sørensen, Erik Ø (2015). "Luck, Choice and Responsibility—An Experimental Study of Fairness Views." *Journal of Public Economics*, 131, 33–40.
- Niederle, Muriel, Segal, Carmit, and Vesterlund, Lise (2013). "How Costly is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness." *Management Science*, 59(1), 1–16.
- Niederle, Muriel and Vesterlund, Lise (2007). "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics*, 122(3), 1067–1101.
- Niederle, Muriel and Vesterlund, Lise (2011). "Gender and Competition." *Annual Review of Economics*, 3, 601–630.
- Petters, Lea and Schroeder, Marina (2020). "Negative Side Effects of Affirmative Action: How Quotas Lead to Distortions in Performance Evaluation." *European Economic Review*, 130, 103500.
- Raven, John and Raven, C.Jean (2008). "Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics." In: Competency Motivation Project. Chap. General Introduction and Overview: The Raven Progressive Matrices Tests: Their Theoretical Basis and Measurement Model.
- Rotter, Julian B (1966). "Generalized Expectancies for Internal Versus External Control of Reinforcement." *Psychological Monographs: General and Applied*, 80(1), 1–28.
- Schotter, Andrew and Weigelt, Keith (1992). "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results." *The Quarterly Journal of Economics*, 107(2), 511–539.
- Schupp, Jürgen and Gerlitz, Jean-Yves (2008). "BFI-S: Big Five Inventory-SOEP." *In Zusammenstellung sozialwissenschaftlicher Items und Skalen. ZIS Version. 12.*
- Sutter, Matthias and Glätzle-Rützler, Daniela (2015). "Gender Differences in the Willingness to Compete Emerge Early in Life and Persist." *Management Science*, 61(10), 2339–23354.
- Sutter, Matthias, Glätzle-Rützler, Daniela, Balafoutas, Loukas, and Czermak, Simon (2016). "Cancelling out Early Age Gender Differences in Competition: An Analysis of Policy Interventions." *Experimental Economics*, 19(2), 412–432.
- Van Veldhuizen, Roel (2022). "Gender Differences in Tournament Choices: Risk Preferences, Overconfidence, or Competitiveness?" *Journal of the European Economic Association*, 20(4), 1595–1618.
- Wagner, Gert G, Frick, Joachim R, and Schupp, Jürgen (2007). "The German Socio-Economic Panel Study (SOEP) - evolution, scope and enhancements." *SOEP papers on Multidisciplinary Panel Data Research*, (1).

Appendix

A Additional results

A.1 Variation in productivity and working time

Figure A.1 and Figure A.2 show the distributions of productivity and working time for all 463 subjects participating in the preparatory session. The median productivity is 7 in the DICE Lab sample and 8 in the BonnEconLab sample. The median working time is 17 min 46 s in the DICE Lab sample and 17 min 17 s in the BonnEconLab sample. In Figure A.2, the spike at 61 minutes is due to the fact that we stopped subjects who still worked on the grid task after 60 minutes. Those who work longer are classified as subjects with long working time anyway and any further measurement of their chosen working time is not necessary.

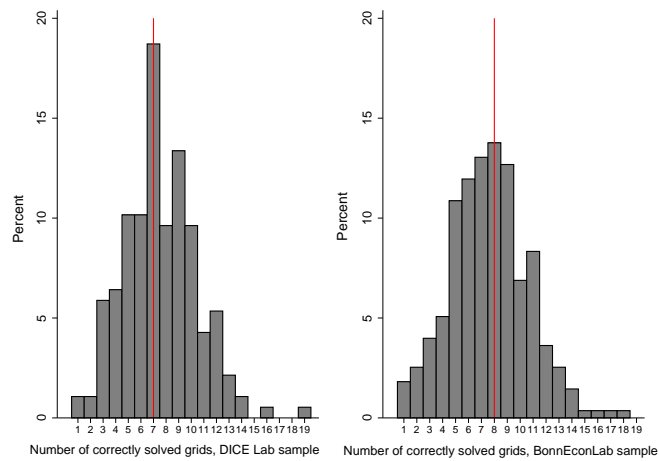


Fig. A.1 Distribution and median split (indicated by the vertical red line) of productivity by sample

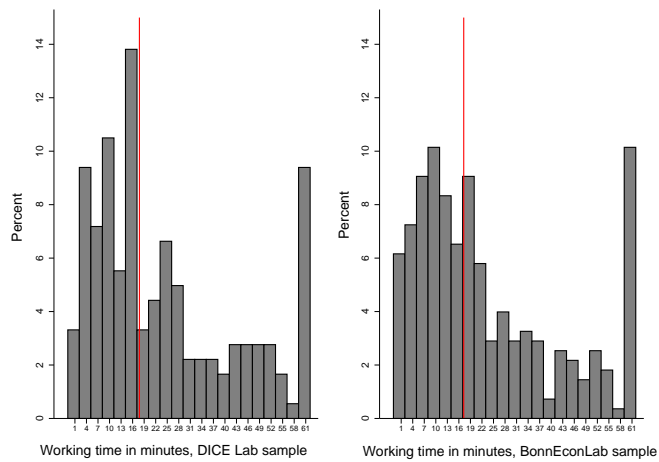


Fig. A.2 Distribution and median split (indicated by the vertical red line) of working time by sample

A.2 Fairness perception

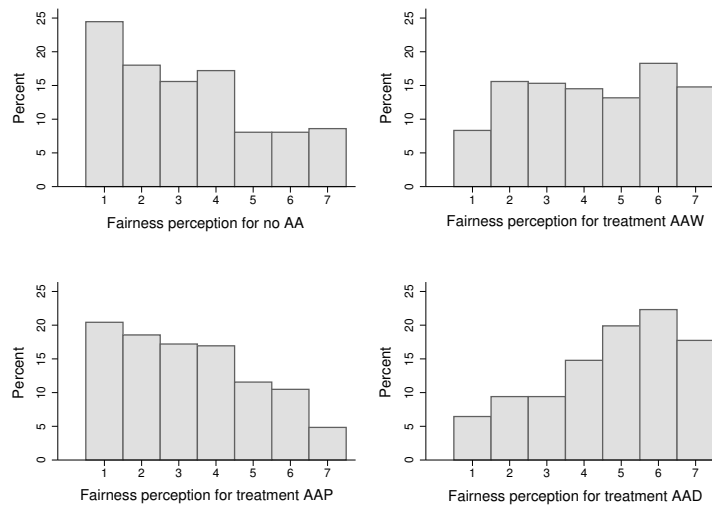


Fig. A.3 Distribution of fairness perception

Notes: Perceived fairness of different policies elicited on a 7-point Likert scale, answering the item "How fair do you perceive the rules of the competition to be?". Higher numbers indicate that a policy is perceived as fairer. No AA (CTR): No affirmative action. AAP: Affirmative action w.r.t. productivity. AAW: Affirmative action w.r.t. working time. AAD: Affirmative action w.r.t. discrimination. $N = 372$.

	(1)
Avg. fairness perception of	
AAP	0.065 (0.129)
AAW	0.976*** (0.138)
AAD	1.452*** (0.134)
Constant	3.250*** (0.100)
Observations	1488

OLS regressions, standard errors in parentheses, clustered at subject level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The dependent variable is perceived fairness of different policies elicited on a 7-point Likert scale, answering the item “How fair do you perceive the rules of the competition to be?”. Higher numbers indicate that a policy is perceived as fairer. The base category is fairness judgment of CTR. Wald tests on the equality of the coefficients yield $p < 0.001$ for all pairwise comparisons among AAP, AAW, and AAD.

Table A.1 Fairness perception - Treatment comparison

A.3 Willingness to compete

	(1)	(2)	(3)
Belief on rank			
2	-0.080 (0.083)	-0.127 (0.096)	-0.104 (0.096)
3	-0.312*** (0.089)	-0.340*** (0.105)	-0.333*** (0.105)
4	-0.313*** (0.090)	-0.354*** (0.102)	-0.337*** (0.101)
5	-0.393*** (0.098)	-0.497*** (0.116)	-0.504*** (0.116)
6	-0.504*** (0.112)	-0.580*** (0.129)	-0.591*** (0.135)
Risk attitude	0.028** (0.011)	0.029** (0.013)	0.098** (0.013)
Female	-0.094* (0.050)	-0.092 (0.058)	-0.093 (0.058)
Fairness perception	0.016 (0.013)	0.016 (0.016)	0.048** (0.022)
Favored		0.215*** (0.063)	0.476*** (0.136)
Favored × Fairness			-0.067** (0.031)
Constant	0.546*** (0.104)	0.472*** (0.118)	0.347*** (0.131)
Observations	372	264	264

Robustness check for specifications (1), (2), (3) of Table 3. OLS regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The binary dependent variable is willingness to compete in stage 3 (1 if tournament, 0 if piece-rate). The base category is non-favored subjects in (2) and (3); base category for Belief on rank is 1. *Belief on rank* is a categorical variable about beliefs regarding own rank (between 1 and 6) in stage 3; *Risk attitude* is the answer to the general risk question elicited on an 11-point scale on which higher numbers indicate a higher willingness to take risks; *Female* is an indicator variable for gender (1 if female, 0 if male); *Fairness perception* reflects fairness rating of own treatment, elicited on a 7-point scale on which higher numbers indicate higher perceived fairness; *Favored* is an indicator variable (1 if favored, 0 otherwise). Due to the inclusion of the interaction effect in (3), we estimate OLS instead of a non-linear model (Ai and Norton (2003)).

