CGS WORKING PAPER

Justification as a Key Determinant of the Success of Affirmative Action

Lea Petters (University of Cologne)

Marina Schröder (University of Cologne)

University of Cologne

Justification as a Key Determinant of the Success of Affirmative Action

Lea Petters[*]   Marina Schröder[†]

This version: 19.02.2018

Abstract

Quotas are a common but controversial instrument to reduce discrimination. We experimentally analyze the impact of awareness of exisiting discrimination on the effectiveness of quotas. Between treatments, we vary whether or not a quota is implemented and whether or not the affirmed group is discriminated against. We find that quotas lead to a decrease in performance of affirmed individuals. Absent of discrimination, quotas increase sabotage and reduce help received by affirmed individuals. We do not observe these undesirable effects in the presence of discrimination. Thus, perceived justification has a crucial impact on the effect of quota interventions.

Key Words: Affirmative action, Quota, Sabotage, Real effort, Peer evaluation, Fairness, Discrimination

JEL Classification: C92, J33, J71, M51

[*] University of Cologne, Seminar of Personnel Economics and Human Resource Management, Albertus-Magnus-Platz, 50923 Cologne, Germany, phone: +49 221 470 6312, Email: lea.petters@uni-koeln.de.
[†] University of Cologne, Seminar of Personnel Economics and Human Resource Management, Albertus-Magnus-Platz, 50923 Cologne, Germany, phone: +49 221 470 6939, Email: marina.schroeder@uni-koeln.de.

1. Introduction

More and more countries pass laws enforcing female quotas in order to correct for the existing gender imbalance in higher management positions and executive boards.[1] The implementation of such quotas and other means of affirmative action (AA) are highly controversial. We argue that an important reason for the difference in attidude towards AA relates to beliefs regarding the mechanisms behind the underrepresentation of females in leading positions. Our goal in this paper is to study the relevance of awareness of existing discrimination for the effectiveness of quota interventions.

Previous research reveals several mechanisms that may be driving the underrepresentation of women in leading positions. Examples include: discrimination (Babcock et al. 2017, Beaman, Keleher, and Magruder 2017, for an overview see Betrand and Duflo 2016), gender differences in career-related preferences (Niederle 2017, for an overview see Croson and Gneezy 2009 and), and gender differences in educational attainments (Goldin, Katz, and Kuziemko 2006). Analyzing data from a representative survey among U.S. citizens (Pew Research Center 2014), we find that around one third of the respondents oppose further measures taken to foster gender equality, while the rest are in favor of such measures. Interestingly, whether or not an individual is in favor of AA correlates with the perception of whether or not differences in the representation of women in leadership positions are perceived to be due to discrimination - circumstances that women cannot be held accountable for (see Appendix 1). Holzer and Neumark (2000), Son Hing, Bobocel and Zanna (2002), Shteynberg et al. (2011) and Bohnet (2016) provide further evidence for the relationship between the support of AA and awareness of existing discrimination. There is evidence for heterogeneity in the awareness of discrimination and some indications for a tendency to underestimate existing discrimination (Billy 2015, Sipe et al. 2016, LeanIn.Org and McKinsey & Company 2017). Thus, the effectiveness of quotas (and other means of AA) is likely to be related to awareness of discrimination. In this paper we provide experimental evidence for this relationship.

While we motivate our work in the gender context, our experimental design is gender-neutral and thus, our experimental findings are applicable to quota regulations in various contexts. We introduce a novel experimental design that allows us to quantify the effect of quotas on behavior in multiple dimensions. In a two-by-two factorial design, we vary whether or not a quota is implemented and whether or not the affirmed group is indeed discriminated against. We assign both discrimination and quotas according to a randomly determined player type, green or yellow. In the experiment, groups of four participants, each consisting of two green and two yellow types, compete for two prizes. Participants perform a real effort

---

[1] Countries that have passed such laws for gender quotas include Norway in 2006, Spain in 2007, Italy in 2011, Belgium in 2011, France in 2011, and Germany in 2016. Furthermore, several countries have adapted affirmative action measures for specific minorities in various contexts, such as for example preferential treatment of African American or Hispanic applicants in employment and education in the U.S. or quotas for the lowest castes and specific tribes in higher education, political positions, and government jobs in India (e.g., Sowell 2004, Bagde, Epple, and Taylor 2016).

task, in which they have to illustrate predefined objects using a given set of materials. Prizes in the tournament are awarded according to peer reviews of these illustrations and - depending on the treatment - a quota. The use of peer reviews to determine tournament winners mimics many real life work situations. It is estimated that up to 90% of companies listed in the Fortune 500 use peer reviews as a tool for performance evaluation (Edwards and Ewen 1996, Bracken 2001). While peer reviews can serve as a good proxy for performance, previous experimental evidence shows that they are often a source of sabotage (Carpenter, Matthews, and Schirm 2010, Chowdhury and Gürtler 2015, Balietti, Goldstone, and Helbing 2016, Leibbrandt, Wang, and Foo 2017).

In treatments with a quota, at least one of the two prizes has to be awarded to a participant of the affirmed type. In settings involving discrimination, we induce discrimination by assigning less working time to discriminated participants.[2] In addition to peer reviews, we elicit a neutral measure of performance through an online experiment in which independent participants, who are blind to the treatments, evaluate the illustrations. We use the mean evaluations from the online experiment as a neutral performance measure. Similar to previous studies (Carpenter, Matthews, and Schirm 2010, Balietti, Goldstone, and Helbing 2016, Leibbrandt, Wang, and Foo 2017), we define sabotage as the difference between this neutral measure and peer reviews. As a third dimension of behavior, we conduct a modified dictator experiment after the main experiment. This allows us to observe whether our treatment interventions have an effect on helping (giving) behavior in a non-competitive environment.

Previous experimental research reveals that quotas effectively increase the proportion of women selecting into higher management career paths. This increase is explained by the provision of role models (Beaman et al. 2009, Bhavnani 2009, Beaman et al. 2012, Banerjee et al. 2013) and an increase in the willingness of women to compete (Gneezy, Niederle, and Rustichini 2003, Balafoutas and Sutter 2012, Niederle, Segal, and Vesterlund 2013). To this point, only little is known about the negative side effects of quota interventions. Leibbrandt, Wang, and Foo (2017) show that quotas increase sabotage directed at the affirmed group and that this increase in sabotage undermines the positive effect of quotas on women's willingness to compete. This previous research provides important insights into how quotas can increase the fraction of women in leading positions in the long run. However, it may take years until those women who select into higher management career paths today climb the ladder high enough to qualify for actual top management positions. Therefore, it is also relevant to study the immediate effects of quotas on behavior in predefined competitive environments. In this paper, we focus on such immediate effects with an emphasize on potential negative effects of quotas.

