

Reward and punishment in a team contest

Citation for published version (APA):

Heine, F. A., & Strobel, M. (2015). *Reward and punishment in a team contest*. Maastricht University, Graduate School of Business and Economics. GSBE Research Memoranda No. 034
<https://doi.org/10.26481/umagsb.2015034>

Document status and date:

Published: 01/01/2015

DOI:

[10.26481/umagsb.2015034](https://doi.org/10.26481/umagsb.2015034)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Florian Heine, Martin Strobel

**Reward and Punishment in a
Team Contest**

RM/15/034

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O. Box 616
NL- 6200 MD Maastricht
The Netherlands

Reward and Punishment in a Team Contest

Florian Heine* and Martin Strobel

Department of Economics, Maastricht University, The Netherlands

November 2015

Abstract

A team contest entails both public good situations within the teams as well as a contest across teams. In an experimental study, we analyse behaviour in such a team contest when allowing to punish or to reward other group members. Moreover, we compare two types of contest environment: One in which two groups compete for a prize and another one in which we switch off the between-group element of the team contest. Unlike what experimental studies in isolated public goods games indicate, we find that reward giving, as opposed to punishing, induces higher contributions to the group project. Furthermore, comparing treatment groups, expenditures on rewarding other co-players are significantly higher than those for punishing. This is particularly pronounced for the between-group contest.

Working paper

*Corresponding author. Tel.: +31 43 388 2772. *E-mail address:* f.heine[at]maastrichtuniversity[dot]nl

1 Introduction

Cooperation problems within groups are often modelled as public goods games, where personal payoffs depend on the actions of the own group. In the field however, groups frequently do not possess this stand-alone character. Their weal and woe also depends on the actions of competing groups. This change in (economic) environment can have major impact on group dynamics. For example, consider a regatta where the final outcome (of winning it or not) depends on the relative functioning of the group. Put simply: Is your group faster than the other one? In this article we investigate how reward and punishment systems work if two groups are in a contest for a group prize.

Cooperation in alliances or social groupings is commonly observed in many applications. By working together, everyone benefits from the efforts of the other, resulting in a more efficient outcome than if everyone was working on their own. From the classical economic point of view, though, rational individuals do not contribute at an optimal level. The reason for this is a combination of coordination problems and individualistic preferences. As a result, positive welfare effects from efficient organisation cannot come to play.

Previous literature suggests that the opportunity to give punishment and / or reward can overcome the free-rider problem in cooperative games ([Rand, Dreber, Ellingsen, Fudenberg, and Nowak, 2009](#); [Sefton, Shupp, and Walker, 2007](#); [Sigmund, Hauert, and Nowak, 2001](#); [Fehr and Gächter, 2000](#)). Furthermore, punishment seems to work better in this regard than reward ([Balliet, Mulder, and Van Lange, 2011](#)). The results are not yet fully conclusive, but already give a first indication for policy advice: Use the stick rather than the carrot.

Isolated tests of public goods games, however, ignore the large class of situations where groups are not detached, but embedded in some form of competition. Examples for this are public tenders, sports team competition or alliances in military conflicts. We name these situations team contests or group contest games. Teams, groups and alliances are often built on a voluntary base. Hence, it is difficult to imagine that punishment provides the glue to hold them together when facing off against a common competitor. Imagine a football team when things are going badly. Does punishment from some team members to others make these others provide more effort? Indeed, studies suggest that “in-group cohesiveness is associated with positive treatment (...) of the in-group” ([Stephan, 1985](#)).

Furthermore, when there is no distinction between an in-group and an out-group, like in public goods games, “people are unlikely to attend to the fact that they are interacting with in-group members” ([Stephan, 1985](#)). The differentiation between an own group and an out group creates a favouring atmosphere towards the own group from the get-go ([Brewer, 1979](#)) and can lead to the “ultimate attribution error” ([Pettigrew, 1979](#)). This is a social schema to explain away positive actions of out-group members as exceptions to the rule, for example.¹

Our goal is to compare reward and punishment mechanisms in team contests. Specifically, we consider, which mechanism leads to more effort and the efficiency implications. We find a higher contribution level to the contest in the reward treatments, compared to the punishment treatments. Also, as well comparing between and within treatments,

¹[Pettigrew \(1979\)](#) mentions that this pattern of attributions is greatest in groups with a history of conflict.

players distribute more reward than punishment. This gives a new perspective on group dynamics in interactive economic games. We argue that the environment a group faces, has significant influence on its social dynamics and preferences within the group. In line with the vast majority of research on rent-seeking or contest games, we find a drastic over-contribution across all our treatments, as compared to the Nash-equilibrium.² This leads to an inefficient outcome and rent over-dissipation.