Table A.2 Willingness to compete

A.4 Efficiency

	(1)	(2)	(3)	(4)	(5)	(6)
	Output			Productivity		
	AAP	AAW	AAD	AAP	AAW	AAD
Low type in AA	-0.360 (1.046)	-0.605 (0.863)	-1.606 (1.141)	-0.019 (0.127)	-0.144 (0.138)	-0.132 (0.141)
High type in CTR	4.830*** (0.951)	7.996*** (0.797)	1.667 (1.019)	0.702*** (0.115)	0.321** (0.128)	0.085 (0.125)
High type in AA	4.289*** (0.999)	7.324*** (0.855)	0.918 (1.141)	0.713*** (0.121)	0.228* (0.137)	0.005 (0.141)
Fairness perception	0.186 (0.210)	-0.025 (0.163)	0.366* (0.218)	-0.002 (0.025)	0.001 (0.026)	0.014 (0.027)
Constant	9.420*** (0.882)	8.413*** (0.687)	10.508*** (0.909)	1.671*** (0.107)	1.853*** (0.110)	1.936*** (0.112)
Observations	198	198	192	198	198	192

OLS regressions, standard errors in parentheses. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The dependent variable in columns (1) to (3) is the number of correctly solved grids in stage 3. The dependent variable in columns (4) to (6) is the number of correctly solved grids per minute in stage 3. The sample is the AA treatment stated in the third row and CTR. The base category is the Low type of the respective AA treatment in CTR. Wald tests on the equality of coefficients of High type in CTR and High type in AA are not significant for any specification (all p-values larger than 0.423). *Fairness perception* reflects fairness rating of own treatment, elicited on a 7-point scale on which higher numbers indicate higher perceived fairness.

Table A.3 Efficiency in stage 3 of all subjects; split by type

A.5 Post-competition measures

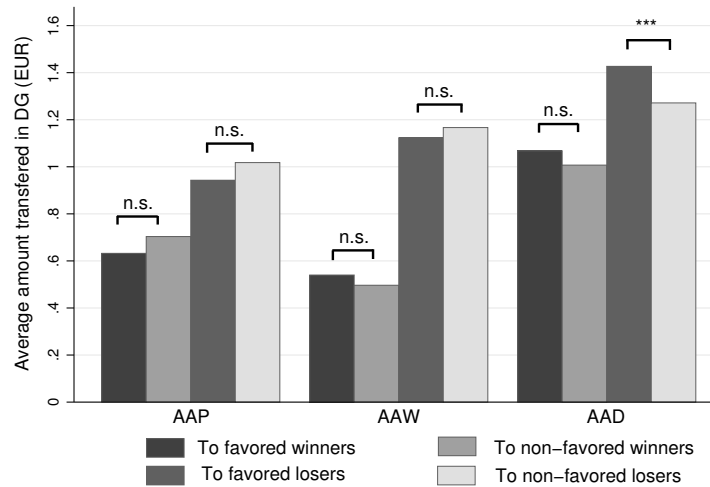


Fig. A.4 Retaliation in dictator games - a broader view

Notes: Average transfer amount in EUR from non-favored subjects to favored winners versus non-favored winners as well as to favored losers versus non-favored losers in the dictator games for each affirmative action treatment. AAP: Affirmative action w.r.t productivity ($N = 45$). AAW: Affirmative action w.r.t. working time ($N = 45$). AAD: Affirmative action w.r.t. discrimination ($N = 42$). The brackets and stars above each bar show results of Wilcoxon signed rank tests: * $p < 0.1$; ** $p < 0.05$, *** $p < 0.01$.

A.6 Power analyses

In this subsection we report power calculations for tests underlying our main results. In all power calculations, we follow standard conventions and fix $\alpha = 0.05$ and power to 80%. We are conservative by assuming two-sided tests and, whenever possible, use the control group as benchmark. The power calculations were run ex post such that we rely on actually realized means, standard deviations, and group sizes to obtain information on detectable effect sizes.

Section 4.1 (Fairness perceptions of affirmative action).

Figure 2: Wilcoxon signed rank tests for all pairwise policy comparisons. At conventional levels of power, we can detect changes in the mean of at least 0.33 to 0.39 points (on a 7-point scale), depending on the exact pairwise comparison. This corresponds to 10%-12% of the mean in CTR of 3.25.

Section 4.2 (Willingness to compete).

Results: aggregate level. Ex-post power calculations reveal that we can only detect effect sizes of at least 19.8pp to 20.2pp, depending on the exact pairwise treatment comparison. As observed differences in willingness to compete between CTR and

the other treatments are small throughout (ranging between 0.9pp and 6.5pp in absolute value), our findings still rule out that the different affirmative policies induce large changes in average willingness to compete.

Section 4.3 (Efficiency).

Table 4. We can detect differences in productivity of at least 0.26-0.28 grids per minute for all subjects and at least 0.55-0.85 grids for winners (depending on which AA treatment we compare to CTR). That corresponds to 13%-14% of the CTR group mean for all subjects (2.02) and 22%-34% of the CTR group mean for winners (2.47). For output, we can detect differences of at least 2.16-2.29 grids for all subjects and at least 3.64-5.46 grids for winners. This corresponds to about roughly 18% for the CTR group mean for all subjects (12.34) and 21%-31% of the CTR group mean for winners (17.65). Actually observed differences both for productivity and output as well as for all subjects and winners only are at most about 7% of the respective control group mean and typically much smaller. Therefore, our data allow us to rule out that the different affirmative action policies have economically large effects on efficiency.

Section 4.4 (Post-competition teamwork and retaliation).

Results on teamwork. For pairwise comparisons of the affirmative action treatments to CTR, we are able to detect effect sizes between 0.70 and 0.79 (roughly a 10% performance change compared to the control group mean of 7.20). Actually observed differences are always smaller than 2% of the control group mean, implying that the different affirmative action policies do not have economically meaningful effects on teamwork.

Results on retaliation. Focusing on transfers from non-favored losers to favored as opposed to non-favored winners within a treatment, we have the power to detect differences in the range of 49-96 Cents (depending on the exact treatment). This corresponds to a 10-18% change on the overall scale. Please note that there is no control group mean that we can compare these effect sizes to since there are no favored or non-favored individuals in the control treatment. As can be seen in [Figure 3](#), actually observed differences are small (between -4 Cents in AAW and +10 Cents in AAP). Non-favored losers even give a little more to favored than non-favored winners in two of the three affirmative action treatments. Both findings clearly contrast the idea of strong retaliation.

Online Appendix

B Measurement of fairness perception

The following text is translated from German and refers to the control treatment (CTR). In all treatments, subjects were first asked how fair they perceive the rules of competition they had actually been exposed to to be, before the other competition rules were described and rated.

In the following, we would like to know how fair you perceived the rules of competition in PART 2 of the experiment to be.

As a reminder, in a group of six members, the two members with the highest overall performance (that is, the number of correct answers \times the multiplier) were the two winners of the competition.

How fair do you perceive the rules of the competition to be? The leftmost box means "completely unfair", the rightmost means "completely fair". With the boxes in between you can graduate your statement.

completely unfair *completely fair*

Now we would like to know, how fair you perceive other possible rules for the competition to be. Just like in the competition you have participated in in PART 2, the following applies in all competition rules: in a group of 6 members, there are two winners who earn a positive amount of money. The other group members earn nothing. With regard to productivity, working time, and multiplier, the group composition is exactly the same as described on page 2 of the instructions.

Each form of competition has one additional special rule:

- Special rule A: At least one winner must be a group member whose productivity lies in the lower half in comparison to all other participants in the experiment. Productivity is defined as the number of correctly solved tasks per minute of a participant in the five-minute counting task last week.
- Special rule B: At least one winner must be a group member whose working time last week was in the lower half in comparison to all other participants in the experiment, and who therefore worked for 4.5 minutes on the task today.
- Special rule C: At least one winner must be a group member with the low multiplier of 0.75.

