Competing for Cruelty

How adjusting incentives can increase destruction and conflict

Florian Kammermeier and Malte Rudolph (University of Passau)

We examine if members of a group that are in a position to exert violence on another group can be stimulated to start a competition for more violent behaviour. Our experiment simulates conflict with a modified Tullock-Contest, where the winner group can destroy part of the loser's endowment. In the baseline a small bonus is randomly allocated within the winner group. In the treatment the group member who proposes the highest destruction gets the bonus. This incentive increases destruction by 18.5 %, conflict by 22,6 % and prevents conflict from settling.

Inhalt

Introduction: Literature and Experiments Review

Destruction seems to be part of human nature. Regularly, there are news about people destroying other people's property or harm each other. This behaviour appears to worsen in groups. A real-life example is bullying children on the playground who often incite each other (Juvonen and Galvan, 2008). A more severe example is the destructions and crimes by the German army during World War II: "Troops of the Wehrmacht were directly involved in the genocide of the Jews and widespread crimes against enemy soldiers and the civilian population, acting both on orders by their superiors and also, in some instances, on their own initiative (p. 101)" (Waller, 2007). This is the starting point of our experiment as there are two contradicting explanations about World War II which are supposed to explain the murder and torture by the German army. The first one is the *intentionalistic* explanation which sees the violent behaviour as a result of top-down incentives by the commanders. They ordered or rewarded destruction and torture by the soldiers. The *functionalistic* explanation states that the destruction evolved from within the soldiers who had too much freedom in their decision making. So, it was not incentivised but part of a group dynamic which lead to a self-radicalization (Mommsen, 1976). This inspired us to look at the question if members of a group that are in a position to exert violence on another group can be stimulated to start a competition for more violent behaviour.

To investigate this question, we conduct a lab experiment. It consists of two stages. The first one is a lottery contest between two groups. We modify the lottery contest proposed by Tullock (1980), so intra-group free riding isn't possible in our setting because the median proposal within the group must be invested by every group member. We adapt it to avoid that group members have different endowments in the second stage. That makes it easier to compare their behaviour. Furthermore, in our setting the groups don't win a price but avoid losing part of their endowment. In the second stage the winner team can destroy a certain amount of the loser's team's endowment without having a strategic use of this destruction. The decision is again taken through the median proposal. In our treatment the subject who proposes the highest destruction is given a small bonus and in the baseline the bonus is randomly allocated amongst the winning group. This is supposed to compare if destruction can be top-down incentivised or might also happen as part of a group dynamic. These two stages are played for four rounds.

The experiment can be seen as a mix of a lottery contest and a Joy of Destruction (JoD) game. Although, in the treatment there is a small incentive to destroy which differs to the JoD designed by Abbink and Sadrieh (2009). We are not aware of any study that has examined the mix of lottery contest and JoD the way we do. However, there are studies which examine these games independently. Regarding the Tullock or lottery contest there has already been made substantial research. By 2015 there were 55 experimental papers examining the lottery contest in several forms. Most studies find significant

overbidding relative to the Nash equilibrium prediction (Dechanaux et al., 2015). There are two studies that have partly a similar setup of the lottery contest as our experiment. Lee (2012) conducts a group contest where the effort of the members within a group is determined by the minimum effort among the group members. Chowdhury et al. (2013) examines the opposite where only the highest amount invested by every group member counts for the lottery contest. This is similar to our experiment as it is not the sum of lottery tickets that are invested by the group, but it is one group member who finally determines the amount of lottery tickets. However, in Lee`s (2012) and Chowdhury et al.'s (2013) setup free-riding is still possible because everyone only pays his/her own proposed amount. The second adjustment of our lottery game is that the winner team doesn't win a prize but only doesn't lose part of its endowment. To our knowledge this adjustment hasn't been used in a lottery game yet.

The Joy-of-Destruction research is much newer and was started by Abbink and Sadrieh (2009) with a simple game between two participants. They have the possibility to destroy the other persons endowment and no benefit or cost ensues from their decision. They found that 20.4 % of the participants endowment was destroyed without strategic reason. The authors interpreted these results as "indication of the pleasure of being nasty". Zizzo and Oswald (2001) similarly show that subjects, under the condition that there is an inequality, are willing to pay to reduce other people's endowment. This is related to our experiment as there is also an indirect cost of destruction. Destroying the other persons endowment can be expected to result in a more intense conflict and may lead to retaliation when the other group wins. Zhang and Ortmann (2013) played with participants the JoD and Dictator game. So, the participants had the option to destroy other people's endowment as well as share their own. Results showed that participants both decided to give money in the Dictator game also decided to destroy in the JoD game. Perhaps the most similar experiment to ours is done by Prediger et al. (2014) who applied the JoD game to specific socioeconomic backgrounds. The game was played by farmers in either resource rich or poor areas. Additionally, they had to pay to reduce the other farmers income. Prediger et al. (2014) found that farmers from resource-scarce areas 17% more often destroyed another's reward than in the resource-rich areas. They reason it with the more competitive environment. Although, we didn't conduct our experiment in a competitive environment, we created one by playing the lottery contest before.

Design and Hypothesis

The experiment is anonymously played for 4 rounds between two groups of 3 persons each. In each round every group member has an endowment of 200 Taler. The first stage is similarly designed as a lottery contest. Each group member makes a proposal within the group of up to 100 Taler how much every member must invest in lottery tickets. One ticket costs one Taler. Then, the median proposal of the group members must be invested by each one. Therefore, free riding within the group isn't

possible. After both groups determined the number of lottery tickets which are bought, all tickets are put in a bowl and one ticket is randomly drawn. So, the chance of winning the contest is the amount of own tickets divided by the total number of tickets. Finally, every member of the loser team loses 40 Taler of her/his endowment. This is the second adjustment of the normal lottery contest. Thereby the game is easier to understand for the participants as there is only one kind of endowment from which all payments are subtracted during the rounds.

In the second stage of each round the winner team can destroy up to 60 Taler of the loser team. The winner team has no strategic use of the destruction, because every round starts with 200 Taler. The amount of destruction is again determined by the median proposal of the team members. In the baseline group one team member is randomly drawn to get the bonus of 5 Taler. In the treatment group the team member who makes the highest proposal gets a bonus of 5 Taler. This is supposed to be a small top-down incentive to create an intra-group competition. After the destruction every player pays the remaining Taler in a bank account which is paid out at the end of the experiment. The complete instructions are in the Appendix.