Previous research provides some indication that the effectiveness of AA policies depends on whether the intervention is perceived as justified (Balafoutas, Davis, and Sutter 2016, Ip, Leibbrandt, and Vecci work in progress). Balafoutas, Davis, and Sutter (2016) compare randomly assigned quotas to quotas

---

[2] In this paper, we define discrimination as an inequality in opportunity that individuals cannot be held accountable for and that does not arise due to affirmative action. Given that we randomly assign player types, discrimination in our design is independent of any individual characteristics.

assigned according to gender and find that randomly assigned quotas are less efficient and less accepted compared to gender quotas. Parallel to our work, Ip, Leibbrandt, and Vecci (work in progress) study gender quotas in gift exchange experiments. They vary the perception of the justification of the quota by providing different information about the average performance of women in the relevant task. Ip, Leibbrandt, and Vecci (work in progress) find that this information has a significant effect on both principal and agent behavior. To the best of our knowledge, we are the first to systematically vary whether the affirmed group is discriminated against or not and to cleanly study the causal impact of (perceived) discrimination on the effectiveness of a quota intervention.[3]

Research in behavioral economics reveals that many individuals have a disutility from inequality in opportunity (see e.g. Konow 1996, 2000, Cappelen et al. 2007). Absent of discrimination, a quota creates an inequality in opportunity since affirmed types are more likely to win a prize, even though they cannot be held accountable for this advantage. Therefore, quotas that are introduced absent of discrimination are likely to be considered as unjustified. To reduce the disutility arising due to an unjustified quota, we predict that participants in our experiment will increase sabotage directed at affirmed types. If the affirmed group is discriminated against, the introduction of a quota – if designed appropriately - reduces inequalities in opportunity (that arise due to discrimination). In a setting involving discrimination, we predict that the quota will be perceived as justified and will not result in an increase in sabotage directed at affirmed types.

We find that the awareness of existing discrimination justifying a quota is crucial to reduce the negative effects that quotas entail. As predicted, we observe that a quota without discrimination results in an increase in sabotage directed at affirmed types. We do not observe such an effect when a quota is implemented in the presence of discrimination. Furthermore, we observe that a quota in a setting without discrimination leads to a decrease in help (giving in the dictator game) received by affirmed types, which is not the case for a quota in a setting involving discrimination. Thus, we find that an unjustified quota entails substantial negative effects for those who are supposed to benefit from this intervention (affirmed types). These negative effects do not arise if the quota is justified by discrimination.

In our experiment, quotas do not lead to a significant increase in the fraction of affirmed types among the prizewinners of the tournament. The reason seems to be that sabotage is effectively used to reduce inequalities arising due to discrimination or unjustified quotas. However, quotas lead to a decrease in the performance of affirmed types. Thus, while quotas do not increase the fraction of affirmed types among the prizewinners in our experiment, we find that quotas lead to a significant decrease in the performance of those affirmed types who are successful and win a prize. Resulting differences in

---

[3] Fallucchi and Quercia (2016) provide an innovative experimental design in which the affirmed group is discriminated against. However, they do not provide a comparison to a situation that does not involve discrimination. Therefore, the results of Fallucchi and Quercia (2016) are not informative on the impact of discrimination on the effectiveness of quota interventions.

performance standards may lead to negative stereotypes (Reuben, Sapienza, and Zingales 2014, Glover, Pallais, and Pariente 2017) and actually reinforce discrimination against affirmed individuals.

Our findings have implications for policy makers and organizations seeking to increase diversity by implementing quota regulations. We provide insighs for undesirable effects of quotas on performance, sabotage and helping behavior. These undesirable effects are especially relevant if the quota is not justified by discrimination. Thus, our findings suggest that before a quota is implemented, it is crucial to inform and convince all individuals affected by this intervention of its justification. Our research provides an important contribution to the research on the effectiveness of quota interventions. To the best of our knowledge, we are the first to show that the immediate effects of a quota crucially depend on whether or not this intervention is justified by (perceived) discrimination.

2. Experimental Design

In the experiment, groups of four participants compete for two prizes each. The experiment consists of two stages: a working stage and a peer review stage. In the working stage, participants perform a real effort task. The real effort task consists of illustrating predefined objects using a given set of materials. It is a modified version of the creative task introduced by Laske and Schröder (2016). Participants receive a set of materials: one string, two O-rings, four wooden sticks, and twelve colored glass pebbles (see Figure 1). In the trial round, participants are instructed to illustrate a butterfly, and in the three payoff-relevant rounds they are asked to illustrate glasses, a flower, and a car. Participants take pictures of their illustrations using a special software and a preinstalled webcam. See Figure 2 for examples of pictures of illustrations and Appendix 2 for the experimental instructions. In each of the four rounds, the time available to work on the task is restricted and depends on the treatment and the type of participant. Participants can take as many pictures as they want within the time frame. Once the time is up, participants choose one of these pictures in each round to be payoff relevant. All other pictures are deleted and not payoff relevant.

Figure 1: Set of materials

In the peer review stage of each round, we show participants the pictures of the illustrations created by all members of their group (including their own illustration) and ask them to rate the illustrations of their peers (participants did not rate their own illustration) on a scale from 0.0 to 10.0 (0.0 being the worst rating and 10.0 being the best rating). Thus, in the peer review stage, the illustrations of all participants are evaluated by the three other members of each group. Prizes are awarded according to the final score from this peer review and depending on the treatment, as described in the following.
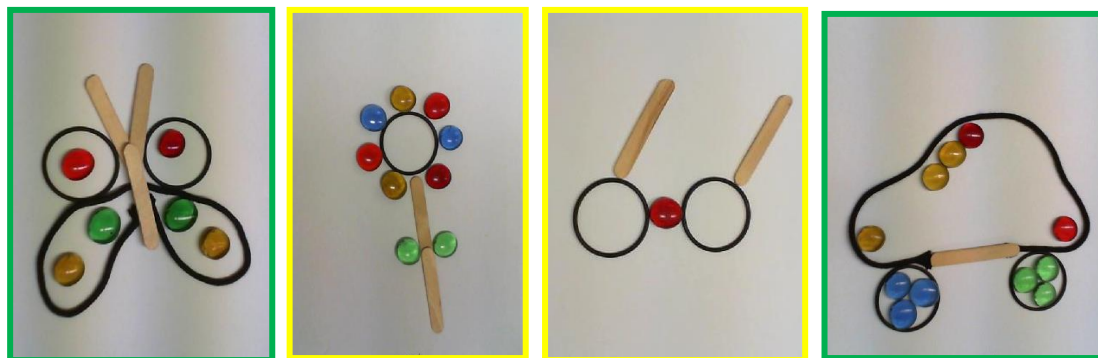

Figure 2: Examples of illustrations

In a two-by-two design, we vary whether or not a quota is implemented and distinguish between an equal setting in which the affirmed group is not discriminated against and a discrimination setting in which the affirmed group is discriminated against. Table 1 provides an overview of the treatments. In all treatments, we randomly assign participants to a type (yellow or green), and each group consists of two yellow and two green types. All participants are informed about their type before the experiment starts. In the peer review stage, the pictures are shown in a randomized sequence. They have a colored frame, which indicates the type (green or yellow) of the respective group member (see frames in example pictures provided in Figure 2). Thus, the types (but not the identities of the other participants) are common information.

In the treatments without a quota, the participants with the highest and the second highest final rating in the peer review receive a prize, independent of their type. In the treatments with a quota, at least one of the two prizes is awarded to a participant of the affirmed type. Thus, in the treatments with a quota, the participant among the affirmed types with the highest final rating from the peer review receives a prize (even if this participant is not among the two participants with the highest final ratings) and the participant among the remaining three participants of either type with the highest final rating from the peer review receives a prize. This procedure is common information. In the equal setting, all types have five minutes time to illustrate the object. In the discrimination setting, one type has a reduced working time of only 2 minutes and 30 seconds. Again, this procedure is common information. In the treatment

6

with a quota and discrimination, the same types are discriminated against and affirmed. We randomized whether green or yellow types are affirmed and/or discriminated against.[4]

Table 1: Treatments

	Equal	Discrimination
No Quota	**Base_equal** No quota Equal working time for all n=160; N=20	**Base_discr** No quota Discriminated type has reduced working time n=160; N=20
Quota	**Quota_equal** At least one prize awarded to affirmed type Equal working time for all n=160; N=20	**Quota_discr** At least one prize awarded to affirmed & discriminated type Affirmed & discriminated type has reduced working time n=152; N=19[5]

n indicates the total number of participants in each treatment.
N indicates the number of independent observations in each treatment.