How fair do you perceive the competition with special rule A? The leftmost box means "completely unfair", the rightmost means "completely fair". With the boxes in between you can graduate your statement.

completely unfair *completely fair*

How fair do you perceive the competition with special rule B?

completely unfair *completely fair*

How fair do you perceive the competition with special rule C?

completely unfair *completely fair*

C Experimental instructions

C.1 Preparatory session

The preparatory session started with an on-screen description of the grid task and an unpaid trial round of the grid task (including feedback on whether each table was solved correctly). We then measured baseline productivity, implemented the questionnaire (for details, see Online [Appendix D](#)), and finally measured the individual choice of working time. Below we provide translated versions of the instructions that were originally in German.

Measurement of baseline productivity: On-screen instructions for the five-minute grid task

You will now start working on the task. Your performance in this task is relevant for your payment. The more counting tasks you solve correctly in the given time, the higher your payment. For each correctly solved table, you receive 0.50 EUR. In the upper right corner of the screen, the remaining time (in seconds) is shown. The task lasts 5 minutes.

Please try hard to solve as many tables correctly as possible in the five minutes, so that we get a realistic idea how good you are in this task.

Measurement of choice of working time: On-screen instructions for the grid task in which subjects choose their working time

You will now again work on a similar task as the previous one. As before, you count the number of zeros (“0”) in each table and receive 0.1 EUR for it, i.e., 10 Euro-cent for each correct table.

However, you are now free to choose how long you like to work on the task. You will start working on the task on the next screen, and can work on it as long as you want.

The tables will appear one after another, until you decide to stop working.

In addition, there is a special feature: Your working time today will also determine your working time in the experiment next week. Next week, you will work on a similar task again in which you will be given a specific amount of time to solve as many tables as possible and get paid accordingly. In this task, the tables will appear one after another until working time is up.

Based on your working time today, we will form two groups of the same size but with different working times. Those who decide to work shorter today, will also work shorter next week (one half of the participants). Those who decide to work longer today, will also work longer next week (the other half of the participants). Individuals who work shorter will, on average, solve fewer tables correctly and therefore earn less. For them, the experiment will be shorter (it will end earlier). Individuals who work longer will, on average, solve more tables correctly and therefore earn more. For them, the experiment will be longer (it will end later).

In case you have any questions, please raise your hand. An experimenter will then come to your seat.

C.2 Main session

General instructions (distributed on paper at the beginning of the main session)

Welcome to today's experiment! Thank you for participating!

During the experiment, you and the other participants will be asked to make decisions. Your own decisions as well as the decisions of the other participants will determine your earnings, according to the rules that will be described in what follows.

The experiment will be conducted on the computer. You make your decisions on the screen. All your decisions and answers will remain confidential and anonymous.

The experiment consists of 5 parts. PART 1, PART 2, PART 3, PART 4, and PART 5. Additionally, you will answer a short questionnaire.

One of the five parts will be selected randomly by the computer to determine your payment. Every part of the experiment is equally likely to be selected. It is therefore in your own interest to make your decisions in each part as if it was the only part. Independent of your decisions you will receive a show-up fee of 4 EUR. This means that your total earnings from today's session will be the payment from the randomly chosen one of the five parts of the experiment plus the show-up fee of 4

EUR. You will receive your earnings at the end of today's session together with the earnings from last week.

All other explanations will be given stepwise at the beginning of each part of the experiment. You will receive the instructions for each part in turn. You will have enough time to read the instructions carefully and to ask questions. Please do not hesitate to ask questions if something is unclear.

Please note that, as the last week, talking is not permitted. If you have questions, please do not ask them loudly but raise your hand. One of the experimenters will come to your seat to answer your question. If you do not comply with these rules you will be excluded from the experiment and you will not receive any payments.

General information regarding today's experiment

In today's experiment, your task is once again to solve as many counting tasks correctly as possible in a given amount of time, i.e., to correctly count the number of zeros ("0") in as many tables as possible. In addition, there is one special feature. **Each participant has three characteristics which remain fixed during the whole experiment: his productivity, his working time (in minutes) and his multiplier.**

- The **productivity** states how many counting tasks per minute the participant has solved correctly in last week's five-minutes-task. For half of the participants, productivity lies **in the lower half**. For the other half of participants, productivity lies **in the upper half**.
- Today's **working time** depends on the self-chosen working time in the task at the end of the session last week. Half of the participants will have **4.5 minutes** per task today to solve as many tables correctly as possible. These are those participants whose working time belonged to the lower half last week. The other half of the participants will have **7.5 minutes** per task today to solve as many tables correctly as possible. These are those participants whose working time belonged to the upper half last week.
- The **multiplier** is a number which is multiplied with the number of correctly solved counting tasks to determine overall performance. The multiplier will be assigned randomly to each participant. For half of the participants, the multiplier will be **0.75**. For the other half, it will be **1.25**.

You will soon receive information about your productivity, your working time and your multiplier on the following screen.

The performance of each participant is determined as follows:

$$\text{Performance} = \text{Number of correctly solved counting tasks in your working time} \times \text{Multiplier}$$

PART 1 – Piece rate (distributed on paper at the beginning of stage 1)

Your task in PART 1 is similar to the one in the first session. Again, the task is to solve as many counting tasks as possible in a given amount of time, i.e., to correctly count the number of zeros (“0”) in as many tables as possible. How much time you have is displayed on the screen. Each table consists of ten rows and ten columns, which contain either a zero (“0”) or a one (“1”). Each table differs from the previous one. You are allowed to use the provided scratch paper if you like. After you have entered your response, please click the “confirm” button. Afterwards, you will learn immediately on the same screen whether your answer is right or wrong.

If PART 1 of the experiment is chosen for payment, you will receive the following payment:

$$\text{Payment} = \underbrace{\text{No. of correctly solved counting tasks in your working time} \times \text{Multiplier}}_{\text{Overall performance}} \times 0.50 \text{ EUR}$$

For example, if you have solved **ten** tables correctly and your multiplier is **1.25**, you receive the following payment:

$$\text{Payment} = 10 \times 1.25 \times 0.50 \text{ EUR} = 6.25 \text{ EUR}$$

If you have answered **ten** questions correctly and your multiplier is **0.75**, you receive the following payment:

$$\text{Payment} = 10 \times 0.75 \times 0.50 \text{ EUR} = 3.75 \text{ EUR}$$

Your payment will not be reduced if you enter a wrong answer. We will refer to this payment as the **piece-rate payment** from now on.

After all questions regarding PART 1 are answered, your working time for PART 1 will start.

PART 2 – Tournament (distributed on paper at the beginning of stage 2)

As in PART 1, you will have a given amount of time to solve as many counting tasks correctly as possible. Again, your working time is displayed on the screen. Different from before, in this part your payment depends on your performance compared to the performance of other participants in your group.

Group allocation:

For the following parts of the experiment, you will be allocated to a **group with 6 members**. The groups were formed randomly and stay the same throughout the whole experiment. This means that you will form a group with the same participants for the rest of the experiment.

Reminder: **Each participant has 3 characteristics: his productivity, his working time, and his multiplier.**

Note that each group consisting of six **members** meets the following criteria regarding productivity, working time and multiplier:

- The **productivity** of **three** group members lies in the **upper half** compared to all participants. The productivity of the **other three** group members lies in the **lower half** compared to all participants.
- The chosen **working time** last week of **three** group members lies in the **upper half** compared to all participants. Therefore, these three group members work for **7.5 minutes** on each counting task today. The chosen **working time** last week of the **other three** group members lies in the **lower half** compared to all participants. Therefore, these three group members work for **4.5 minutes** on each counting task today.
- The randomly drawn **multiplier** of **three** group members is **0.75**. The number of correctly solved tables of these three group members will thus be multiplied with **0.75** to calculate overall performance. The randomly drawn **multiplier** of the **other three** group members is **1.25**. The number of correctly solved tables of these three group members will thus be multiplied with **1.25** to calculate overall performance.

Rules of the tournament:

If PART 2 is chosen for payment, your payment depends on how high your performance is compared to the other five members of your group.

The **two** group members with **highest** overall performance (i.e., number of correctly solved tasks in the total individual working time × Multiplier) are the two winners of the tournament.

(The content of the following part in gray differs across treatments. There is no further content for the control treatment (CTR).)

Affirmative action w.r.t. productivity (AAP):

In addition, the following **special rule** is applied:

At least one winner must be a group member whose productivity lies in the lower half in comparison to all other participants in the experiment. Productivity is the number of correctly solved counting tasks per minute last week.

If this is not automatically the case given the overall performance of the group members, then the group member with the best performance among the three group members whose productivity lies in the lower half will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

Affirmative action w.r.t. working time (AAW):

In addition, the following **special rule** is applied:

At least one winner must be a group member whose working time last week lied in the lower half in comparison to all other participants in the experiment, and who therefore works for 4.5 minutes on the task today.