Factors that influence the decision to destroy the other player's endowment without own strategic use, as it is constructed in the baseline, are contradictory. On the one hand there seems to be a joy of destruction or being nasty. Abbink and Sadrieh (2009) found that 20.4 % of the participants endowment was destroyed without strategic reason. Mommsen (1976) also argued that the group dynamics might lead to destruction, although there is no top-down incentive. Furthermore, negative reciprocity might lead to destruction. If the other player has destroyed some of the endowment in the last round, participants might retaliate. Although, they are aware that this might lead to further negative reciprocity. On the contrary, there are factors to not destroy the other players endowment. The literature on dictator games has shown that there is a sense of altruism (Bardsley, 2008). So, participants include the other player's utility function in their own. Applied on destruction, this predicts that participants don't destroy. From a rational game theoretical perspective there should be no destruction in the baseline because it neither increases the individual's payoff nor there is a higher societal payoff. This includes also that there is no efficiency gain from destruction. Third, participants might fear negative reciprocity in the next round if they destroy the other team's endowment. Weighting these different arguments, we state in our first hypothesis that in the baseline group, the members of the winner team propose low destruction (H01).

In the treatment group there is a small incentive of 5 Taler for the member of the winner group who proposes the highest destruction. This is derived from the *intentionalistic* explanation for the violence of the German Army during WW II. The bonus can be seen as a top-down incentive because it is determined in the instructions by the experimenters. There is now a small strategic benefit of

proposing a destruction of 60 Taler. Furthermore, participants can hand over the responsibility of their behaviour as they can argue to just decide in line with the incentives of the experiment. Factors that might lead to a lower level of destruction are the same as in the baseline: altruism and fear of negative reciprocity. Hence, our second hypothesis is that the destruction in the treatment group will be higher than in the baseline group (H02).

If participants have a similar expectation as we have in H01 and H02, they might adjust their decisions in the lottery contest. In the baseline there is a lower fear of high destruction for the losing team and therefore a lower incentive to invest in the lottery contest. In the treatment, the opposite is the case and participants might try to avoid the increased expected cost by investing more in the lottery contest. Hence, our third hypothesis is that in the treatment the participants invest more in the lottery contest than in the baseline (H03).

Procedure of the Experiment

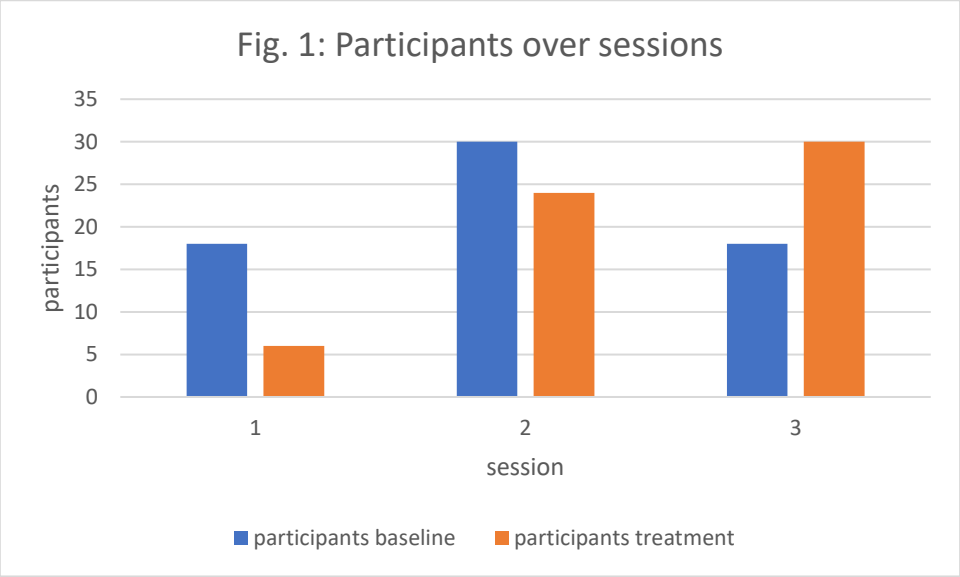
The experiment was conducted with bachelor students in the tutorials of the courses "Makroökonomie" and "Public Finance" at the University of Passau in lecture halls. It was played at the end of the tutorials and took about 20 minutes. The first session was held in "Public Finance" at the 8th of July 2019. 53 bachelor students participated. The second session was held on the 10th of July in "Makroökonomie". 94 bachelor students participated. The third session was held on the 11th of July in "Makroökonomie". 138 bachelor students participated. In total there were 258 participants. We didn't ask for personal details of the student as gender, subject of study and age. The incentivisation for the participants was that over the three sessions 9 participants were randomly drawn and paid out. In the first session 2 participants were drawn, in the second session 4 and in the third session 3. The Taler these participants had at the end of the experiment were paid out in Euros by an exchange rate of 30 Taler = 1 Euro. The experiment was programmed in classEx (Giamattai and Lambsdorff, 2019) and the participants used their own smartphones. The matching was done randomly and anonymously by classEx and the groups and its opponent groups stayed constant for the whole experiment. There were no trial sessions.

The experiment was usually announced at the end of the tutorials and the lecturers asked the students to stay for the experiment. The lecturers also announced that there was money to win as an incentive. On average about 10 % of the students left before the experiment started. Afterwards the lecturers gave over to us (the experimenters) but stayed in the room. Before the experiment started we announced that the required language was German, the participants weren't allowed to talk during the experiment, some participants were randomly drawn to be paid out, the pay-out was probably between 8 to 25 Euros, the pay-out was done double-blind. Furthermore, the experiment would take about 20 minutes and we explained how to log-in at classEx. The remaining instructions to participate

were shown in classEx. During the experiment, one experimenter was responsible for starting the next stages in classEx, while the other experimenter announced how much time the participants had left in the current stage. At the end of the experiment, we asked them to not talk about the experiment to students outside their course and the winners were paid out outside the lecture hall.

Results

Some participants dropped out of the experiment due to problems with the internet connection and some groups were not matched correctly[1]. We took out all groups that were matched incorrectly and cancelled groups with inactive players and their enemy groups. Before the participants started with their decisions they had to go through a questionnaire. Those who did not fill it out were marked in the data and later we flagged groups where one person was marked as inactive. 126 participants remained. 66 of them in the baseline and 60 in the treatment. As figure 1 shows they were not distributed equally across sessions, which causes limitations. We will address these in the chapter "Discussion".[2] In the following, we will test our hypotheses by using the cleaned data sheet.