The main experiment is repeated three times with no feedback between rounds and with random rematching within groups of eight participants between rounds. Only one round is randomly chosen for payment. The two winners of the tournament in this randomly determined round receive a prize of 16 euros each. Additionally, all participants receive a show-up fee of 4 euros. Before the experiment, we conduct a trial round, which is not payoff relevant and consists only of the working stage. Upon arrival, each participant is randomly seated and assigned to one of two types (yellow or green). We conducted the experiments at the Cologne Laboratory for Economic Research (CLER). Overall, 632 subjects participated in our experiment and we ran 40 experimental sessions. We used Java and SoSci Survey (Leiner 2014) to program our experiments and recruited the participants through the online recruitment software ORSEE (Greiner 2015).

After the main experiment, we conduct a modified dictator game (see Appendix 3 for the instructions). This modified dictator game serves as a proxy to measure the effect of quotas on helping behavior in a non-competitive environment. The instructions for this experiment are given after the main experiment.

---

[4] The quota is implemented according to a randomly assigned player type and is independent of gender or any other participant characteristics. In this way, our experiment differs from most of the previous experimental studies that focus specifically on the effect of gender quotas (see e.g. Leibbrandt, Wang, and Foo 2017, Niederle, Segal, and Vesterlund 2013). The advantages of randomly assigning the quota are that this procedure allows us to cleanly induce discrimination and that our findings can also be applied to other types of quotas, i.e. quotas based on race or family background.

[5] One group had to be eliminated from the analysis because one subject from this group did not pass the control questions.

In this dictator game, dictators allocate 2 euros between themselves and an anonymous recipient. Dictators are informed about the type of the recipient (yellow or green) from the previous tournament experiment. We repeat the dictator game four times with random rematching, so that each participant acts twice as a dictator and twice as a recipient. In both roles (dictator and recipient), each participant is matched to both a yellow and a green type player. We randomized the sequence of the four rounds. Only one of the rounds is randomly chosen for payment, and no feedback is given between the rounds. After the dictator game, we distributed a paper and pen questionnaire (see Appendix 4 for the questionnaire).

Our treatment interventions can have an effect on the way participants evaluate their peers, i.e., on sabotage, but it can also have an effect on performance per se. To be able to distinguish between these two effects, we elicit a quality measure that is independent of the treatments. We conduct an online experiment, in which we ask a total of 400 independent raters to evaluate the illustrations from the experiment on the same scale as in the laboratory experiment (0.0 to 10.0). To avoid overload of the raters, each rater evaluates a subset of pictures. Each screen shown to the raters displays four pictures from one group in one round. Thus, the composition of pictures on one screen shown to the independent raters was identical to that shown to the participants of the experiment. The pictures also have the yellow and green frames as in the laboratory setting. One important difference to the main experiment is that the evaluators are blind to the treatment and do not know what the frames mean. Each set of pictures is evaluated by ten different raters. The average of these ten evaluations constitutes our neutral quality measure. Participants in the online experiment receive a fixed payment of 2 euros and can earn an additional bonus of up to 4 euros, which is awarded according to the quadratic deviation from the mean evaluation for one randomly chosen picture (we follow the procedure suggested by Selten 1998). Online raters were also recruited through the online recruitment software ORSEE (Greiner 2015) and did not participate in any previous related experiments.

3. Results

Figure 3 provides summary statistics on performance (height of bars), sabotage (black color) and peer-reviews (white color). The upper graph displays results from the equal setting that does not involve discrimination and the lower graphs from the discrimination setting. In each graph, a comparison of the left two bars displays the effect of a quota on affirmed types, while a comparison of the right two bars displays the effect of a quota on non-affirmed types. Focusing first on performance, we find that in the equal setting, the introduction of a quota leads to a significant decrease in the performance of the affirmed types by 5 percent (U-test, p=0.10),[6] while we find no effect on the performance of

---

[6] For the non-parametric analysis in this paper, we use mean values for the matching groups of eight participants and over all rounds of the experiment. Thus, we make sure that we only compare independent observations. We always report p-values for two-sided tests.

Equal
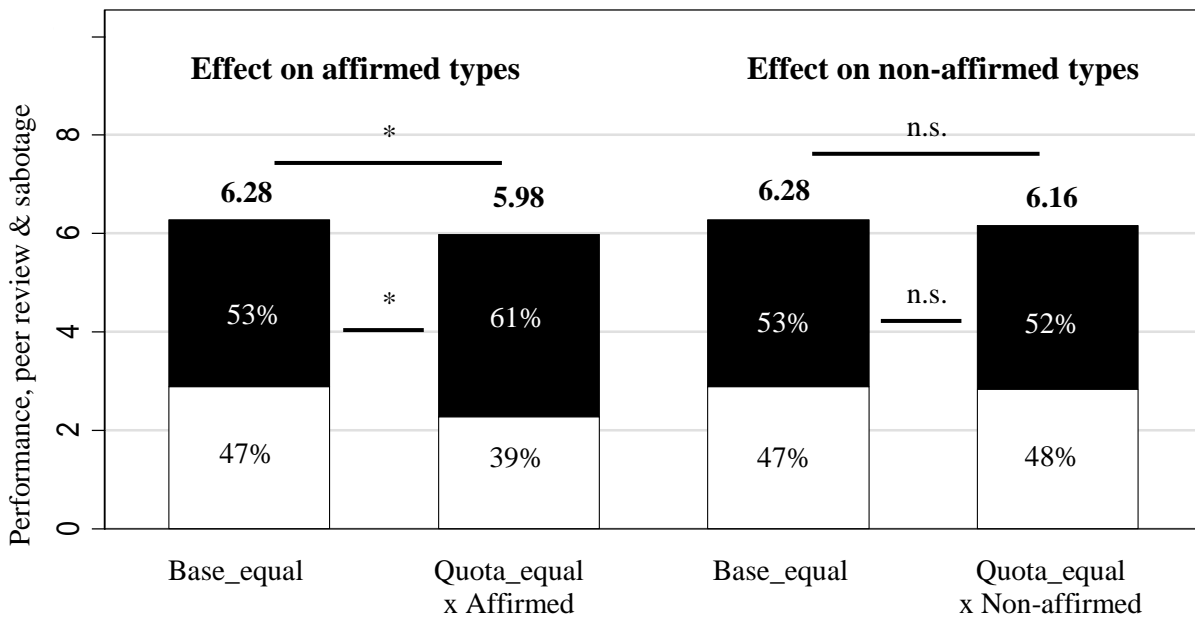
Discrimination
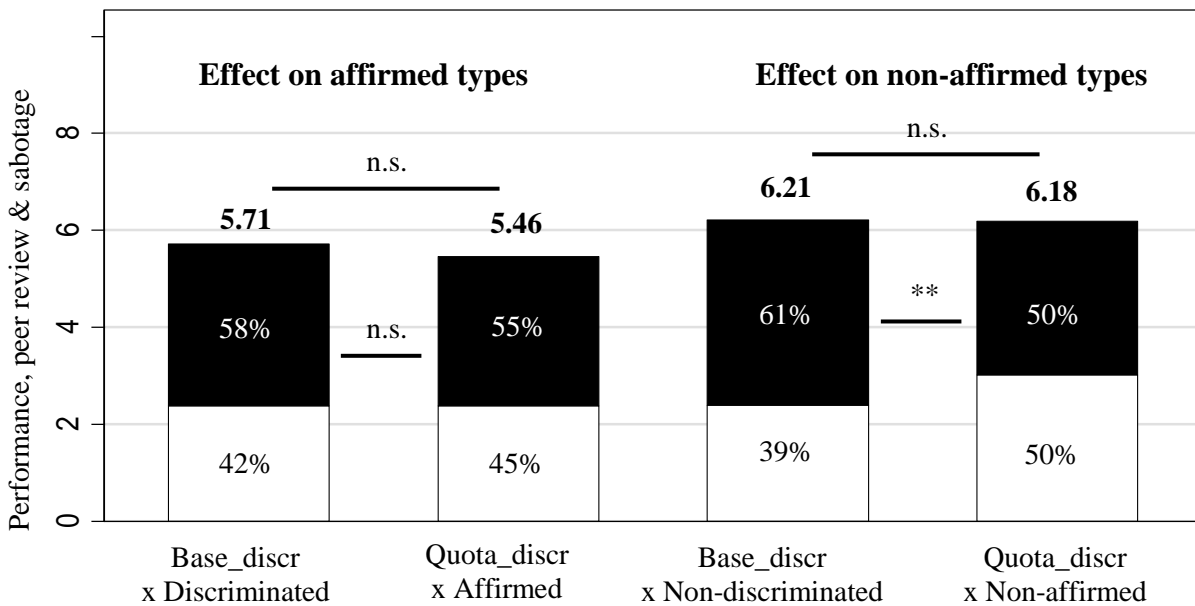
Height of bars indicates mean performance   ☐ Peer review   ■ Sabotage