If this is not automatically the case given the overall performance of the group members, then the group member with the best performance among the three group members whose working time is 4.5 minutes will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members is no longer a winner.**

Affirmative action w.r.t. discrimination (AAD):

In addition, the following **special rule** is applied:

At least one winner must be a group member with the low multiplier of 0.75.

If this is not automatically the case given the overall performance of the group members, then the group member with the highest performance of the three group members with the low multiplier of 0.75 will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members is no longer a winner.**

The payment of the two winners is as follows:

$$\text{Payment} = \underbrace{\text{Correctly solved counting tasks in their working time} \times \text{Multiplier}}_{\text{Overall performance}} \times 1.50 \text{ EUR}$$

For example, a winner with 10 correct answers and a multiplier of **1.25** receives the following payment:

$$\text{Payment} = 10 \times 1.25 \times 1.50 \text{ EUR} = 18.75 \text{ EUR}$$

A winner with 10 correct answers and a multiplier of **0.75** receives the following payment:

$$\text{Payment} = 10 \times 0.75 \times 1.50 \text{ EUR} = 11.25 \text{ EUR}$$

The **other four** members of your group get **no payment**.

If there is a tie between two group members, the winner will be determined randomly. We will refer to this payment as **tournament payment** from now on. At the end of today's session, you will be informed about the outcome of the tournament.

PART 3 – Choice between piece-rate and tournament payment

(distributed on paper at the beginning of stage 3)

Similar to PART 1 and PART 2, you will have a given amount of time to solve as many counting tasks correctly as possible. Your working time will be shown on the screen. However, now you choose by yourself which payment scheme you prefer for your performance in PART 3. You can choose either the piece-rate payment (same rules as in PART 1) or the tournament payment (same rules as in PART 2).

If PART 3 is chosen for payment, your earnings will be determined as follows:

- If you choose the **piece-rate payment**, your payment is:

$$\text{Payment} = \underbrace{\text{No. of correctly solved counting tasks in your working time} \times \text{Multiplier}}_{\text{Overall performance}} \times 0.50 \text{ EUR}$$

- If you choose the **tournament payment**, your earnings depend on the level of your overall performance in PART 3 compared to the overall performance of your five group members in PART 2 (tournament). Reminder: PART 2 is the part you have just finished.

(The content of the following part in gray differs across treatments.)

Control treatment (CTR):

If your overall performance (i.e., number of correctly solved counting tasks in the individual working time \times Multiplier) is higher than that of at least four other members of your group in PART 2, your payment is as follows:

Affirmative action w.r.t. productivity (AAP):

In general, the two group members with the highest overall performance, (i.e., number of correct answers in the total individual working time) \times (Multiplier), are the two winners of the competition.

The following **special rule** is still applied:

At least one winner must be a group member whose productivity lies in the lower half in comparison to all other participants in the experiment. Productivity is the number of correctly solved counting tasks per minute last week.

If this is not automatically the case given the overall performance of the group members, then the group member with the best performance among the three group members whose productivity lies in the lower half compared to all participants in the experiment will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members of your group is no longer a winner.**

If your overall performance in PART 3 compared to the overall performance of your group members in PART 2 implies you are a winner, your payment is as follows:

Affirmative action w.r.t. working time (AAW):

In general, the two group members with the highest overall performance, (i.e., number of correct answers in the total individual working time) \times (Multiplier), are the two winners of the competition.

The following **special rule** is still applied:

At least one winner must be a group member whose working time last week lied in the lower half in comparison to all other participants in the experiment, and who therefore works for 4.5 minutes on the task today.

If this is not automatically the case given the overall performance of the group members, then the group member with the best performance among the three group members whose working time is **4.5 minutes** will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members is no longer a winner.**

If your overall performance in PART 3 compared to the overall performance of your group members in PART 2 implies you are a winner, your payment is as follows:

Affirmative action w.r.t. discrimination (AAD):

The following **special rule** still is applied: **At least one winner must be a group member with the low multiplier of 0.75.**

If this is not automatically the case given the overall performance of the group members, then the group member with the highest performance of the three group members with the low multiplier of 0.75 will **replace** the initial second-best winner. **In this case the group member with the second highest performance of all six group members is no longer a winner.**

If your overall performance in PART 3 compared to the overall performance of your group members in PART 2 implies you are a winner, your payment is as follows:

$$\text{Payment} = \frac{\text{No. of correctly solved counting tasks in your working time} \times \text{Multiplier}}{\text{Overall performance}} \times 1.50 \text{ EUR}$$

That means it is three times as high as the piece-rate payment.

If your overall performance in PART 3 compared to the overall performance of the other group members in PART 2 implies that **you are not a winner**, you get **no payment**.

If there is a tie between two group members, the winner will be randomly determined.

The group composition is the same as in PART 2. If you choose the tournament payment, you will be informed about the outcome of the tournament at the end of the experiment.

On the next screen, you will decide whether you choose the piece-rate payment or the tournament payment for your performance in PART 3. Then the task will begin.

PART 4 (displayed on screen at the beginning of stage 4)

In the following, you will work on a new task in which you have to place slider markers in a certain position.

You will see six sliders on each screen. They can be placed on a scale from 0 to 100. As soon as you click on a slider marker, the current position will be displayed on the screen. You can change the position using the mouse.

Your task is to move all six slider markers on a screen to the position of “50”. Only then a screen is finished correctly and you can proceed to the next screen by clicking the “Continue” button. You have five minutes to correctly finish as many screens as possible. In this task, all participants work for the same amount of time and there is no multiplier.

Your payment in this part depends on the number of screens that you and the other five members of your group finish correctly. The group composition is the same as before.

Precisely, your payment is determined as follows: You will receive 10 Euro-cent for each correctly finished screen by each member of your group (including yourself). The other members of your group will also receive 10 Euro-cent for each screen that

any group member (including yourself) has finished correctly. This means each correctly finished screen by each player yields 60 Euro-cent for the group (i.e. all six group members together).

In addition, the members of the group who won the tournament in PART 2, will receive an endowment of 5 EUR. The other members will receive an endowment of 2 EUR.

If PART 4 is chosen for payment, your payment is the sum of your individual endowment and your earnings from the sum of all correctly finished screens of your group members.

At the end of the experiment, you will be informed about the performance of your group.

If you have any questions, please raise your hand.

PART 5 (displayed on screen at the beginning of stage 5)

In this part, you are asked to make five decisions which will affect you and one of the five other members of your group, respectively. In order to be able to attribute decisions, each group member will be randomly assigned a number from 1 to 6. You are group member number X.¹⁸

For each decision, you will get an initial endowment of 5 EUR. Your task is to decide how to split this endowment between you and the other member of your group. You may choose an amount between 0 and 5 EUR (in steps of 10 Euro-cents) which you want to pass on to the other group member. You will keep the rest for yourself. You will not get any information about the identity of the other group member and the other group member will not get any information concerning your identity. The only thing you will get to know about the respective other group member before you will make your decision is whether (s)he has won the tournament in PART 2 or not.

(The content of the following part in gray differs across treatments. There is no further content for the control treatment (CTR).)

¹⁸ The exact number differs for each subject.

Affirmative action w.r.t. productivity (AAP):

... and whether his productivity lies in the upper or in the lower half.

Affirmative action w.r.t. working time (AAW):

... and his working time.

Affirmative action w.r.t. discrimination (AAD):

... and his multiplier.

If this part is chosen for payment, your payment will be determined as follows: In each group, three pairs are chosen randomly and their decisions will determine payments. This means each group member is assigned to exactly one pair which is relevant for the payment. In each pair, it is randomly assigned who will be the donor and the recipient. The decision of the donor determines the payment of both. This means that each decision is paid out with the same probability and therefore you should make each decision as if it was the only one.

At the end of the experiment, you will be informed about the number of the group member you paired with, who is the donor and who is the recipient in this pair, and what the donor has decided. You will not get any information about the decisions made in the other pairs (that you do not belong to).

If you have any questions, please raise your hand.

D Questionnaire

The questionnaire in the preparatory session contains the following items:

1. **Risk preference, general risk question:** same wording as in German Socio-Economic Panel questionnaire (SOEP, see, for example, Wagner et al. (2007))

How do you evaluate yourself? Are you generally a risk-seeking person or do you try to avoid risks? The leftmost box means "not at all risk-seeking" and the rightmost "very risk-seeking". With the boxes in between, you can graduate your statement.

not at all risk-seeking *very risk-seeking*

2. **Risk preference, incentivized choice list:** Subjects make eleven, pairwise decisions between a lottery with a fifty-fifty chance of winning either 2 EUR or 7 EUR and a safe payment. The safe payment increases in 0.5 EUR increments, ranging from 2 EUR to 7 EUR.