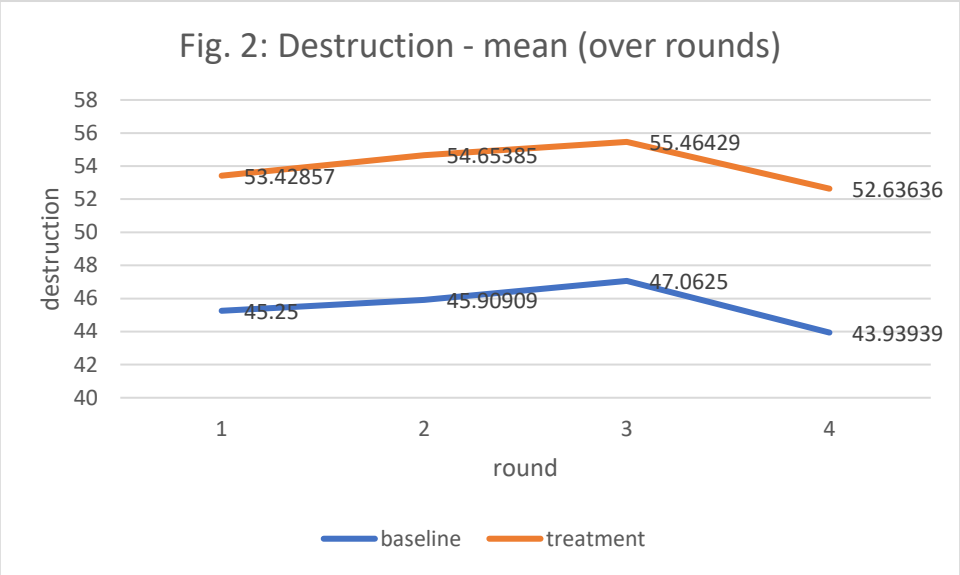


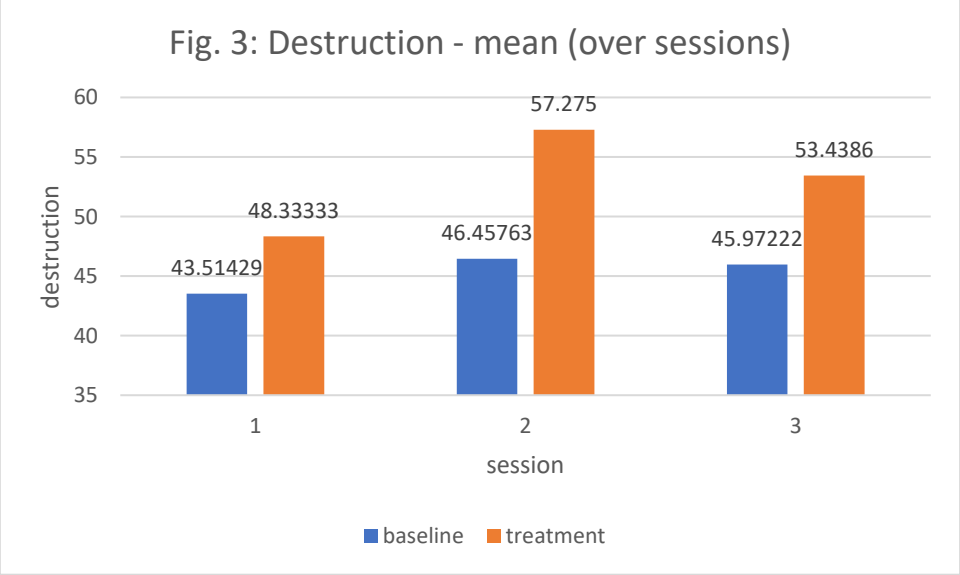
(H01) In the baseline group, the members of the winner team propose low destruction.

[1] Due to a software error some groups were not matched one to one, but a third group was matched to the duo. So e.g. group A reacted to the decisions of group B and vice versa, while group C also reacted to the results of group A, but no group saw the results of group C. Therefore, we had to drop groups that were matched like group C. Also 7 groups had only two participants and had to be dropped together with their enemy groups. We cannot exactly explain which part of the software caused the errors and they never appeared in pre-tests.

[2] Our data still contains three participants that dropped out after the opening questions and entered the data mainly with 0 conflict and destruction per round. Erasing them and their groups and enemies only slightly alters the results into a more extreme direction.

Figure 2 shows that the average destruction in the baseline is constantly over 43, so over two third of the highest possible destruction of 60 Taler. The baseline data in figure 3 confirms that these results are stable over all three sessions. Thus, we have to reject H01 and state that there is a severe level of destruction. This suggests that the joy of destruction and negative reciprocity outweigh the countering factors mentioned in chapter “Design and Hypothesis”.





(H02) The destruction in the treatment group will be higher than in the baseline group.

Figures 2 & 3 strongly support H02 as they show that the average destruction in the treatment is higher than in the baseline. This holds for each round and session. The overall means are 45,53 in the baseline

and 53,97 in the treatment. The differences are highly significant[3] for round 2 and session 2 as well as the overall means. Differences for round 4 and session 3 are significant and for round 1 and 3 weakly significant. Although differences in session 1 are not significant, we see strong evidence for H02. This suggests that the bonus has a strong impact on the participants destruction decision.

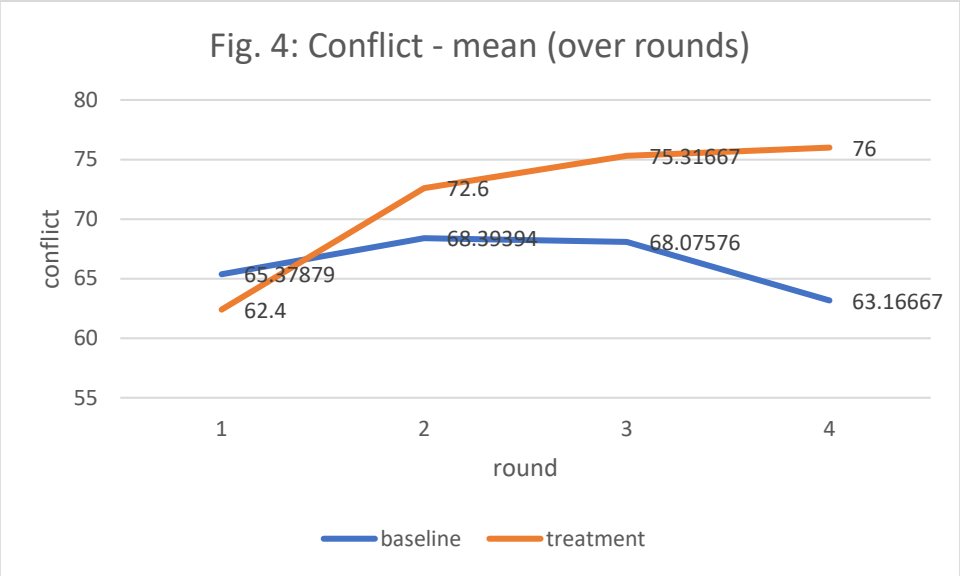
We can also observe that the slopes of treatment and baseline in figure 2 move almost parallel with differences only ranging from 8.15 to 8.8. This suggests that the bonus doesn't alter the slope but the level of destruction decisions.