The remainder of this article is structured as follows: First, we discuss the nature of reward and punishment systems in economic games; in Section 3 we explain the setup of the experiment and the equilibrium prediction; Section 4 contains our results and interpretation; and in Section 5 we provide concluding comments and some suggestions for further research.

2 Reward and Punishment in Economic Games

Providing the option to punish and / or reward other group members in public good games has repeatedly led to higher contributions in both experimental (Fehr and Gächter, 2000; Sefton et al., 2007; Rand et al., 2009; O’Gorman, Henrich, and Van Vugt, 2009) and theoretical (Sigmund et al., 2001; Sasaki, Brännström, Dieckmann, and Sigmund, 2012) studies. Furthermore, it seems that punishing works slightly better than rewarding in rendering higher contribution. Balliet et al. (2011) conduct a meta-analysis of the effectiveness of reward or punishment in social dilemmas. Their analysis includes 187 effect sizes from 76 papers. They find a higher effect of punishment for generating cooperation as compared to reward. Although being just marginally significant, a definite trend is evident towards a higher effectiveness of punishment in this regard.

Things might look differently though, when groups compete for a group prize. Previous studies suggest that engagement in a project or a group prize is higher if there is a conflict or a competition with another group (cf. Sherif, 1961; Bornstein and Ben-Yossef, 1994; Tan and Bolle, 2007; Goette, Huffman, Meier, and Sutter, 2012). At the same time, in-group favouritism and out-group spite increases, a desire to benefit players from the own group and harm those from the competing group (Konrad and Morath, 2012). This can change the relative effectiveness of rewarding and punishing, as well as players’ preference for either rewarding or punishing as means to reciprocate other group mates’ actions.

There are two groups of psychological motivations for an in-group – out-group bias in a competitive game like the contest game. The first one considers cognitive factors. Doise (1976) argues that being categorised in competitive groups, an anticipatory-justification process is active, devaluing one’s antagonists. Alternative explanations for the in-group – out-group bias consider motivational factors. According to Tajfel (1978), people desire to compare the in-group in a favourable way towards the out-group.

These studies suggest that the economic environment shapes social preferences. In changing the economic environment of a public goods game, introducing a contest for a group prize between groups, we expect to find contrasting results. In public goods games there is no competition between groups, but investment is socially beneficial, not unproductive effort, as in the contest game. Furthermore, reciprocity is one social preference

²Cf. Dechenaux, Kovenock, and Sheremeta (2012); Öncüler and Croson (2005)

and it is a prominent cause for conditional cooperation (play cooperatively if others do) (Gächter and Herrmann, 2009).³ The degree at which reciprocity is endogenous towards the economic environment determines whether a change in players' behaviour can be found.

Another major issue in experimental studies on contest games is players' persistent tendency to over-contribute and the ensuing effects on efficiency (cf. Öncüler and Croson, 2005). In an experimental study, Abbink, Brandts, Herrmann, and Orzen (2010) find that allowing for the opportunity to punish other group members in a contest game renders expenditure levels that are as far as 60 % above those in the control treatment. Although it would be socially beneficial to reduce contributing to the contest, players incentivise group mates to increase their spendings to the contest. As a consequence, introducing the option to punish other group members leads to even higher over-contribution than in the control treatment with negative effects on total welfare. Including the costs for punishment, subjects' material losses are almost 900 % of the equilibrium level.

3 Setup

Our project brings together the two concepts of interaction in a contest game on the one hand and reward and sanctioning systems on the other. Experimental research in this field has mainly focused on either of the concepts. In combining these systems, insight can be gained into the interaction of the institutional system of rewarding / punishing with the contest / non-contest-environment. More specifically, we also consider what effect the presence of another competing group has in this regard. In related experiments by Sefton et al. (2007), agents are not in a competitive situation and in Abbink et al. (2010), participants have no opportunity to give rewards.

Comparison of contest vs. non-contest requires a design that makes both comparable. To this end we modify the classical contest environment. In the non-contest environment we substitute the rival group by an equilibrium playing virtual mechanism. To subjects of the prevailing group this means that they are acting in a public good game in which overall effort decides about the likelihood to win the prize. Players are automatically sorted in groups of four. In the field, agents repeatedly interact with the same set of others; hence also throughout the experiment participants play with the same players both in the own and in the opponent group.⁴ Each player also keeps the same label throughout the experiment. The experiment is conducted over 15 rounds. This and all other features of the experiment are disclosed and are known to the participants.