3. **Social preference** (survey question, Falk et al., 2018)

Question 1: Imagine the following situation: Today you unexpectedly received 1000 EUR. How much of this amount would you donate to a good cause? (Values between 0 and 1000 are allowed).

Question 2: Please think about what you would do in the following situation. You are in an area you are not familiar with, and you realize that you lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 EUR in total. However, the stranger says he or she does not want any money from you. You have six presents with you. The cheapest present costs 5 EUR, the most expensive one costs 30 EUR. Do you give one of the presents to the stranger as a “thank you” gift?

Which present do you give to the stranger?

1. No, would not give present
2. The present worth 5 EUR
3. The present worth 10 EUR
4. The present worth 15 EUR
5. The present worth 20 EUR
6. The present worth 25 EUR
7. The present worth 30 EUR

4. **Big Five:** we use the 15-item Big Five scale developed for the SOEP (Schupp and Gerlitz, 2008) to measure personality traits.

5. **Locus of control:** we use 10 different items adapted from Rotter (1966) which have been used in the 2005 wave of the SOEP.

6. **Questions on general fairness ideals:** all using the same scale

completely disagree *completely agree*

To what extent do you personally agree with the following statements?

It is unfair for someone who does a strenuous job to earn little.

Who performs better, should earn more.

If someone is naturally good at something, it is right to reward him/her for it.

It is wrong to favor somebody just because he/she may have experienced discrimination elsewhere.

7. **Cognitive ability:** Raven matrices from the Wechsler IQ test (J. Raven and C. Raven, 2008).

Before subjects start the test, we elicit their belief about individual rank as follows:

Before you begin, we would like to ask you to assess how well you will score in the IQ test compared to the other participants in the experiment. For example, 0-10% means that you are among the 0-10% participants with the fewest correct answers, and at least 90% of the participants have more correct answers than you. 41-50% means that at least 40% of participants have fewer correct answers and at least 50% have more correct answers than you. 91-100% means that at least 90% have fewer correct answers than you. What do you think? How do you compare to the rest of the group?

- 0-10%
- 11-20%
- 21-30%
- 31-40%
- 41-50%
- 51-60%
- 61-70%
- 71-80%
- 81-90%
- 91-100%

8. **Cognitive reflection test:** see Frederick (2005).
9. **Socio-demographics:** age, gender, final grade point average at academic high school, last math grade at academic high school, field of study, monthly disposable amount of money, political orientation, number of experiments already participated in the same lab.

E Experimental instructions in German

E.1 Preparatory session

Measurement of baseline productivity: On-screen instructions for the five-minute grid task

Sie beginnen nun mit der Aufgabe. Ihre Leistung in dieser Aufgabe ist für Ihre Auszahlung relevant. Je mehr Zählaufgaben Sie in der vorgegebenen Zeit korrekt lösen, desto höher ist Ihre Auszahlung. Pro richtig ausgezählter Tabelle erhalten Sie 0,50 EUR. In der rechten oberen Ecke Ihres Bildschirms wird Ihnen die verbleibende Zeit (in Sekunden) angezeigt. Die Aufgabe dauert 5 Minuten.

Bitte strengen Sie sich an, in den 5 Minuten so viele Tabellen wie möglich korrekt auszuzählen, damit wir ein realistisches Bild davon bekommen, wie gut Sie in dieser Aufgabe sind.

Measurement of choice of working time: On-screen instructions for the grid task in which subjects choose their working time

Sie nehmen nun noch einmal an einer ähnlichen Zählaufgabe wie der vorherigen teil. Genau wie zuvor zählen Sie die Anzahl Nullen ("0") in jeder Tabelle und erhalten dafür 0,1 EUR, also 10 Euro-Cent pro Tabelle.

Allerdings dürfen Sie nun frei wählen, wie lange Sie an der Aufgabe arbeiten möchten. Sie beginnen auf dem nächsten Bildschirm mit der Aufgabe und bearbeiten diese, solange Sie möchten.

Es wird sich eine Tabelle nach der anderen öffnen, bis Sie sich entscheiden die Arbeit zu beenden.

Außerdem gibt es folgende Besonderheit: Ihre Arbeitszeit heute bestimmt auch Ihre Arbeitszeit im Experiment nächste Woche. Nächste Woche bearbeiten Sie wieder eine ähnliche Aufgabe, bei der Sie in einer vorgegebenen Zeit so viele Tabellen wie möglich auszählen und dafür bezahlt werden. Dabei wird sich eine Tabelle nach der anderen öffnen, bis die Arbeitszeit abgelaufen ist.

Auf Grundlage Ihrer Arbeitszeit heute werden wir zwei gleich große Gruppen mit verschiedenen Arbeitszeiten bilden. Diejenigen, die heute kürzer arbeiten, werden auch nächste Woche kürzer arbeiten (eine Hälfte der Teilnehmer). Diejenigen, die heute länger arbeiten, werden auch nächste Woche länger arbeiten (die andere Hälfte der Teilnehmer). Wer kürzer arbeitet, wird im Durchschnitt weniger

Tabellen korrekt auszählen und weniger verdienen. Das Experiment wird aber auch kürzer dauern (es endet früher). Wer länger arbeitet, wird im Durchschnitt mehr Tabellen korrekt auszählen und mehr verdienen. Das Experiment wird aber auch länger dauern (es endet später).

Sollten Sie dazu Fragen haben, heben Sie die Hand! Einer der Experimentatoren kommt dann zu Ihnen an den Platz.

E.2 Main session

Allgemeine Anweisungen

(distributed on paper at the beginning of the main session)

Willkommen zum heutigen Experiment! Vielen Dank für Ihre Teilnahme!

Während des Experiments werden Sie und die anderen Teilnehmer gebeten, Entscheidungen zu treffen. Sowohl Ihre eigenen Entscheidungen als auch die der anderen Teilnehmer bestimmen Ihre Auszahlung entsprechend den im Folgenden erklärten Regeln.

Das Experiment wird am Computer durchgeführt. Sie treffen Ihre Entscheidungen am Bildschirm. Alle Ihre Entscheidungen und Antworten bleiben vertraulich und anonym.

Das Experiment besteht aus fünf Teilen: TEIL 1, TEIL 2, TEIL 3, TEIL 4, und TEIL 5. Zusätzlich werden Sie einen kurzen Fragebogen beantworten.

Einer der fünf Teile wird vom Computer zufällig zur Auszahlung ausgewählt. Jeder Teil hat dabei die gleiche Wahrscheinlichkeit ausgewählt zu werden. Es ist also in Ihrem Interesse, Ihre Entscheidungen in jedem Teil so zu treffen als wäre dies der einzige Teil.

Außerdem erhalten Sie pauschal 4 EUR für die heutige Teilnahme. Das heißt, Ihr gesamter Verdienst aus der heutigen Sitzung ist die Summe aus Ihren Auszahlungen aus dem zufällig ausgelosten Teil und der pauschalen Entlohnung von 4 EUR. Ihr Verdienst wird Ihnen am Ende der heutigen Sitzung zusammen mit dem Verdienst aus der letzten Woche ausbezahlt.

Alle weiteren Erklärungen bekommen Sie schrittweise jeweils zu Beginn eines Experimentteils. Sie erhalten nacheinander die Anleitungen zu den jeweiligen Teilen. Wir werden Ihnen genug Zeit geben, diese in Ruhe zu lesen und Fragen zu stellen. Bitte zögern Sie nicht, Fragen zu stellen, falls es Unklarheiten gibt.

Bitte beachten Sie, dass wie letzte Woche während des gesamten Experiments ein Kommunikationsverbot herrscht. Sollten Sie Fragen haben, stellen Sie diese

bitte keinesfalls laut, sondern heben Sie die Hand. Einer der Experimentatoren wird dann zu Ihnen kommen, um Ihre Frage zu beantworten. Die Nichtbeachtung dieser Regeln führt zum Ausschluss aus dem Experiment und von allen weiteren Zahlungen.

Allgemeine Informationen zum heutigen Experiment

Im heutigen Experiment wird es wieder darum gehen, möglichst viele Zählaufgaben in der vorgegebenen Zeit korrekt zu lösen, d.h., in möglichst vielen Tabellen die Anzahl der Nullen („0“) korrekt zu zählen. Dabei gibt es allerdings eine Besonderheit. **Jeder Teilnehmer hat 3 Merkmale, die ihn charakterisieren und während des gesamten Experiments gleich bleiben: seine Produktivität, seine Arbeitszeit (in Minuten) und seinen Multiplikator.**

- Die **Produktivität** gibt an, wie viele Zählaufgaben pro Minute ein Teilnehmer letzte Woche in der fünfminütigen Aufgabe korrekt gelöst hat. Für die Hälfte der Teilnehmer liegt die Produktivität **in der unteren Hälfte**. Für die andere Hälfte der Teilnehmer liegt die Produktivität **in der oberen Hälfte**.