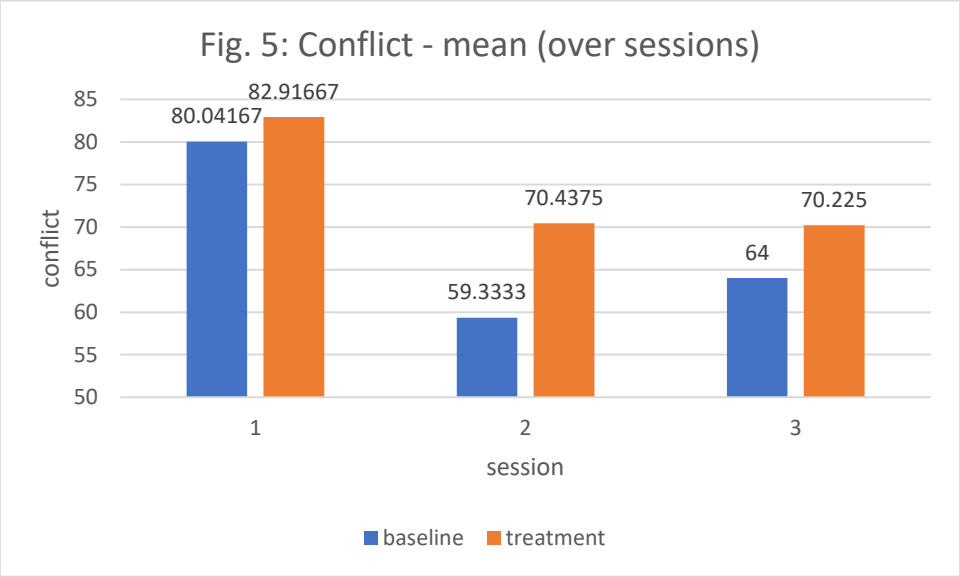
(H03) In the treatment, participants invest more in the conflict.

Figure 4 shows that on average treatment participants invested more in the conflict than baseline participants in rounds 2, 3 and 4. The total mean increased from 58.38 to 71.58. As figure 5 shows, this pattern holds for all sessions but to a quite volatile extent. Only session 2 and total means are highly significant. Round four is significant and round 3 and session 3 are weakly significant. Session 1, round 1 and 2 are not significant. Contrary to the other rounds, treatment participants invested less in round 1 than baseline participants. However, the Mann-Whitney test shows that with a value of 0.7998 this is insignificant. In conclusion we see H03 supported.

Furthermore, figure 4 gives some more information. It seems that participants in the baseline decreased conflict spending already after round 2, while conflict in the treatment escalated further, but with decreasing slope.

[3] All tests for significance are made with the Mann-Whitney test. We will refer to a value as highly significant if the test returns a value equal or below 0.01, as significant at a value of 0.05 or below and as weakly significant for 0.1 or below.





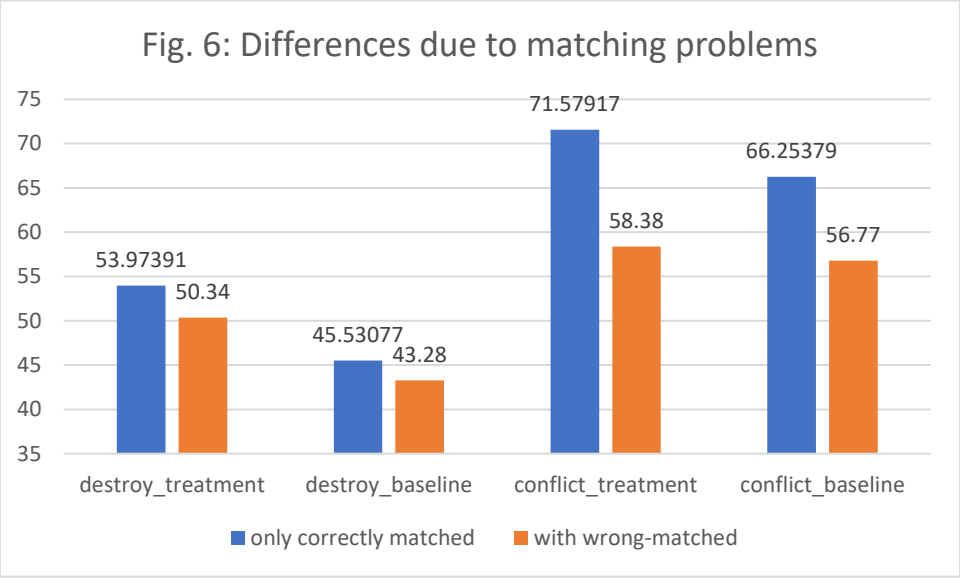
Discussion

As we conducted the experiment with classEx in lecture halls, we face some limitations. First, people were able to communicate with each other or look at mobile devices of neighbours. This could cause e.g. spillover effects or participants to feel more pressure to behave in a socially accepted way. Second, further spillovers might have occurred as participants were able to talk to future participants about the experiment. Fourth, lecturers were not the same in the classes, which might have biased the results.

Fifth, some participants dropped out of the game due to problems with the internet connection or the error-prone matching algorithm. As figure 1 shows, these dropouts were asymmetrically distributed over sessions and treatments, which might have biased our results as we already saw that there are session differences concerning conflict and destruction. For the first round we can

investigate the effect of the broken matching algorithm as participants did not know if other players were active or not[4]. Figure 6 shows that the main results for the entire data sheet (258 participants) with wrongly matched groups point in the same direction as the clean data sheet (126 participants). All the values are highly significant. This suggests that our findings include a bias due to asymmetrical dropouts.

Most of these limitations can be fixed by conducting the experiment in a lab setting. However, we decided to accept these possible problems, because the lecture hall increased the sample. This was important as we could only work with a minimum of six people or multiples of six. Moreover, anonymity would be endangered if only a small number of people participated in a session.



There is also a wide range of possible adaptions to the experiment. One that would be straightforward is asking for background information like age, gender and type of studies. Another one would be asking for expectations about other participants' behaviour. We decided not to integrate them as the experiment is already complex and conducting it within 20 minutes in a lecture hall does not provide opportunities for further explanations. A controlled lab setting again would provide this opportunity.

In addition, the lab would hopefully make it more possible to produce regression results that can be interpreted straightforward as they include less noise. In our case, regression outputs were highly volatile - with size, significance and goodness-of-fit changing strongly over rounds and sessions. We therefore decided to rely on the analysis of means.

[4] Inactive group members always played peacefully and contributed 0 to the conflict and did not destroy.

It also appears beneficial to adjust our lottery contest and include further treatments. By including different sized bonuses, we could try to disentangle the pure effect of implementing a bonus and the effect of its size. Finally, many participants voted for 60 Taler of destruction and we might have lost some observational power at the upper end of the scale. By reducing the amount that the losing team automatically loses we might cool down conflict and destruction. This would give us a wider range of observations.

Conclusion

We see that our adaption of the lottery contest managed to produce destruction as well as conflict decisions by the participants. The conflict did not completely settle or escalate, which we see as evidence that our contest design presents a slender alternative to the common lottery contest design in terms of observing decisions on violent behaviour.