The experiment is set up in a 2 x 4 design, with each four treatments with and four treatments without opponent group. In the treatments with opponent group, every group K of four subjects competes against another group M of four subjects. They compete by buying lottery tickets for their own group. Then one ticket will be drawn by the computer and this group wins a fixed prize. For the groups without opponent group, the other group is replaced by a fixed number of 25 blank tickets which do not result in rent payments if drawn. Furthermore, in the latter treatments, there is no mentioning of other groups at any time. Every round is partitioned into three distinct stages as described in what follows:

³Another cause for conditional cooperation would be conformity – players just do what others do.

⁴Not all treatments have an opponent group though.

1. Each player receives an endowment (budget) of $B = 100$ tokens and decides how much of it to invest in order to buy lottery tickets. The price for a ticket is one token. Investment of subject $k \in K$ of group K is labelled v_k and $m \in M$ of group M is v_m . All tokens that a player does not invest will be added to her private account.
2. The winning probability, or contest success function, is as in [Tullock \(1980\)](#); [Katz, Nitzan, and Rosenberg \(1990\)](#). For the contest environment:

$$p_K((v_k)_{k \in K}, (v_m)_{m \in M}) = \begin{cases} \frac{\sum_{k \in K} v_k}{\sum_{k \in K} v_k + \sum_{m \in M} v_m} & \text{if } \max_{i \in K \cup M} \{v_i\} > 0 \\ 1/2 & \text{otherwise} \end{cases} \quad (1)$$

For the non-contest environment:

$$p_K((v_k)_{k \in K}) = \frac{\sum_{k \in K} v_k}{\sum_{k \in K} v_k + 25} \quad (2)$$

where p_K is the probability that group K wins (over group M). For the treatments without opponent, group M is replaced by a total of 25 blank tickets, which is known to all players. As will be discussed in Section 3.1, this represents the Nash equilibrium group contribution. For this, imagine $\sum_{m \in M} v_m = 25$ constantly in equation (1) to represent the blank tickets as in equation (2). The winning probability is the sum of all lottery tickets of the own group divided by one of the two (depending on the environment): *a*) the total lottery tickets bought by both groups, or *b*) the sum of total lottery tickets and blank tickets. After each investment phase, one ticket is drawn by the computer. Hence, the more tickets a group buys, the higher the chances of winning the lottery (*ceteris paribus*). Players of the winning group each receive a prize of $z = 100$ tokens, the other group gets nothing. In the non-opponent treatments, if a blank ticket is drawn, the prize is forfeited.

3. The players get to know whether their group has won or not and how much the others in their group, each, contributed. They also get to know the number of lottery tickets bought by the opponent team, but not the opponent players' individual contributions (if applicable). As well, participants get to know, what their probability of winning was, which takes into account the sum of contributions of the opponent group / blank tickets. They now receive another $F = 50$ tokens (response tokens),⁵ which they can either keep in their own account or spend to give a response to their own group members.⁶ To this end there are four treatments, each played with and without opponent (thus 2 x 4 setup):

[Baseline:] Players receive the aforementioned 50 response tokens at the end of each round to be added to their account directly. As in this treatment there is no further interaction between the players, these tokens are presented as extra tokens.

[Reward:] Players can reward co-players in their own group. For this, they assign response tokens to one or more players of their own group. Each response token

⁵These tokens were called feedback tokens in the instructions (see Appendix A).

⁶Throughout the article we regularly use the expression response when referring to the reward and punishment mechanism in order to ease the reading flow. As well, we make use of the term sanctioning as synonymous term for punishment.

player k assigns to another player l is subtracted from k 's account and added to l 's account. However, k can also keep all response tokens for herself and let them be added to her own account. The players can at most invest all 50 response tokens per period (as sum of all response assigned to all co-players) and they cannot save tokens to be spent in upcoming rounds.

[*Punish:*] This is similar to the reward treatment, but with punishment instead. Consider again two players, k and l . k can assign deduction points to any other player of her own group, say l . For this, the same amount of response tokens spent by k , is deducted from l 's account. In this treatment, it is possible that a player gets punished such that her round payoff would theoretically turn negative. However, if a player gets punished by more tokens than there are on her account for this period, the excess punishment is discarded and the round payoff set to zero.⁷

[*Reward and Punishment (R&P):*] In this treatment each player can choose to either reward or to sanction another co-player of the same group. Players can give both rewarding and sanctioning response to different co-players in the same round;⁸

The prize to be won, z , constitutes a sort of local public good, as its consumption is non-rival and non-excludable within the group: It is independent of group size, all group members receive it and no-one can be excluded.