- Die **Arbeitszeit** heute hängt von der selbst gewählten Arbeitszeit in der Aufgabe am Ende der Sitzung letzte Woche ab. Die Hälfte der Teilnehmer wird heute in jeder Aufgabe **4,5 Minuten Zeit** haben, so viele Tabellen wie möglich korrekt auszuzählen. Dies sind diejenigen, deren Arbeitszeit letzte Woche in der unteren Hälfte lag. Die andere Hälfte der Teilnehmer wird heute in jeder Aufgabe **7,5 Minuten Zeit** haben, so viele Tabellen wie möglich korrekt auszuzählen. Dies sind diejenigen, deren Arbeitszeit letzte Woche in der oberen Hälfte lag.

- Der **Multiplikator** ist ein Wert, mit dem die Anzahl korrekt gelöster Zählaufgaben multipliziert wird, um die Gesamtleistung zu bestimmen. Er wird jedem Teilnehmer zufällig zugewiesen. Für die Hälfte der Teilnehmer beträgt er **0,75**. Für die andere Hälfte der Teilnehmer beträgt er **1,25**.

Wie Ihre Produktivität, Arbeitszeit und Ihr Multiplikator lauten, erfahren Sie gleich auf dem nächsten Bildschirm.

Für jeden Teilnehmer gilt:

Gesamtleistung = Korrekt gelöste Zählaufgaben in Ihrer Arbeitszeit × Multiplikator

TEIL 1 – Stücklohn (distributed on paper at the beginning of stage 1)

Ihre Aufgabe in TEIL 1 ist ähnlich wie die Aufgabe in der ersten Sitzung. Die Aufgabe besteht wieder darin, möglichst viele Zählaufgaben in der vorgegebenen Zeit korrekt zu lösen, d.h., in möglichst vielen Tabellen die Anzahl der Nullen („0“) korrekt zu zählen. Wie viel Zeit Sie haben, sehen Sie noch einmal auf dem Bildschirm. Jede Tabelle besteht dabei aus 10 Zeilen und 10 Spalten, die entweder eine Null („0“) oder eine Eins („1“) enthalten. Jede Tabelle ist anders als die vorherige. Sie dürfen das bereitliegende Schmierpapier benutzen, wenn Sie möchten. Nachdem Sie eine Antwort eingegeben haben, klicken Sie bitte auf „Bestätigen“. Sie erfahren dann gleich auf dem Bildschirm, ob die Antwort richtig ist oder nicht.

Wenn TEIL 1 der zur Auszahlung ausgewählte Teil ist, sieht Ihre Auszahlung wie folgt aus:

$$\text{Auszahlung} = \underbrace{\text{Korrekt gelöste Zählaufgaben in Ihrer Arbeitszeit} \times \text{Multiplikator}}_{\text{Gesamtleistung}} \times 0,50 \text{ EUR}$$

Haben Sie also zum Beispiel **zehn** richtige Antworten gegeben und Ihr Multiplikator ist **1,25**, erhalten Sie folgende Auszahlung:

$$\text{Auszahlung} = 10 \times 1,25 \times 0,50 \text{ EUR} = 6,25 \text{ EUR}$$

Haben Sie zum Beispiel **zehn** richtige Antworten gegeben und Ihr Multiplikator ist **0,75**, erhalten Sie folgende Auszahlung:

$$\text{Auszahlung} = 10 \times 0,75 \times 0,50 \text{ EUR} = 3,75 \text{ EUR}$$

Ihre Auszahlung wird nicht reduziert, wenn Sie eine falsche Antwort eingeben. Diese Auszahlung nennen wir ab jetzt die **Stücklohn-Auszahlung**.

Sobald alle Fragen zu TEIL 1 beantwortet sind, beginnt die Arbeitszeit für TEIL 1.

TEIL 2 – Wettbewerb (distributed on paper at the beginning of stage 2)

Wie in TEIL 1 haben Sie eine bestimmte Anzahl Minuten Zeit, um möglichst viele Zählaufgaben korrekt zu lösen. Ihre Arbeitszeit wird Ihnen nun noch einmal auf dem Bildschirm angezeigt. Anders als zuvor hängt Ihre Auszahlung für diesen Teil von Ihrer Leistung relativ zur Leistung der anderen Teilnehmer in Ihrer Gruppe ab.

Einteilung in Gruppen:

Für die folgenden Teile des Experiments werden Sie einer **Gruppe mit sechs Mitgliedern** zugeteilt. Die Gruppen wurden zufällig gebildet und bleiben während des gesamten Experiments gleich. Das heißt, dass Sie während des gesamten

restlichen Experiments mit den gleichen Teilnehmern eine Gruppe bilden.

Zur Erinnerung: **Jeder Teilnehmer hat 3 Merkmale, die ihn charakterisieren: seine Produktivität, seine Arbeitszeit (in Minuten) und seinen Multiplikator.**

Bei der Auslosung der Gruppen wurde darauf geachtet, dass jede Gruppe von sechs **Mitgliedern** hinsichtlich der Produktivität, der Arbeitszeit, und des Multiplikators der Gruppenmitglieder wie folgt aussieht:

- Die **Produktivität** von **drei** Gruppenmitgliedern liegt im Vergleich zu allen Teilnehmern in der **oberen Hälfte**. Die Produktivität der **anderen drei** Gruppenmitglieder liegt im Vergleich zu allen Teilnehmern in der **unteren Hälfte**.
- Die gewählte **Arbeitszeit** von **drei** Gruppenmitgliedern lag letzte Woche im Vergleich zu allen Teilnehmern in der **oberen Hälfte**. Diese drei Gruppenmitglieder arbeiten deshalb heute an jeder Zählaufgabe **7,5 Minuten**. Die **Arbeitszeit** der **anderen drei** Gruppenmitglieder lag letzte Woche im Vergleich zu allen Teilnehmern in der **unteren Hälfte**. Diese drei Gruppenmitglieder arbeiten deshalb heute an jeder Zählaufgabe **4,5 Minuten**.
- Der zufällig ausgeloste **Multiplikator** von **drei** Gruppenmitgliedern beträgt **0,75**. Die Anzahl der korrekt ausgezählten Tabellen dieser drei Gruppenmitglieder wird bei der Berechnung der Gesamtleistung also mit **0,75** multipliziert. Der zufällig ausgeloste **Multiplikator** der **anderen drei** Gruppenmitglieder beträgt **1,25**. Die Anzahl der korrekt ausgezählten Tabellen dieser drei Gruppenmitglieder wird bei der Berechnung der Gesamtleistung also mit **1,25** multipliziert.

Regeln des Wettbewerbs:

Wenn TEIL 2 der zur Auszahlung ausgewählte Teil ist, hängt Ihr Gewinn davon ab, wie hoch Ihre Gesamtleistung im Vergleich zu den anderen fünf Mitgliedern in Ihrer Gruppe ist.

Grundsätzlich sind die **zwei** Gruppenmitglieder mit der **höchsten** Gesamtleistung (also Anzahl richtiger Antworten in der gesamten individuellen Arbeitszeit \times Multiplikator) die zwei Sieger des Wettbewerbs.

(The content of the following part in gray differs across treatments. There is no further content for the control treatment (CTR).)

Affirmative action w.r.t. productivity (AAP):

Allerdings gibt es folgende **Sonderregel**:

Mindestens ein Sieger muss ein Gruppenmitglied sein, dessen Produktivität im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte liegt. Produktivität bedeutet dabei, wie viele Zählaufgaben pro Minute jemand letzte Woche korrekt gelöst hat.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung, dessen Produktivität im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte liegt, den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.**

Affirmative action w.r.t. working time (AAW):

Allerdings gibt es folgende **Sonderregel**:

Mindestens ein Sieger muss ein Gruppenmitglied sein, dessen Arbeitszeit letzte Woche im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte lag und das daher heute 4,5 Minuten an der Aufgabe arbeitet.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung unter den drei Gruppenmitgliedern mit der kürzeren Arbeitszeit von 4,5 Minuten den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.**

Affirmative action w.r.t. discrimination (AAD):

Allerdings gibt es folgende **Sonderregel**:

Mindestens ein Sieger muss ein Gruppenmitglied mit dem niedrigen Multiplikator 0,75 sein.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung unter den drei Gruppenmitgliedern mit dem niedrigen Multiplikator von 0,75 den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.**

Die Auszahlung der zwei Sieger sieht wie folgt aus:

$$\text{Auszahlung} = \underbrace{\text{Korrekt gelöste Zählaufgaben in ihrer Arbeitszeit} \times \text{Multiplikator}}_{\text{Gesamtleistung}} \times 1,50 \text{ EUR}$$

Hat ein Sieger also zum Beispiel zehn richtige Antworten gegeben und sein Multiplikator ist **1,25**, so erhält er folgende Auszahlung:

$$\text{Auszahlung} = 10 \times 1,25 \times 1,50 \text{ EUR} = 18,75 \text{ EUR}$$

Hat ein Sieger zum Beispiel zehn richtige Antworten gegeben und sein Multiplikator ist **0,75**, so erhält er folgende Auszahlung:

$$\text{Auszahlung} = 10 \times 0,75 \times 1,50 \text{ EUR} = 11,25 \text{ EUR}$$

Die **anderen vier** Mitglieder in der Gruppe bekommen **keine Auszahlung**.