The height of the bars refers to the mean performance. We indicate mean sabotage and peer review as percentages of performance. A comparison of the left two bars relates to the effect of a quota on affirmed types, while a comparison of the right two bars refers to the effect of a quota on non-affirmed types. In the Quota_equal treatment, the baseline group for the affirmed and non-affirmed types is the same. In the Quota_discr treatment, the affirmed types are discriminated against and the non-affirmed types are non-discriminated. Significance levels of two-tailed U-tests indicated by stars * $p<0.1$, **$p<0.05$. For the non-parametric analysis in this paper, we use mean values for the matching groups of eight participants over all rounds of the experiment.

Figure 3: Performance, peer review & sabotage

non-affirmed types (U-test, p=0.30). In the discrimination setting, a quota leads to a decrease in the performance of affirmed types by 4 percent. This decrease is, however, not significant using the non-parametric analysis (U-test, p=0.21). Again, we observe no effect on the performance of non-affirmed types (U-test, p=1.00). In all treatments, we find that the average performance (height of bars) is significantly greater than the average peer reviews, which are indicated by the white shading of the bars (Wilcoxon rank-sum test, p<0.01, for all treatments and types). The difference between actual performance and peer reviews amounts to 50 to 61 percent of performance. In the following, we will refer to this difference as sabotage. We find that introducing a quota in the equal setting (as displayed in the upper graph) leads to a significant increase in sabotage directed at affirmed types (U-test, p=0.07), while it does not have a significant effect on sabotage directed at non-affirmed types (U-test, p=0.73). In the discrimination setting, we find that introducing a quota does not have a significant effect on sabotage directed at affirmed types (U-test, p=0.45), while it leads to a significant decrease in sabotage directed at non-affirmed types (U-test, p=0.02).

Given the structure of our data, we perform the non-parametric analysis in a very conservative manner and only compare mean values of matching groups of eight participants over all rounds. As an alternative approach to analyzing the data, Table 2 provides the results from two random-effects regression models. In column (1), we present the effect of quotas and discrimination on performance (measured by evaluations from independent online raters) introducing random effects for the creator of an illustration. Column (2) of table 2 provides the results from a multilevel mixed-effects linear regression model with sabotage received in percent of actual performance as the dependent variable. In this model, we include random effects for the victim (creator of an illustration) and the saboteur (peer reviewer of an illustration). In both specifications, we cluster standard errors by matching groups of eight participants. We decompose the effect of introducing a quota into its effect on the affirmed types and the non-affirmed types in the equal setting and the discrimination setting (independent variables are labled as Quota_equal x Affirmed, Quota_equal x Non-affirmed, Quota_discr x Affirmed, Quota_discr x Non-affirmed, correspondingly). Furthermore, we control for the effect of discrimination on discriminated (Discr x Discriminated) and non-discriminated (Discr x Non-discriminated) types.

With respect to performance (Col (1) of Table 2), we find that in the equal setting, the introduction of a quota leads to a significant decrease in performance of the affirmed types (coefficients of the independent variable Quota_equal x Affirmed). In the discrimination setting, we find a similar but insignificant negative effect on the performance of affirmed types (coefficients of the independent variable Quota_discr x Affirmed). Furthermore, we find no significant effect of a quota on the performance of non-affirmed types, independent of whether the quota is introduced in the equal or the discrimination setting. Confirming that our discrimination intervention is relevant, we find that discrimination has a significant negative effect on the performance of the discriminated types

(independent variable Discr x Discriminated), while it has no significant effect on the performance of non-discriminated types.

Table 2: Regression analysis performance & sabotage received

	(1) Performance	(2) Sabotage (%)
Quota_equal x Affirmed	-0.303*	8.108**
	(0.160)	(3.868)
Quota_equal x Non-affirmed	-0.125	-0.956
	(0.149)	(4.246)
Quota_discr x Affirmed	-0.257	-3.718
	(0.218)	(5.428)
Quota_discr x Non-affirmed	-0.029	-11.035**
	(0.210)	(5.129)
Discr x Discriminated	-0.566***	5.569
	(0.175)	(4.909)
Discr x Non-discriminated	-0.072	7.776
	(0.153)	(5.154)
Constant	6.281***	53.177***
	(0.083)	(2.975)
Observations	1,896	5,673
Number of participants	632	632
Number of groups	79	79

(1): Random-effects model, allowing for creator random effects. The dependent variable is performance (derived from the evaluations of independent raters who are blind to treatments).

(2): Two-way error component linear model, allowing for victim and saboteur random effects. The dependent variable is sabotage received as percentage of actual performance.

Independent variables: Quota_equal x Affirmed (dummy equal to one for affirmed types in the Quota_equal treatment), Quota_equal x Non_affirmed (dummy equal to one for non-affirmed types in the Quota_equal treatment), Quota_discr x Affirmed (dummy equal to one for affirmed types in the Quota_discr treatment), Quota_discr x Non_affirmed (dummy equal to one for non-affirmed types in the Quota_discr treatment), Discr x Discriminated (dummy equal to one for discriminated types in all treatments involving discrimination), Discr x Non-discriminated (dummy equal to one for non-discriminated types in all treatments involving discrimination).

The reference group is the Base_equal treatment.