The result shows that it is possible to stimulate a group, which is in the position to exert violence on another group, to start a competition for more violent behaviour. A small bonus on the highest proposal on destruction increased the average destruction from 45,53 to 53,97 Taler and seems to have the power to start an in-group competition for being cruel. A back on the envelope calculation shows that a bonus of 5 Taler on the highest destruction proposal leads to an increased destruction of 25.32 Taler[5] per round (not mentioning losses created by increased conflict spending). We also see that the proposal on the highest destruction increased the conflict over the rounds, while in baseline the conflict cooled down. The treatment overall increased conflict spending and prevented the conflict from settling.

Our results show that group conflicts and their confounding destruction can be stimulated by adjusting the incentives. Therefore, in behavioural economics these effects should be taken into account when scientists analyse conflicts.

[5] The three players of the losing group lost on average 8,44 more Taler in the treatment compared to the baseline.

References:

Abbink, Klaus, and Abdolkarim Sadrieh. "The pleasure of being nasty." *Economics Letters* 105.3 (2009): 306-308.

Abbink, Klaus, and Benedikt Herrmann. "The moral costs of nastiness." *Economic inquiry* 49.2 (2011): 631-633.

Andreoni, James. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The economic journal* 100.401 (1990): 464-477.

Bardsley, Nicholas. "Dictator game giving: altruism or artefact?." *Experimental Economics* 11.2 (2008): 122-133.

Chowdhury, Subhasish M., Dongryul Lee, and Roman M. Sheremeta. "Top guns may not fire: Best-shot group contests with group-specific public good prizes." *Journal of Economic Behavior & Organization* 92 (2013): 94-103.

Giamattei, M.;Lambsdorff, J. (2019): classEx —an online tool for lab-in-the-field experiments with smartphones, *Journal of Behavioral and Experimental Financ*e 22,pages 223-231.

Gneezy, Uri, and Rann Smorodinsky. "All-pay auctions—an experimental study." *Journal of Economic Behavior & Organization* 61.2 (2006): 255-275.

Juvonen, Jaana, and Adriana Galvan. "Peer influence in involuntary social groups: Lessons from research on bullying." (2008).Lee, Dongryul. "Weakest-link contests with group-specific public good prizes." *European Journal of Political Economy* 28.2 (2012): 238-248.

Mommsen, Hans. "Der Nationalsozialismus. Kumulative Radikalisierung und Selbstzerstörung des Regimes." *Meyers Enzyklopädisches Lexikon* 16 (1976): 785-790.

Prediger, Sebastian, Björn Vollan, and Benedikt Herrmann. "Resource scarcity and antisocial behavior." *Journal of Public Economics* 119 (2014): 1-9.

Waller, James E. *Becoming evil: How ordinary people commit genocide and mass killing*. Oxford University Press, 2007.

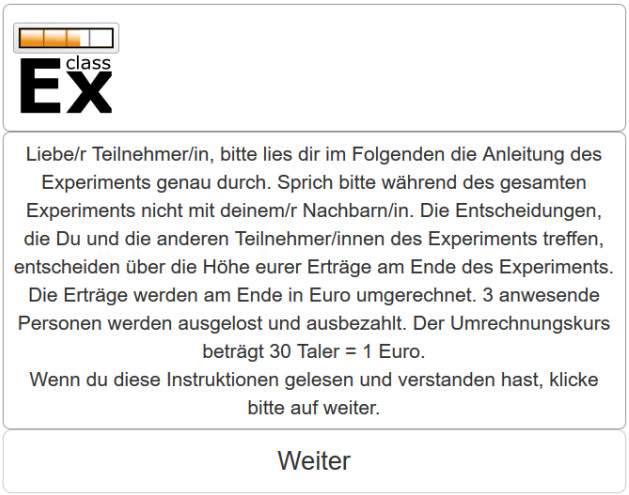
Tullock, Gordon. "Efficient Rent Seeking. JM Buchanan, RD Tollison, G. Tullock, eds., Toward a Theory of the Rent Seeking Society." *A & M University Press, College Station, TX: Texas* (1980).

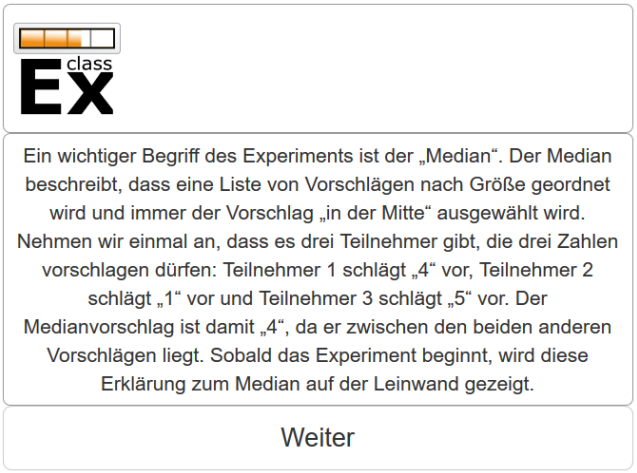
Zizzo, Daniel John, and Andrew J. Oswald. "Are people willing to pay to reduce others' incomes?." *Annales d'Economie et de Statistique* (2001): 39-65.

Appendix

ClassEx:

Instructions:





Das ist falsch. Richtig wäre „2“, da dieser Vorschlag zwischen den Vorschlägen „1“ und „7“ liegt.

Weiter

Das ist richtig, da Vorschlag „2“ zwischen den Vorschlägen „1“ und „7“ liegt.

Weiter

Das Experiment besteht aus 4 Runden. Du wirst in eine Gruppe mit zwei anderen Teilnehmer/innen gelost und ihr spielt gegen eine andere Dreiergruppe hier im Raum.

Am Anfang jeder Runde verfügst Du über 200 Taler. Davon können 100 Taler in Lose für einen Wettbewerb mit der anderen Gruppe investiert werden. Die Lose beider Gruppen werden in einen virtuellen Topf geworfen. Der Computer zieht zufällig ein Los daraus. Die Gruppe, deren Los gezogen wird, gewinnt den Wettbewerb. Die Chance, den Wettbewerb zu gewinnen, setzt sich wie folgt zusammen:

$$\text{Gewinnchance} = \frac{\text{Lose deiner Gruppe}}{\text{Lose beider Gruppen}}$$

Das heißt, die Gewinnchance steigt mit der Anzahl an Losen deiner Gruppe. Bei gleicher Anzahl an Losen (auch 0), beträgt die Gewinnchance 50 %.

Die eingesetzten Taler werden nicht zurückerstattet.