Falls es einen Gleichstand zwischen zwei Gruppenmitgliedern gibt, wird der Sieger zufällig bestimmt. Die Auszahlung aus diesem Teil nennen wir ab jetzt die **Wettbewerbs-Auszahlung**. Über das Ergebnis des Wettbewerbs werden Sie am Ende der heutigen Sitzung informiert.

TEIL 3 - Wahl zwischen Stücklohn-Auszahlung und Wettbewerbs-Auszahlung

(distributed on paper at the beginning of stage 3)

Wie in TEIL 1 und TEIL 2 haben Sie wieder eine bestimmte Anzahl Minuten Zeit, um möglichst viele Zählaufgaben korrekt zu lösen. Ihre Arbeitszeit wird Ihnen nun noch einmal auf dem Bildschirm angezeigt.

Allerdings wählen Sie nun selbst, welches Auszahlungsschema Sie für Ihre Leistung in TEIL 3 bevorzugen. Sie können entweder die Stücklohn-Auszahlung (Regeln wie in TEIL 1) oder die Wettbewerbs-Auszahlung (Regeln wie in TEIL 2) wählen. Wenn TEIL 3 der zur Auszahlung ausgewählte Teil ist, dann wird Ihr Gewinn folgendermaßen bestimmt:

- Wählen Sie die **Stücklohn-Auszahlung**, ist Ihre Auszahlung:

$$\text{Auszahlung} = \underbrace{\text{Korrekt gelöste Zählaufgaben in Ihrer Arbeitszeit} \times \text{Multiplikator}}_{\text{Gesamtleistung}} \times 0,50 \text{ EUR}$$

- Wählen Sie die **Wettbewerbs-Auszahlung**, hängt Ihr Gewinn davon ab, wie hoch Ihre Gesamtleistung in TEIL 3 im Vergleich zur Gesamtleistung der anderen fünf Mitglieder Ihrer Gruppe ist, die diese in TEIL 2 (Wettbewerb) erbracht haben. Zur Erinnerung: Das ist der Teil, den Sie gerade abgeschlossen haben.

(The content of the following part in gray differs across treatments.)

Control treatment (CTR):

Haben Sie eine höhere Gesamtleistung, also (Anzahl korrekt gelöster Zählaufgaben in der individuellen Arbeitszeit) \times (Multiplikator), als mindestens vier der anderen Mitglieder Ihrer Gruppe in TEIL 2, dann sieht Ihre Auszahlung wie folgt aus:

Affirmative action w.r.t. productivity (AAP):

Grundsätzlich sind die zwei Gruppenmitglieder mit der höchsten Gesamtleistung, also (Anzahl richtiger Antworten in der gesamten individuellen Arbeitszeit) \times (Multiplikator), die zwei Sieger des Wettbewerbs.

Dabei gilt weiterhin die **Sonderregel**:

Mindestens ein Sieger muss ein Gruppenmitglied sein, dessen Produktivität im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte liegt. Produktivität bedeutet dabei, wie viele Zählaufgaben pro Minute jemand letzte Woche korrekt gelöst hat.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung unter den drei Gruppenmitgliedern, deren Produktivität im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte liegt, den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.**

Gewinnen Sie also mit Ihrer TEIL 3-Gesamtleistung im Vergleich zur TEIL 2- Gesamtleistung Ihrer Gruppenmitglieder, dann sieht Ihre Auszahlung wie folgt aus:

Affirmative action w.r.t. working time (AAW):

Grundsätzlich sind die zwei Gruppenmitglieder mit der höchsten Gesamtleistung, also (Anzahl richtiger Antworten in der gesamten individuellen Arbeitszeit) \times (Multiplikator), die zwei Sieger des Wettbewerbs.

Dabei gilt weiterhin die **Sonderregel:**

Mindestens ein Sieger muss ein Gruppenmitglied sein, dessen Arbeitszeit letzte Woche im Vergleich zu allen anderen Teilnehmern am Experiment in der unteren Hälfte lag und das daher heute 4,5 Minuten an der Aufgabe arbeitet.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung unter den drei Gruppenmitgliedern mit der kürzeren Arbeitszeit von **4,5 Minuten** den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.**

Gewinnen Sie also mit Ihrer TEIL 3-Gesamtleistung im Vergleich zur TEIL 2- Gesamtleistung Ihrer Gruppenmitglieder, dann sieht Ihre Auszahlung wie folgt aus:

Affirmative action w.r.t. discrimination (AAD):

Dabei gilt weiterhin die **Sonderregel:**

Mindestens ein Sieger muss ein Gruppenmitglied mit dem niedrigen Multiplikator 0,75 sein.

Ist dies aufgrund der Gesamtleistungen der Gruppenmitglieder nicht automatisch der Fall, so **ersetzt** das Gruppenmitglied mit der höchsten Gesamtleistung unter den drei Gruppenmitgliedern mit dem niedrigen Multiplikator **0,75** den ursprünglichen zweiten Sieger. **In diesem Fall ist das Gruppenmitglied**

mit der zweithöchsten Gesamtleistung von allen sechs Mitgliedern in Ihrer Gruppe also kein Sieger mehr.

Gewinnen Sie also mit Ihrer TEIL 3-Gesamtleistung im Vergleich zur TEIL 2- Gesamtleistung Ihrer Gruppenmitglieder, dann sieht Ihre Auszahlung wie folgt aus:

$$\text{Auszahlung} = \underbrace{\text{Korrekt gelöste Zählaufgaben in Ihrer Arbeitszeit} \times \text{Multiplikator}}_{\text{Gesamtleistung}} \times 1,50 \text{ EUR}$$

Das heißt, sie ist dreimal so hoch wie die Stücklohn-Auszahlung.

Gewinnen Sie mit Ihrer TEIL 3-Gesamtleistung im Vergleich zur TEIL 2- Gesamtleistung Ihrer Gruppenmitglieder **nicht**, erhalten Sie **keine Auszahlung**.

Falls es einen Gleichstand zwischen zwei Personen gibt, wird der Sieger zufällig bestimmt.

Die Gruppenzusammensetzung ist die gleiche wie in TEIL 2. Wählen Sie die Wettbewerbs-Auszahlung, werden Sie am Ende des Experiments über das Ergebnis des Wettbewerbs informiert.

Auf dem nächsten Bildschirm entscheiden Sie sich, ob Sie für Ihre Leistung in TEIL 3 nach der Stücklohn-Auszahlung oder der Wettbewerbs-Auszahlung entlohnt werden wollen. Danach beginnt die Aufgabe.

TEIL 4 (displayed on screen at the beginning of stage 4)

Im Folgenden werden Sie eine neue Aufgabe bearbeiten, bei der es darum geht, Schieberegler in eine bestimmte Position zu bringen.

Auf jedem Bildschirm sehen Sie sechs Schieberegler. Diese können auf einer Skala von 0 bis 100 eingestellt werden. Sobald Sie auf einen der Schieberegler klicken, erscheint daneben die aktuelle Position. Sie können diese durch Ziehen mit der Maus verändern.

Ihre Aufgabe ist es, alle sechs Schieberegler auf einem Bildschirm auf "50" zu stellen. Nur dann ist ein Bildschirm korrekt bearbeitet und Sie können durch Klicken auf Weiter zum nächsten Bildschirm gelangen. Sie haben fünf Minuten Zeit, um so viele Bildschirme wie möglich korrekt zu bearbeiten. Bei dieser Aufgabe arbeiten alle Teilnehmer gleich lange und es gibt auch keinen Multiplikator. Ihre Auszahlung in diesem Teil hängt davon ab, wie viele Bildschirme Sie und die fünf anderen Mitglieder Ihrer Gruppe korrekt bearbeiten. Die Gruppenzusammensetzung ist nach wie vor dieselbe.