In discrimination treatments involving a quota (Quota_discr) the affirmed type is discriminated and the non-affirmed type ist non-discriminated.

In both specifications robust standard errors are clustered by matching groups of eight participants. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Result 1: *Abesent of discrimination, quotas lead to a significant decrease in performance of the affirmed types. In the discrimination setting, the performance of affirmed types decreases insignificantly.*

As a next step, we focus on the effect of quotas on sabotage activity (Col (2) in Table 2), specifically on the difference between the peer-review and the independent rating measured in percent of actual performance. As in the non-parametric analysis above, we find that in the equal setting, a quota leads to

a significant increase in sabotage directed at affirmed types (coefficient for Quota_equal x Affirmed). In the discrimination setting, however, we find no significant effect of a quota on sabotage directed at affirmed types. With respect to non-affirmed types, we find that in the equal setting, a quota does not have a significant effect on sabotage targeted at these types. In the discrimination setting, a quota leads to a significant decrease in sabotage directed at non-affirmed types. We find that discrimination leads to an insignificant increase in sabotage directed at both discriminated and non-discriminated types.

Overall, we find that a quota, that is introduced in a setting of equal opportunity, leads to an increase in sabotage directed at affirmed types while a quota that is introduced in a setting involving discrimination does not entail this undesirable effect. These findings are in line with the predictions of a fairness model in which individuals care about fairness in opportunities (Konow, 2000, Cappelen et al. 2007). Due to the introduction of the quota absent of discrimination, the affirmed types have a higher probability of winning a prize. To reduce the arising inequality, subjects in our experiment increase sabotage directed at affirmed types. We do not observe this effect when a quota is implemented in a setting involving discrimination. However, in the discrimination setting, we observe that a quota leads to a significant decrease in sabotage targeted at non-affirmed/non-discriminated types.[7]

Result 2: *In the equal setting, quotas lead to a significant increase in sabotage directed at affirmed types. In the discrimination setting, quotas do not have an effect on sabotage directed at affirmed types.*

Next, we study the effect of quotas on the fraction of affirmed and/or discriminated types among the prizewinners of the tournament. Table 3 summarizes the fraction of affirmed and/or discriminated types who win a prize according to the peer review mechanisms and the fraction of affirmed and/or discriminated types who would have won a prize if winners would have been determined according to the independent rating. As there is neither affirmative action nor discrimination, we do not have data for the fraction of affirmed and/or discriminated types who successfully win a prize in the Base_equal treatment. Our baseline in this setting is an equal split in which both types win the prize with equal probability. In the Quota_equal treatment, 54 percent of the prizes are awarded to affirmed types. This fraction is not significantly different from 50 percent (Chi^2 test, p=0.41). However, the fraction of affirmed types winning a prize according to the independent rating would have been significantly larger than the equal split (Chi^2 test, p=0.07). Thus, it seems that sabotage reduces the effectiveness of the quota in the equal setting. In the Base_discr treatment, we find that 48 percent of the winners are from the discriminated type. This fraction is not significantly different from 50 percent (Chi^2 test, p=0.72)

---

[7] Discrimination leads to an insignificant increase in sabotage activity targeted at both discriminated and non-discriminated indiviudals. When we introduce a quota in a setting involving discrimination, we overcompensate for this inequality. Per design of our quota, the affirmed and discriminated types in the Quota_discr treatment have higher chances of winnig a prize compared to non-discriminated and non-affirmed types and compared to the Base_equal treatment. Introducing a quota in a setting involving discrimination thus neutralizes the increase in sabotage directed at non-discriminated types in the discrimination treatment. However, the quota does not reduce sabotage directed at affirmed and discriminated types.

despite the fact that the performance of discriminated types is on average 8 percent lower compared to that of non-discriminated types. If prizes would have been awarded according to independent ratings, this fraction would have been substantially lower but not significantly different from an equal split of prizes (Chi^2 test, p=0.14). Again, it seems that sabotage reduces the effect of discrimination on the assignment of prizes. Last, in the Quota_discr treatment, we find that both the actual fraction of winners from the affirmed type determined through the peer review and the fraction of winners from the affirmed type that would have resulted from the independent ratings are equal to 54 percent and not significantly different from the equal split (Chi^2 test, p=0.34). Comparing the fraction of affirmed types winning a prize in the Base_discr treatment and the Quota_discr treatment, we do not find a significant difference in the fraction of affirmed and/or discriminated types winning a prize (Chi^2 test, p=0.19). Overall, these results suggest that sabotage is effectively used to reduce inequalities in opportunity of winning a prize arising due to quotas or discrimination. Both an unjustified and a justified quota lead to only a slight and insignificant increase in the fraction of affirmed types winning a prize.

Table 3: Fractions of affirmed and/or discriminated types among prizewinners

	Equal		Discrimination	
	Base_equal	Quota_equal	Base_discr	Quota_discr
Peer review	-	54%	48%	54%
Independent rating	-	58%	43%	54%

Percentage of affirmed and/or discriminated types among prizewinners under peer review and independent rating.

Result 3: *In both the equal and the discrimination settings, quotas do not have a significant effect on the fraction of the affirmed types among prizewinners in the tournament.*

While we find no significant effect of quotas on the fraction of affirmed types who actually succeed in the tournament, we do find that the introduction of a quota leads to a decrease in the performance of those affirmed types who win a prize. In the Quota_equal treatment those affirmed types who win a prize have a 6 percent lower performance compared to the Base_equal treatment (U-test, p=0.10). In the Quota_discr treatment, the performance of winning affirmed types is 5 percent lower compared to discriminated types in the Base_discr treatment. However, this difference is not significant (U-test, p=0.19). We find that the performance of winning non-affirmed types remains unaffected by the introduction of a quota in both settings (U-test, p>0.53, pairwise comparisons).

Result 4: *In the equal setting, a quota has a significant negative effect on the performance of affirmed types who win a prize. We find a slight but insignificant negative effect of quotas on the performance of affirmed prizewinners in the discrimination setting.*

After the main experiment, we conduct a dictator experiment to elicit the effect of quotas on helping behavior in a non-competitive environment. We keep types fixed and inform dictators about the recipient's type from the main experiment. Table 5 summarizes help received in this dictator experiment

for the four treatments and for affirmed and/or discriminated and non-affirmed and/or non-discriminated types separately. Our results of this dictator experiment mirror those of sabotage behavior. In the equal setting, we find that a quota results in a significant decrease in help received by affirmed types (U-test, p=0.02). In the discrimination setting, we do not find such an effect of a quota on help received by affirmed types (U-test, p=0.86). However, in the discrimination setting, the introduction of a quota leads to a slight but insignificant increase in help received by the non-affirmed (and non-discriminated) types (U-test, p=0.15).

Table 4: Summary statistics for help received

	Equal		Discrimination	
	Base_equal	Quota_equal	Base_discr	Quota_discr
Affirmed/ Discriminated	0.32 (0.08)	0.26 (0.11)	0.30 (0.12)	0.29 (0.08)
Non-affirmed/ Non-discriminated	0.32 (0.08)	0.35 (0.09)	0.25 (0.08)	0.29 (0.08)

Mean values, standard errors in parentheses.
Overall values reported for the base_equal treatment as there is no distinction between affirmed/discriminated and non-affirmed/non-discriminated types.