Weiter

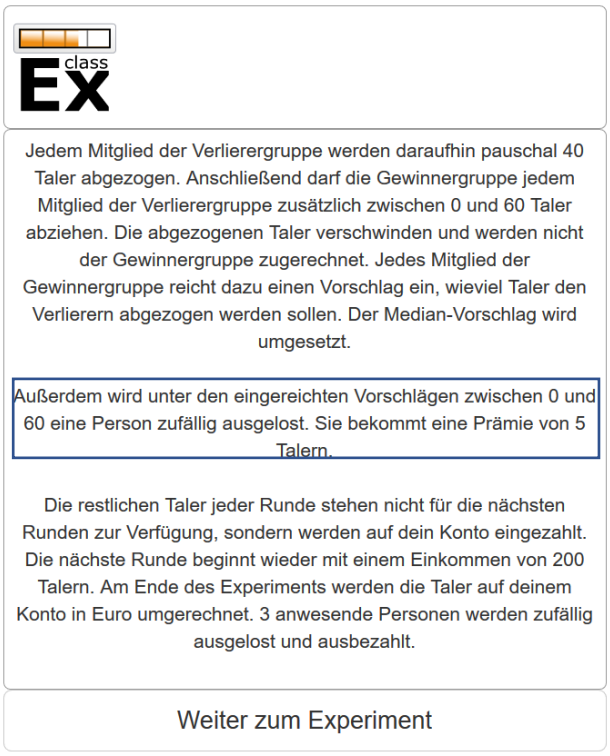


Figure 2: Instructions Baseline

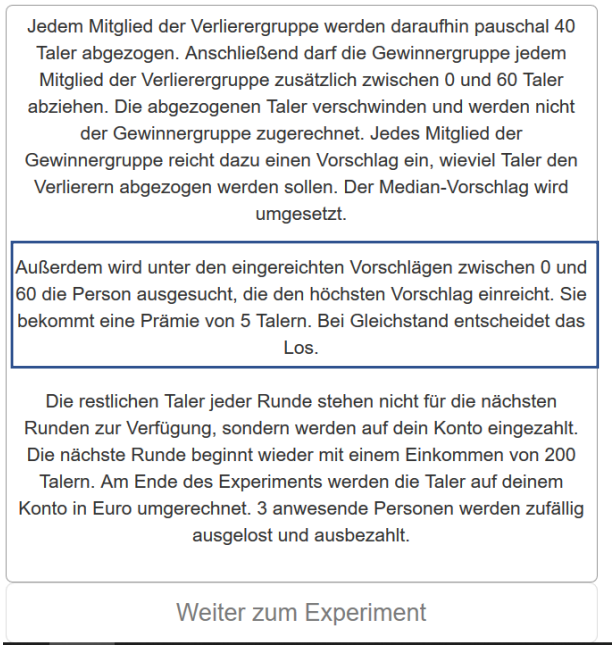
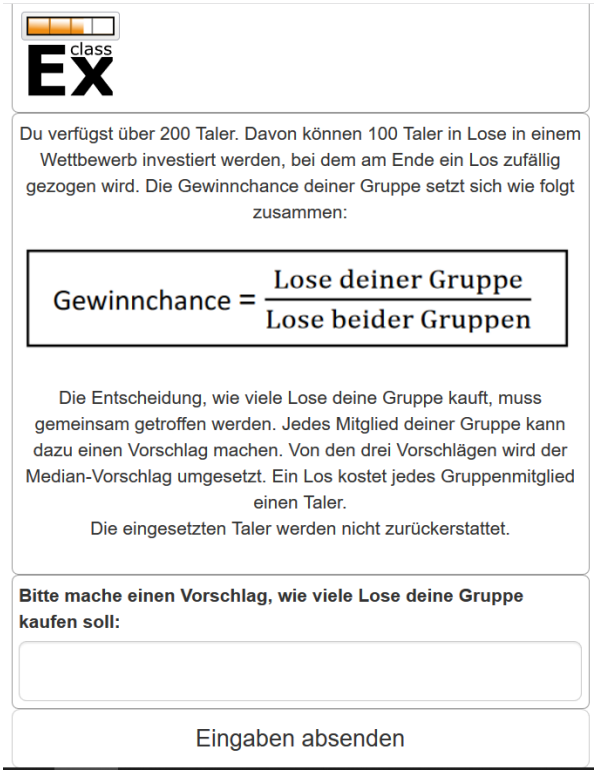


Figure 1: Instructions Treatment

First Stage: Lottery Contest



Second Stage: Destruction

Figure 3: Baseline – Gewinner

Figure 4: Treatment – Gewinner

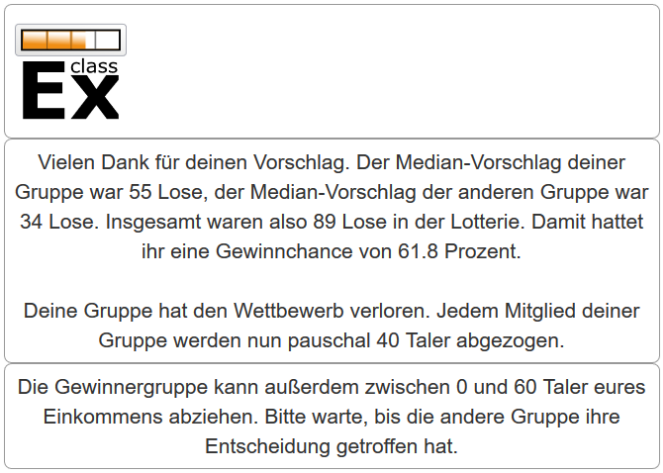


Figure 5: Traetment and Baseline: Verlierer

Start of the next round:

Die vorige Runde ist beendet. Im Folgenden siehst du das Rundenergebnis. Wenn du die Ergebnisse gelesen hast, scrolle bitte nach unten, um wieder einen Vorschlag zum Kauf von Losen für die neue Runde abzugeben.

Median-Einsatz deiner Gruppe: 34
Median-Einsatz der anderen Gruppe: 34
Pauschalabzug durch Niederlage: 0
Median-Abzugsentscheidung der anderen Gruppe: 0

Jemand anderes in deiner Gruppe wurde ausgelost und hat den Bonus von 5 Talern bekommen.

Deine Taler in dieser Runde: 166
Dein Kontostand, der am Ende des Spiels ausbezahlt wird: 166

Die neue Runde beginnt wieder mit einem Einkommen von 200 Talern.

Du verfügst über 200 Taler. Davon können 100 Taler in Lose in einem Wettbewerb investiert werden, bei dem am Ende ein Los zufällig gezogen wird. Die Gewinnchance deiner Gruppe setzt sich wie folgt zusammen:

$$\text{Gewinnchance} = \frac{\text{Lose deiner Gruppe}}{\text{Lose beider Gruppen}}$$

Die Entscheidung, wie viele Lose deine Gruppe kauft, muss gemeinsam getroffen werden. Jedes Mitglied deiner Gruppe kann dazu einen Vorschlag machen. Von den drei Vorschlägen wird der Median-Vorschlag umgesetzt. Ein Los kostet jedes Gruppenmitglied einen Taler.
Die eingesetzten Taler werden nicht zurückerstattet.