Konkret wird Ihr Gewinn folgendermaßen bestimmt: Sie erhalten 10 Euro-Cent für jeden korrekt bearbeiteten Bildschirm jedes Mitglieds Ihrer Gruppe. Die an-

deren Mitglieder Ihrer Gruppe erhalten ebenfalls jeweils 10 Euro-Cent für jeden Bildschirm, den Sie korrekt bearbeitet haben. Das heißt, jeder korrekt bearbeitete Bildschirm jedes Spielers ist für die Gruppe (also alle 6 Mitglieder zusammen) 60 Euro-Cent wert.

Zusätzlich erhalten die Gruppenmitglieder, die in TEIL 2 den Wettbewerb gewonnen haben, eine Ausstattung von 5 EUR. Die anderen Gruppenmitglieder erhalten eine Ausstattung von 2 EUR.

Wenn TEIL 4 der zur Auszahlung ausgewählte Teil ist, dann ist Ihre Auszahlung die Summe aus Ihrer persönlichen Ausstattung und Ihrem Gewinn aus der Summe aller korrekt bearbeiteten Bildschirme der Mitglieder Ihrer Gruppe.

Am Ende des Experiments werden Sie über die Leistung Ihrer Gruppe informiert. Sollten Sie Fragen haben, strecken Sie bitte die Hand aus der Kabine.

TEIL 5 (displayed on screen at the beginning of stage 5)

In diesem Teil treffen Sie fünf Entscheidungen, die Sie selbst und jeweils eines der anderen fünf Gruppenmitglieder betreffen. Um die Entscheidungen zuordnen zu können, wird jedem Gruppenmitglied zufällig eine Nummer von 1 bis 6 zugelost. Sie sind Gruppenmitglied Nummer X.¹⁹

Für jede Entscheidung erhalten jeweils eine Anfangsausstattung von 5 EUR. Ihre Aufgabe ist es zu entscheiden, wie Ihre Anfangsausstattung zwischen Ihnen und dem anderen Gruppenmitglied aufgeteilt werden soll. Sie können also einen Betrag zwischen 0 und 5 EUR (in 10 Cent-Schritten) wählen, den Sie dem anderen Gruppenmitglied geben möchten. Den Restbetrag behalten Sie selbst. Sie erfahren zu keinem Zeitpunkt, wer das andere Gruppenmitglied ist und auch es erfährt auch nicht, wer Sie sind. Das Einzige, was Sie über das jeweils andere Gruppenmitglied vor Ihrer Entscheidung mitgeteilt bekommen, ist, ob es den Wettbewerb in TEIL 2 gewonnen hat oder nicht.

(The content of the following part in gray differs across treatments. There is no further content for the control treatment (CTR).)

Affirmative action w.r.t. productivity (AAP):

... und ob seine Produktivität in der oberen oder in der unteren Hälfte liegt.

¹⁹ The exact number differs for each subject.

Affirmative action w.r.t. working time (AAW):

... und wie lang seine Arbeitszeit war.

Affirmative action w.r.t. discrimination (AAD):

... und wie sein Multiplikator lautet.

Falls dieser Teil zur Auszahlung ausgewählt wird, wird Ihre Auszahlung wie folgt ermittelt: In jeder Gruppe werden zufällig drei Paare ausgelost, deren Entscheidungen für die Auszahlung relevant sind. Das heißt jedes Gruppenmitglied ist genau einem auszahlungsrelevanten Paar zugeordnet. In jedem Paar wird ausgelost, wer der Geber und wer der Empfänger ist. Die Entscheidung des Gebers ist dann für die Auszahlung beider relevant. Das bedeutet, dass jede Ihrer Entscheidungen mit der gleichen Wahrscheinlichkeit ausgezahlt wird und Sie daher jede Entscheidung so treffen sollten, als wäre es die einzige.

Am Ende des Experiments werden Sie darüber informiert, welche Nummer das Gruppenmitglied hat, mit dem Sie ein Paar bilden, wer in diesem Paar der Geber und wer der Empfänger ist, und wie die Entscheidung des Gebers lautet. Sie erfahren dagegen nichts über die Entscheidungen in den anderen Paaren (zu denen Sie nicht gehören). Sollten Sie Fragen haben, strecken Sie bitte die Hand aus der Kabine.

F Questionnaire in German

The questionnaire in the preparatory session contains the following items:

1. **Risk preference, general risk question:** same wording as in German Socio-Economic Panel questionnaire (SOEP, see, for example, Wagner et al. (2007))

Wie schätzen Sie sich persönlich ein: Sind Sie im Allgemeinen ein risikobereiter Mensch oder versuchen Sie, Risiken zu vermeiden? Das äußerst linke Kästchen bedeutet "gar nicht risikobereit", und das äußerst rechte "sehr risikobereit". Mit den Kästchen dazwischen können Sie Ihre Einschätzung abstimmen.

gar nicht risikobereit *sehr risikobereit*

2. **Risk preference, incentivized choice list:** Subjects make eleven, pairwise decisions between a lottery with a fifty-fifty chance of winning either 2 EUR or 7 EUR and a safe payment. The safe payment increases in 0.5 EUR increments, ranging

from 2 EUR to 7 EUR.

3. **Social preference** (survey question, Falk et al., 2018)

Question 1: Stellen Sie sich die folgende Situation vor: Heute haben Sie unerwartet 1000 EUR erhalten. Wie viel von dem Geld würden Sie für einen guten Zweck spenden? (Werte zwischen 0 und 1000 sind möglich).

Question 2: Bitte denken Sie darüber nach, was Sie in der folgenden Situation tun würden. Sie sind in einer Gegend, in der Sie sich nicht auskennen, und merken, dass Sie sich verlaufen haben. Sie fragen einen Fremden oder eine Fremde nach dem Weg. Der/die Fremde bietet Ihnen an, Sie zum Ziel zu bringen. Ihnen zu helfen kostet den/die Fremde/n insgesamt ungefähr 20 EUR. Jedoch sagt der/die Fremde, dass er bzw. sie kein Geld von Ihnen möchte. Sie haben sechs Geschenke bei sich. Das billigste Geschenk kostet 5 EUR, das teuerste Geschenk kostet 30 EUR. Würden Sie eines der Geschenke dem/der Fremden als Dankeschön geben? Welches Geschenk würden Sie dem/der Fremden geben?

1. Nein, ich würde kein Geschenk geben.
2. Das Geschenk im Wert von 5 EUR
3. Das Geschenk im Wert von 10 EUR
4. Das Geschenk im Wert von 15 EUR
5. Das Geschenk im Wert von 20 EUR
6. Das Geschenk im Wert von 25 EUR
7. Das Geschenk im Wert von 30 EUR

4. **Big Five:** we use the 15-item Big Five scale developed for the SOEP (Schupp and Gerlitz, 2008) to measure personality traits.

5. **Locus of control:** we use 10 different items adapted from Rotter (1966) which have been used in the 2005 wave of the SOEP.

6. **Questions on general fairness ideals:** all using the same scale

trifft überhaupt nicht zu *trifft voll zu*

In welchem Maß stimmen Sie persönlich den einzelnen Aussagen zu?

Es ist unfair, wenn jemand, der eine anstrengende Tätigkeit ausübt, wenig verdient.

Wer mehr leistet, sollte auch mehr verdienen.

Wenn jemand etwas von Natur aus gut kann, ist es richtig, ihn dafür zu belohnen.

Es ist falsch, jemanden zu bevorzugen, nur weil er anderswo vielleicht Diskriminierung erfahren hat.

7. **Cognitive ability:** Raven matrices from the Wechsler IQ test (J. Raven and C. Raven, 2008).

Before subjects start the test, we elicit their belief about individual rank as follows:

Bevor Sie anfangen, möchten wir Sie noch bitten, einzuschätzen, wie Sie beim Intelligenztest im Vergleich zu den anderen Teilnehmern am Experiment abschneiden werden. Dabei bedeutet zum Beispiel "0-10%", dass Sie zu den 0-10% Teilnehmern mit den wenigsten richtigen Antworten gehören, während mindestens 90% der Teilnehmer mehr richtige Antworten gegeben haben als Sie. "41-50%" bedeutet, dass mindestens 40% der Teilnehmer weniger richtige Antworten und mindestens 50% mehr richtige Antworten als Sie gegeben haben. "91-100%" bedeutet, dass mindestens 90% weniger richtige Antworten gegeben haben als Sie. Was glauben Sie? Wie schneiden Sie im Vergleich zum Rest der Gruppe ab?

- 0-10%
- 11-20%
- 21-30%
- 31-40%
- 41-50%
- 51-60%
- 61-70%
- 71-80%
- 81-90%
- 91-100%

8. **Cognitive reflection test:** see Frederick (2005).
9. **Socio-demographics:** age, gender, final grade point average at academic high school, last math grade at academic high school, field of study, monthly disposable amount of money, political orientation, number of experiments already participated in the same lab.