Result 5: *In the equal setting, quotas lead to a significant decrease in help received by the affirmed types. In the discrimination setting, quotas do not have a significant effect on help received by affirmed types.*

4. Discussion and Conclusion

We find evidence that the justification of quotas matters. In line with models of liberal egalitarianism (Konow, 1996, 2000, Cappelen et al. 2007), we find that quotas without discrimination lead to a significant increase in sabotage activity directed at affirmed types, and to a decrease in help received by affirmed types. We do not find similar effects for quotas that are justified by discrimination. These undesired effects of unjustified quotas on sabotage and helping are likely to have an impact on the effectiveness of quotas not only through their immediate implications, but also through a reduction of women selecting into relevant career paths. Leibbrandt, Wang, and Foo (2017) show that the anticipation of increased sabotage activity directed at affirmed individuals undermines the desirable effect of quotas on the selection of affirmed individuals into competitive environments. Furthermore, our results imply that women who are in doubt of having benefitted from a quota rule, which is not perceived as justified, might receive less help from colleagues and might find it harder to establish cooperative networks. Similar to the effect of anticipated sabotage, anticipating a reduction in help received may also affect career choices of affirmed types in an undesirable way.

In our experimental setting, quotas do not lead to an increase in the fraction of affirmed types among the prizewinners of the tournament. The reason seems to be that sabotage is effectively used to reduce inequalities arising due to a quota or due to discrimination. Nevertheless, we find that quotas lead to a

decrease in the performance of the winning affirmed types. That is, while we do not observe the desirable increase in the representation of the affirmed types as prizewinners, we do observe that quotas (especially in the equal setting) lead to a situation where affirmed types have to meet lower performance standards in order to be successful. While this itself causes inefficiencies, the negative impact is likely to go beyond these immediate effects. Previous research shows that differences in performance standards lead to negative stereotypes (Reuben, Sapienza, and Zingales 2014, Glover, Pallais, and Pariente 2017). Thus, a quota could reinforce discrimination against affirmed individuals.

Our findings provide an important contribution to the exisiting literature on the effects of quotas as a mean of AA. To the best of our knowledge, we are the first to experimentally show that the effectiveness of a quota crucially depends on the perceived justification of a quota. This finding has broad implications not only for research but also for policy makers and organizations seeking to implement quotas as a tool of AA. If those affected do not perceive the intervention as justified, a quota has negative immediate effects, which are especially harmful for the affirmed group. It seems that a top-down enforcement of a quota - without information about the reasons behind such a quota - may actually do more harm than good. Our results suggest that quotas entail fewer undesirable effects if they are perceived as justified. To increase perceived justification of quotas, it is important to inform those affected of existing discrimination. According to our findings, providing such information can positively affect diversity through two channels. First, in the absence of AA, the perception of discrimination induces distortions of peer reviews which counteract discrimination. That is, if participants are aware of existing discrimination, they adapt their peer reviews in a way as to reduce discrimination even when no means of AA are taken. Second, negative effects of AA can be avoided by convincing those who are affected of the need for these measures.

References

Babcock, L., Recalde, M. P., Vesterlund, L., and L. Weingart. 2017. "Gender differences in accepting and receiving requests for tasks with low promotability". *American Economic Review*, *107*(3), 714–747.

Bagde, S., Epple, D., and L. Taylor. 2016. "Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India". *American Economic Review*, *106*(6), 1495–1521.

Balafoutas, L., Davis, B. J., and M. Sutter. 2016. "Affirmative action or just discrimination? A study on the endogenous emergence of quotas". *Journal of Economic Behavior and Organization*, *127*, 87–98.

Balafoutas, L., and M. Sutter. 2012. "Affirmative Action Policies Promote Women and Do Not Harm Efficiency in the Laboratory". *Science*, *335*(6068), 579–582.

Balietti, S., Goldstone, R. L., and D. Helbing. 2016. "Peer review and competition in the Art Exhibition Game". *Proceedings of the National Academy of Sciences*, *113*(30), 8414–8419.

Banerjee, A. V., Duflo, E., Imbert, C., and R. Pande. 2013. "Entry, Exit and Candidate Selection: Evidence from India". 3ie Draft Grantee Final Report.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and P. Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, *124*(4), 1497–1540.

Beaman, L., Duflo, E., Pande, R., and P. Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India". *Science*, *335*(6068), 582–587.

Beaman, L., Keleher, N., and J. Magruder. 2017. „Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi". *Journal of Labor Economics*, *forthcoming.*

Betrand, M., and E. Duflo. 2016. "Field Experiments on Discrimination". NBER Working Paper No. 22014, National Bureau of Economic Research.

Bhavnani, R. R.. 2009. "Do electoral quotas work after they are withdrawn? Evidence from a natural experiment in India". *American Political Science Review*, *103*(1), 23–35.

Billy, I. A. 2015. "Gendered perceptions of the glass ceiling by generation and educational levels: A quantitative exploratory study of us middle-level managers". PhD diss., Capella University.

Bohnet, I. 2016. *What works: Gender equality by design*. Harvard University Press.

Bracken, D. W. 2001. *The Handbook of Multisource Feedback: The Comprehensive Resource for Designing and Implementing MSF Processes*. San Francisco: Jossey-Bass.

Cappelen, A. W., Hole, A. D., Sorensen, E., and B. Tungodden. 2007." The pluralism of fairness ideals: An experimental approach". *American Economic Review*, *97*(3), 818–827.

Carpenter, J., Matthews, P. H., and J. Schirm. 2010. "Tournaments and Office Politics: Evidence from a Real Effort Experiment". *The American Economic Review*, *100*(1), 504–517.

Chowdhury, S. M. and O. Gürtler. 2015. "Sabotage in contests: a survey". *Public Choice,* 164 (1-2), 135-155.

Croson, R., and U. Gneezy. 2009. "Gender Differences in Preferences". *Journal of Economic Literature*, *47*(2), 448–474.

Edwards, M. R., and A. J. Ewen. 1996. "How to manage performance and pay with 360-degree feedback". *Compensation and Benefits Review*, 28(3), 41-46.

Fallucchi, F., and S. Quercia. 2016. "Affirmative Action and Retaliation in Experimental Contests". CBESS Discussion Paper No. 16–3, Centre for Behavioural and Experimental Social Science, University of East Anglia.

Glover, D., Pallais, A., and W. Pariente. 2017. "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores". *The Quarterly Journal of Economics*, *132*(3), 1219–1260.

Gneezy, U., Niederle, M., and A. Rustichini. 2003. "Performance in competitive environments: gender differences". *Quarterly Journal of Economics*, *118*(3), 1049–1074.

Goldin, C., Katz, L. F., and I. Kuziemko. 2006. „The Homecoming of American College Women: The Reversal of the College Gender Gap". *Journal of Economic Perspectives*, *20*(4), 133–156.

Greiner, B. 2015. "Subject pool recruitment procedures: organizing experiments with ORSEE". *Journal of the Economic Science Association*, *1*, 114–125.

Holzer, H., and D. Neumark. 2000. "Assessing Affirmative Action". *Journal of Economic Literature*, *38*(3), 483–568.

Ip, E., Leibbrandt, A., and J. Vecci. "Affirmative Action in Principal-Agent Relationships: Evidence from Labour Market Experiments". Work in progress.

Konow, J. 1996. "A positive theory of economic fairness". *Journal of Economic Behavior and Organization*, *31*(1), 13–35.