Bitte mache einen Vorschlag, wie viele Lose deine Gruppe kaufen soll:

Figure 6: Baseline - Gewinner - nicht ausgelost

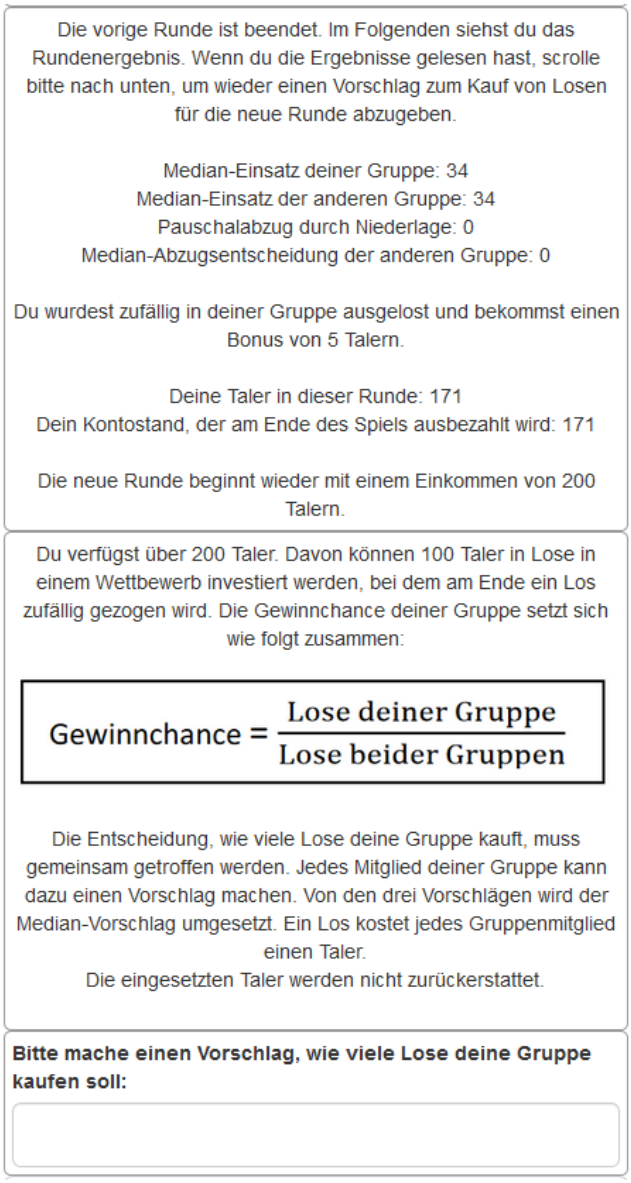


Figure 7: Baseline - Gewinner – gezogen

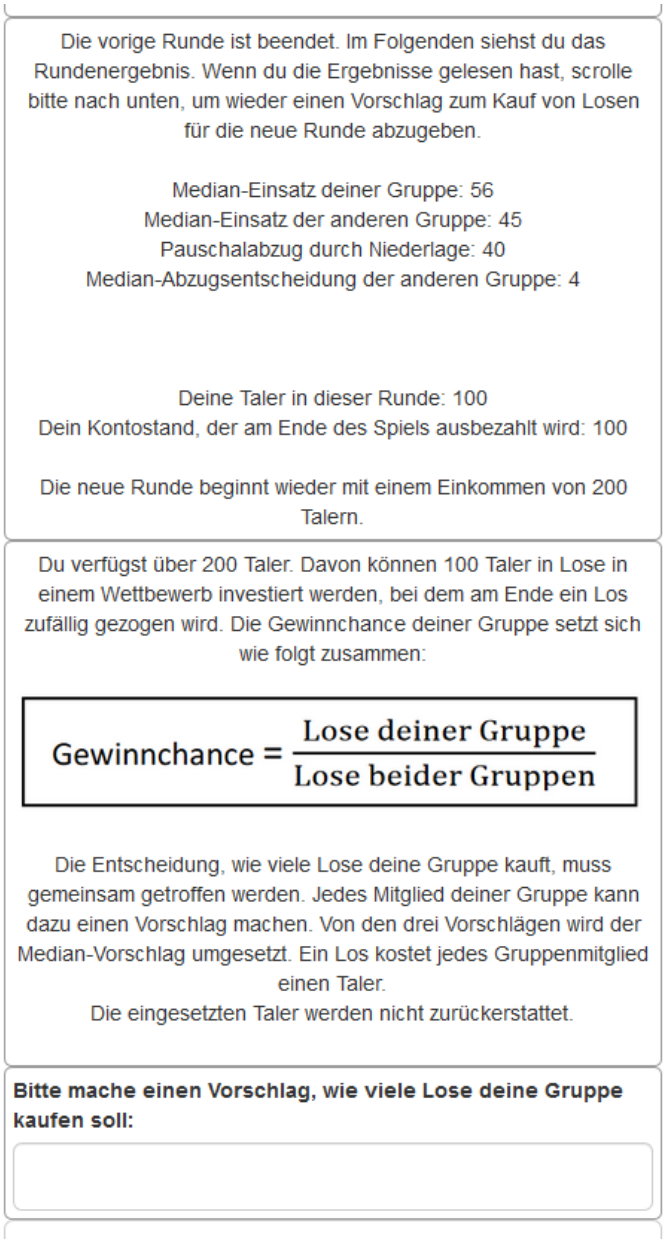


Figure 8: Treatment und Baseline – Verlierer

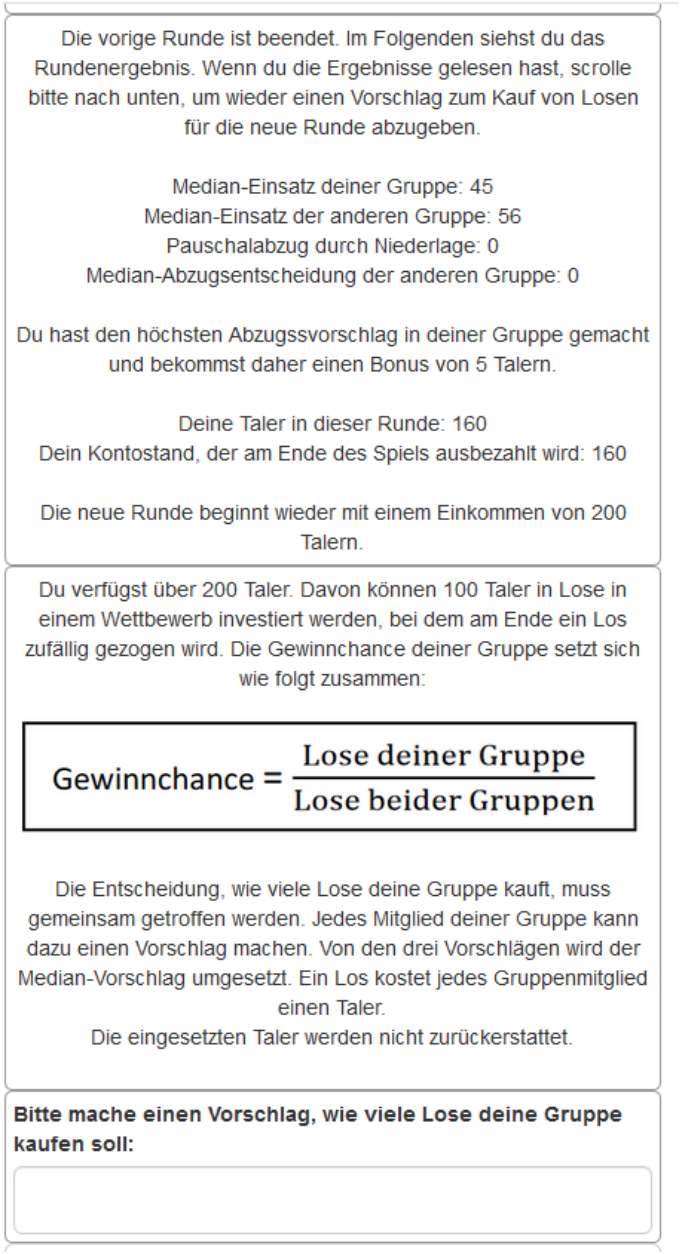