Note that players do not get to know from whom they get response points assigned. So neither do they know in the punishment treatment, which of the co-players sanctioned them, nor do participants get revealed by whom they receive reward tokens in the reward setup (equivalent for the R&P treatment). This is to circumvent that players react on response behaviour of particular co-players – think hereby of retaliation or gift exchanging. Furthermore, participants do not get to know about the response other players received. It also needs to be pointed out, that by rewarding other group members, participants shift around tokens in a way that leaves total welfare unchanged. Punishing others, in contrast, reduces overall welfare.

We apply a cost for response-giving of one (as in [Sefton et al., 2007](#)). This means if player k punishes player l by 6, k has to pay 6 for this action. Some studies incentivise response-giving by relatively cheapening it (cf. [Abbink et al., 2010](#); [O’Gorman et al., 2009](#)).⁹ [Fehr and Gächter \(2000\)](#) apply a non-linear cost function, where low punishments cost one and higher punishments are more costly. We opt for a constant cost of one in order to keep both the punishment and the reward leverage small to mitigate confounds from efficiency preferences. Also, in case we applied unequal leverage for the two types of response, the difference could be confounding as well.

3.1 Equilibrium Strategies

After having cast the decision about how much to invest in the contest, players get to know whether they have won, what their winning probability was and how much the group members of the own group contributed. Now they can assign response points to other

⁷This case never occurred in the experiment.

⁸It is not possible to both reward and sanction the same player in one round.

⁹For convenience, imagine the cost to be one half. Sticking to the aforementioned example, if player k punishes player l by 6 again, k now has to pay 3 for this action.

group members, as discussed above. Under the Nash-logic, costly response only reduces own payoff, so individualistic subjects do not give response.

Before that, players decide on how many lottery tickets to buy. In standard economic theory, players individually maximise their expected payoffs, being

$$\pi(v_l) = \frac{v_l + \sum_{\substack{k \neq l \\ k \in K}} v_K}{v_l + \sum_{\substack{k \neq l \\ k \in K}} v_K + \sum_{m \in M} v_M} \cdot z - v_l$$

in this experiment. This is the [Tullock \(1980\)](#) winning probability (as introduced in [Section 3](#)) times the prize minus the cost for investing in the group project. For treatments without opponent, set $\sum_{m \in M} v_M = 25$. In this type of contest game with homogeneous groups, first order conditions only determine a unique equilibrium with respect to the aggregate group contribution. On the individual level multiple equilibria exist (as explained in [Konrad, 2009, 2007; Katz et al., 1990](#)).

For all treatments the equilibrium investment level for group K is $\sum_{k \in K} v_k = \frac{z}{4}$.¹⁰ For this experiment with $z = 100$, this results in a Nash-equilibrium of 25 tokens per group.¹¹ One conceivable solution to the within-group “burden-sharing” would be for every player to contribute equally. This would result in an individual contribution of $v_l = \frac{z}{4 \cdot |K|}$ or 6.25 tokens, given a group size of $|K| = 4$. The same solution would also emerge if we assume symmetry as in [Katz et al. \(1990\)](#). Of course, this is only one example among other possible solutions to the within-group problem. Note that only integer amounts of investment are allowed, so 6.25 can best be viewed as a theoretical benchmark.

We start the analysis at players’ last decision in period 15 and work our way back to the beginning. Going further in our line of reasoning, also in the penultimate period, players would not give response. The same argument applies as above. In the same sense, contribution would be the same as in the last round. This logic proceeds until the first period, giving the same values as in period 15 as Nash equilibrium.

This, however, is substantially different from equilibria that apply for conventional public goods games, as in [Sefton et al. \(2007\)](#), for example. In a public goods game, the Nash equilibrium would be not to contribute at all to the group project. The following thought experiment clarifies the mechanics leading to this difference: Consider the case where no-one contributes anything to the group project. In a public goods game, by unilaterally deviating from zero contribution, players only decrease own payoff. In a contest game, however, winning probabilities change from a 50:50 chance (default option if no-one contributes) to 100 %.¹² In terms of a lottery game, as presented above, there is only one group buying a ticket, hence it will definitely win. This is not only beneficial for the own group, but also individually rational. Intuitively speaking, in a contest game, Nash equilibrium justifies investing “a bit”, in order to have a few lottery tickets in the game.