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

Laske, K., and M. Schröder. 2016. "Quantity, Quality, and Originality: The Effects of Incentives on Creativity". CGS Working Paper No. 7–1, Cologne Graduate School in Management, Economics, and Social Sciences, University of Cologne.

Leibbrandt, A., Wang, L. C., and C. Foo. 2017. „Gender Quotas, Competitions, And Peer Review: Experimental Evidence on the Backlash Against Women". *Management Science*, *forthcoming*.

Leiner, D. J. 2014. SoSci Survey (Version 2.5.00-i) [Computer software].

LeanIn.Org, and McKinsey & Company. 2017. "Women in the workplace 2017". Retrieved from: https://womenintheworkplace.com (October 2017)

Niederle, M. 2017. "A Gender Agenda: A Progress Report on Competitiveness". *American Economic Revies: Papers & Proceedings 2017, 107*(5), 115-119

Niederle, M., Segal, C., and L. Vesterlund. 2013. „How Costly Is Diversity? Affirmative Action in Light of Gender Differences in Competitiveness". *Management Science*, *59*(1), 1–16.

Pew Research Center. 2014. "Gender and Leadership Dataset". Retrieved from: http://www.pewsocialtrends.org (September 2017)

Reuben, E., Sapienza, P., and L. Zingales. 2014. „How stereotypes impair women's careers in science". *Proceedings of the National Academy of Sciences*, *111*(12), 4403-4408.

Selten, R. 1998. "Axiomatic characterization of the quadratic scoring rule". *Experimental Economics*, *1*, 43–62.

Shteynberg, G., Leslie, L. M., Knight, A. P., and D. M. Mayer. 2011. "But Affirmative Action hurts Us! Race-related beliefs shape perceptions of White disadvantage and policy unfairness". *Organizational Behavior and Human Decision Processes, 115*(1), 1-12.

Sipe, S. R., Larson, L., Mckay, B. A., and J. Moss. 2016. "Taking off the blinders: A comparative study of university students' changing perceptions of gender discrimination in the workplace from 2006 to 2013". *Academy of Management Learning and Education*, *15*(2), 232–249.

Son Hing, L. S., Bobocel, D. R., and M. P. Zanna. 2002. "Meritocracy and Opposition to Affirmative Action: Making Concessions in the Face of Discrimination". *Journal of Personality and Social Psychology, 83*(3), 493-509.

Sowell, T. 2004. *Affirmative action around the world: an empirical study*. New Haven & London: Yale University Press.

For Online Publication

Appendix 1: PEW Gender and Leadership Study

The Pew Research Center survey on Gender and Leadership (Pew Research Center, 2014) is a representative study among 1,835 (921 females, 914 males) U.S. residents and was conducted in November 2014 by the Pew Research Center. The study collected data on the perceptions women in leadership positions along with further demographical information. Our analysis focused on a set of nine questions on the perception of discrimination in the population (Q4a), the support of further gender equality measures (Q8) and the reasons for female underrepresentation in top executive positions (Q16a-g). We change the original coding for the purpose of our analysis. The complete questions and some summary statistics can be found in the following table.

Table A1: Pew Research Center survey questions and summary statistics

No.	Question	Mean	Std. dev.	N
Q4a	How much discrimination would you say there is against WOMEN in our society today? 3 A lot 2 Some 1 Only a little 0 None at all . Refused	1.58	0.85	1,825
Q8	Which of these two statements comes closer to your own view, even if neither is exactly right? 0 This country has made the changes needed to give men and women equality in the workplace 1 This country needs to continue making changes to give men and women equality in the workplace . Refused	0.71	0.45	1,807

Q16	Very few top executive positions in business in this country are filled by women. Here is a list of some possible reasons why.
	
	For each one, please indicate whether you think it is a major reason, a minor reason, or not a reason why there aren't more women in TOP EXECUTIVE BUSINESS POSITIONS.
	
	    2 Major reason
	    1 Minor reason
	    0 Not a reason
	   . Refused			

- a	Generally speaking, women don't make as good managers as men.	0.27	0.58	1,784
- b	Women's responsibilities to family don't leave time for running a major corporation.	0.84	0.78	1,786
- c	Generally speaking, women aren't tough enough for business.	0.39	0.64	1,788
- d	Women don't have access to the same kinds of personal networks and connections that men have.	0.7	0.8	1,785
- e	Generally speaking, women are less likely than men to ask for promotions and raises.	0.71	0.75	1,785
- f	Women are held to higher standards than men and have to do more to prove themselves.	1.13	0.87	1,787
- g	Many businesses are not ready to hire women for top executive positions.	1.15	0.86	1,784

In Table A2, we show the spearman correlation between the support for the introduction of further gender equality measures (Q8) and the different reasons (Q16a-g), summarized into the three categories *preferences*, *characteristics* and *discrimination*.

Table A2: Correlation table gender equality measures support and female underrepresentation

Spearman correlation between support for further gender equality measures (0/1) and the opinion about the importance of reasons for female underrepresentation in top executive positions in business (0-2)		Correlation coefficient	
Preferences	- Women's family responsibilities don't leave time *(b)*	-0.024	0.043*
	- Women are less likely to ask for promotions *(e)*	0.095***	
Characteristics	- Women lack manager qualities *(a)*	-0.059**	-0.040*
	- Women are not tough enough *(c)*	-0.013	
Discrimination	- Women don't have access to networks *(d)*	0.181***	0.365***
	- Women are held to higher standards *(f)*	0.360***	
	- Businesses are not ready to hire women *(g)*	0.321***	

Appendix 2: Experimental Instructions – Tournament

Welcome to this experiment!

Please carefully read through the following instructions. If you have a question, please raise your hand. We will then come to your desk and answer your question.

All of your decisions are anonymous. Communication with other participants is not permitted for the duration of the experiment. We would like to ask you to switch off your mobile phone and place it in your bag.

You will receive a show-up fee of 4 euros for your participation. You can earn additional money in the following experiment.

Instructions

Experimental Procedure

- This experiment consists of **multiple rounds**.
    - Initially you will take part in a test round that is not relevant for your payment.
    - After that, 3 experimental rounds will be conducted.
- At the beginning of the experiment you will be randomly assigned a type.
    - You are either **type "green"** or **type "yellow."**
    - You can recognize your type based on the color of the frame of your display.
    - This type assignment remains constant for the entire experiment.
- In each round, you will be randomly assigned to a group of four participants.
    - Each group consists of 2 type "green" and 2 type "yellow" participants.
    - You will be assigned to a new group in each round.
    - However, your type (green or yellow) remains constant for the entire experiment.

Your Task

- Your task is to **illustrate an object using given materials**.
    - Group members of both types are provided with the same materials to illustrate the object (see images below)
        - The materials provided to type "yellow" members are pictured in the bottom left (yellow frame).
        - The materials provided to type "green" members are pictured in the bottom right (green frame).
    - The object that you are supposed to illustrate will be displayed on your screen.
    - You will illustrate a different objects in each round.
    - All members of your group have to illustrate the same object in the respective rounds.

Baseline and Quota Treatments:

    - You have **5 minutes** time available in each round.

Discrimination and Discrimination & Quota Treatments:

    - The time available to illustrate the object in each round are limited
        - **Type "yellow"/"green"** group members have 2 minutes and 30 seconds available in each round.