Figure 9: Treatment - Gewinner - höchster Vorschlag

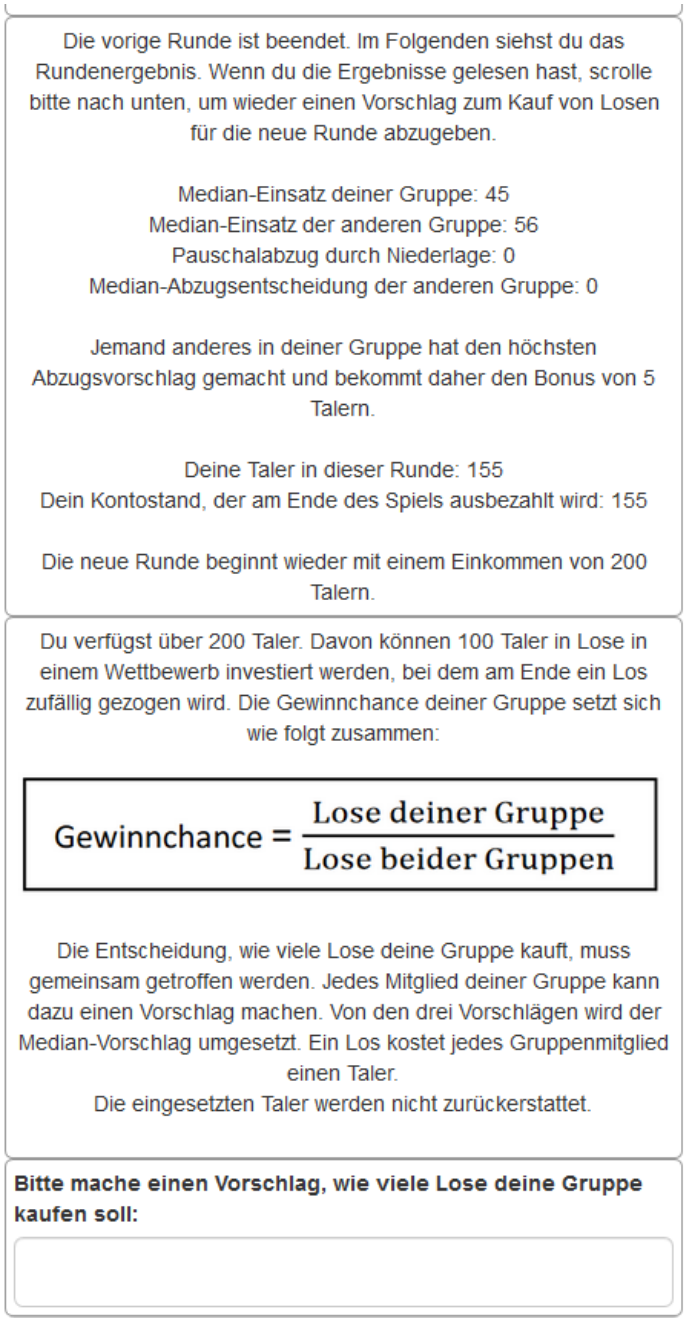


Figure 10: Treatment - Gewinner - nicht höchster Vorschlag

End:

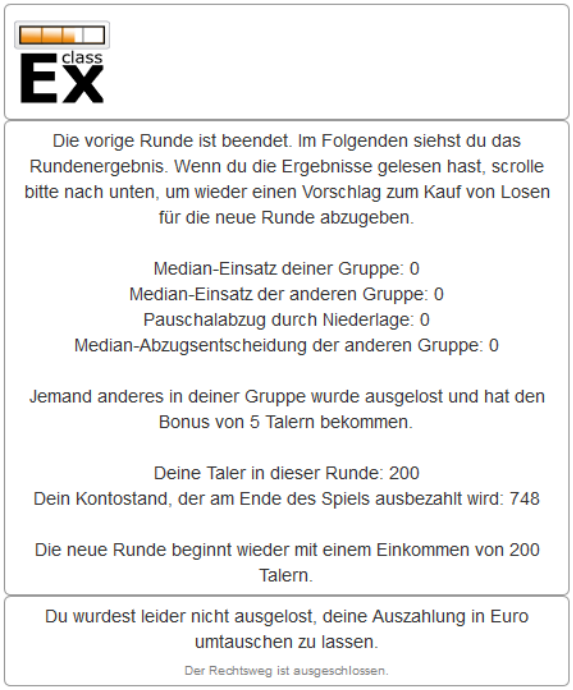


Figure 11: Nicht ausgelost

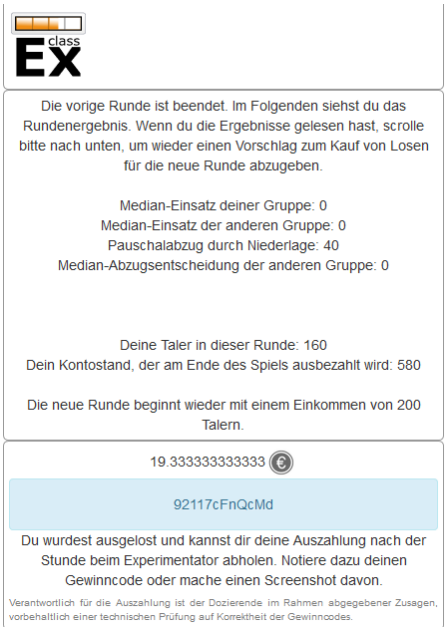


Figure 12: Ausgelost