Rating the Illustrations

- Each group member rates the other 3 group members' illustrations
- The color of the frame of the respective illustrations indicates the type (green or yellow) to which the group member who produced the illustration belongs.
- The rating is conducted on a scale from **0.0 to 10.0 points**.
    - 0.0 points correspond to the worst rating.
    - 10.0 points correspond to the best rating.
    - Please always specify **exactly one decimal place** (please use a dot as a decimal sign).
- The sum of the points awarded equals the final rating.
    - For each illustration, this **final rating** lies **between 0.0 and 30.0 points**.

Payment

- At the end of the experiment, **one of the three experimental rounds** will be randomly **chosen**.
    - Only this round is relevant for the payment.
- Based on the final rating and the assigned type, exactly **two prizes** in the amount of **16 euros** each will be paid out to **two** different group members.

Baseline and Discrimination Treatments:

- The group member with the highest final rating **among all** group members of **both types** receives a prize.
- The group member with the highest final rating among the **remaining three** group members of **both types** receives a prize.
- The other two group members do not receive a prize.
- This means that **at most two prizes** are awarded to **type "green"** group members.
- This means that **at most two prizes** are awarded to **type "yellow"** group members.

Discrimination and Discrimination & Quota Treatments:

- The **type "green" ("yellow")** group member with the higher final rating between **the two** type "green" ("yellow") group members receives a prize.
- The group member with the highest final rating among the **remaining three** group members of **both types** receives a prize.
- The other two group members do not receive a prize.
- This means that **at least one prize and at most two prizes** are awarded to **type "green" ("yellow")** group members.
- This means that at **most one prize** are awarded to the **type "yellow" ("green")** group members.

<u>Procedure</u>

Please proceed with the illustration of the object as follows:

1. Illustrate the object in the marked area using the provided materials.
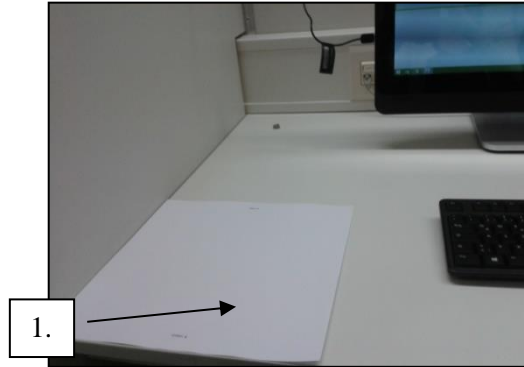
2. Take a photo of the illustrated object by clicking on "take photo." If the photo meets your expectations, save this photo by clicking on "save photo." If a photo does not meet your expectations, you can delete it by clicking on "delete photo."
3. You can take and save as many photos as you wish during the time available. You cannot take or save any additional photos after time has expired.
4. After time has expired, you have the opportunity to choose one of your saved photos. Only this photo will be rated by the other group members in the next step. None of the other photos will be rated.
5. Note that only previously saved photos can be chosen.

Please note the following when illustrating the objects:

- **Only** use the provided materials.
- For each illustration, you may use all materials or a selection of materials.
- Place the object only **within** the area marked with the piece of paper (only this area will be captured by the camera)
- Pay attention to the **direction** of your illustration (the piece of paper is labeled with "top" and "bottom").
- Pay attention that **your hands** are **not visible** in the marked area.
- Keep the **unused materials outside** of the marked area.
- Please **do not write or draw** on the piece of paper representing the marked area.
- Pay attention to the **time limit** for the task; after time has expired you **cannot** take or save any **new photos**.

Appendix 3: Experimental Instructions – Dictator Game

Instructions

Welcome to this part of the experiment! Please carefully read through the following instructions. If you have questions, please raise your hand. We will then come over to you and answer your question. As before, all of your decisions are anonymous. Communication with other participants is not permitted for the duration of the experiment.

The Experiment

This experiment consists of four rounds. In each round you will be **anonymously assigned** to another participant and also **assigned** a role. At the beginning of each round, you will learn the other participant's type (green or yellow) as well as your role (active or passive) in this round. Assignment to the player types (green or yellow) corresponds with the assignment from the first experiment. **The roles will be newly assigned in each round**.

- The **active** participant has an **endowment of 2.00 euros** in each round.
- The **passive** participant has an **endowment of 0.00 euro** in each round.
- The **active** participant **chooses** how many euros (in 0.10 euro increments) he/she would like to **transfer to** the **passive** participant.

A total of **4 rounds** will be conducted in this experiment. The assignment of roles (active and passive) varies between each round, and you will be assigned to a different participant in each round.

At the end of the experiment, **one round** will be **randomly** determined to be relevant for the payment.

Payment

- The active participant's payment equals 2.00 euros minus the amount transferred in the round relevant for the payment, i.e., **the active participant's payment = 2.00 euros – transferred amount**.
- The passive participant's payment equals 0.00 euro plus the amount transferred in the round relevant for the payment, i.e., **the passive participant's payment = 0.00 euro + transferred amount**.

Appendix 4: Questionnaire

Questionnaire

Please, answer the following questions while we prepare the payment.
Thank you!

		female	male
1.	Please, indicate your gender:	☐	☐
2.	How old are you?		
3.	How many siblings do you have?		
4.	Do you have a red-green colorblindness?	Yes ☐	No ☐

Please answer the following questions using the provided scale.

5. How satisfied are you with the experiment overall?

Not satisfied at all				Very satisfied
☐	☐	☐	☐	☐

6. How much would you like to participate in an experiment like this one again?

Not at all				Very much
☐	☐	☐	☐	☐

7. How likely is it that you would recommend participating in an experiment like this one to a friend?

Not likely at all				Very likely
☐	☐	☐	☐	☐

8. How fair did you consider the payment structure of this experiment?

Not fair at all   Very fair
☐   ☐   ☐   ☐   ☐

9. Did you feel disadvantaged or advantaged by the payment structure of this experiment?

Very
disadvantaged   Very advantaged
☐   ☐   ☐   ☐   ☐

10. How justified did you consider this disadvantage or advantage?

Not justified at all   Very justified
☐   ☐   ☐   ☐   ☐

11. How much fun did you have solving the task?

No fun at all   A lot of fun
☐   ☐   ☐   ☐   ☐

12. How creative are you?

Not creative at all   Very creative
☐   ☐   ☐   ☐   ☐

13. How difficult did you find the task?

Not difficult at all   Very difficult
☐   ☐   ☐   ☐   ☐

14. How well do you think you performed compared to the other participants in this room?

Far below average				Far above average
☐	☐	☐	☐	☐

15. In how many rounds do you think you received the highest or second highest final rating?

0	1	2	3
☐	☐	☐	☐

16. In how many of these three rounds do you think you received a prize?

0	1	2	3
☐	☐	☐	☐

17. How likelyis it, do you think, that the other group members rated your illustration according to its quality?

Not likely at all				Very likely
☐	☐	☐	☐	☐

18. How do see yourself: Are you someone who is willing to take risks or do you try to avoid them?

Not willing at all to take risks				Very willing to take risks
☐	☐	☐	☐	☐

Please indicate for the following questions how much the respective statements apply to you.

1. I get upset when someone is better off for no reason.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐

2. I feel guilty when I am better for no reason.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐

3. If someone harms me on purpose, I will try to take revenge even when I have to bear the costs.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐

4. If someone does me a favor, I am happy to return it.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐

5. I like to compete with others.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐

6. It is important to me to be the best.

Doesn't apply at all				Fully applies
☐	☐	☐	☐	☐