Unlike for the Nash-equilibrium, the social optimal strategy differs somewhat between

¹⁰Find a derivation in [Appendix C](#).

¹¹Note that the equilibrium investment level is independent of the endowment with tokens or group size. The only factor influencing equilibrium behaviour at the group level is the prize at stake.

¹²The same also holds for the alternative default option that no group wins if both contribute zero. Winning probability would switch from 0 % to 100 %.

the cases with or without an opponent and whether one regards the opponent team as part of the social system. In the treatments without opponent, the prize is forfeited and goes to nobody if it is not won. Hence, the Nash equilibrium coincides with the optimal strategy for maximising expected total monetary welfare in these treatments. Things look differently, however, if two groups compete against each other. Now the optimal strategy to maximise total monetary welfare across groups would be for both groups to invest nothing and face a 50:50 chance of winning. The proof for this is straightforward: First of all no reward or punishment will be distributed, as this is either welfare neutral (rewarding) or reducing welfare (punishing). Buying lottery tickets is unproductive effort that only influences winning probabilities. One of the two groups wins the prize anyway and regarding total welfare of both groups, it does not matter, which of the two wins it. These steps of analysis are true for all periods of the game. If the opponent group is not regarded as part of the social system, then the Nash equilibrium would constitute the strategy that maximises expected total monetary welfare for the group.

3.2 Realisation of the Experiment

We used the system ORSEE by Greiner (2004) to recruit a total of 372 participants (most of them students) for our experiment. Each participant received a financial compensation for taking part in the experiment. This compensation was dependent on the total amount of tokens earned over the fifteen rounds. The experiment took about one hour, including reading the instructions, a trial period, the contest game as such, a questionnaire and payment. The mean income was €16.94 across all treatments.^{13,14} While players in treatments without opponent received an average of €19.61, players in treatments with opponent received €15.60 on average.

The experiment was conducted as computerised laboratory experiment¹⁵ in the facilities of the BEElab (Behavioral & Experimental Economics Laboratory, Maastricht University, Netherlands). Each participant sat in their own cubicle, where she made her decisions, physically and visually separated from other participants. Upon entering the cubicle, each participant found the instructions at her place.¹⁶ After reading the instructions, participants answered some trial questions, in order to become familiar with both the conceptual setup of the game, as well as with its user interface. When the experiment was finished, players were asked to fill out a short questionnaire about personal features to be used in the analysis of the game.

In the trial phase, players were confronted with randomly generated game situations and answered control questions. They were aware of the randomness of the numbers presented and also knew that they were not yet interacting with other participants. In general, participants got to see three different screens per period: First there was a screen where each player decided how much to contribute for the group account. Then the player got to know whether her group has won, own winning probability and the contributions of fellow group members. She then had the opportunity to assign response tokens to co-players. In the baseline treatment, this was just an overview page, without the opportunity to assign response tokens. On the third screen, the total profit (in tokens) for this period was

¹³About £14.05 or \$22.48 under the exchange rate at the time of the experiment.

¹⁴Consider 4.4 for the effects of over-contribution on final payoff.

¹⁵For this we made use of the experimental software “z-Tree” by Fischbacher (2007).

¹⁶Find a copy of the instructions in Appendix A.

displayed in a detailed overview for the participant.¹⁷

4 Results

We begin by comparing contribution levels and response over time. Then we examine, which effects drive the differences in behaviour at the individual level. We find that players' actions are heavily influenced by what happened in previous rounds. Therefore we take a closer look at the dynamics of this experiment in Section 4.6. There is a strong relationship between contribution to the group account and response-giving. We analyse the relationship between contributions and receiving reward or punishment in Section 4.7.

4.1 Statistical Methodology

For the analysis we make use of the statistical software Stata. The data is not independently distributed, as the actions of other players and previous rounds influence own behaviour. In line with this, contribution shows a high degree of autocorrelation and heteroscedasticity.¹⁸ This is why we use Newey-West standard errors for our regressions (as devised by Newey and West, 1987) with a lag length of two periods.¹⁹

As alternative regression method, we also apply Tobit regressions. This is indicated as the dependent variables are censored from below at 0 and from above at 100 for contribution and at 50 for response tokens, respectively. As in Sefton et al. (2007) we use standard errors that cope for heteroscedasticity and clustering between subjects, as described in Angrist and Pischke (2009, Chapter 8).

Since our data is not normally distributed²⁰ we use non-parametric methods to test the hypotheses: Mann-Whitney U tests (MWU) (Mann and Whitney, 1947) for independent sample tests and Wilcoxon signed-rank test (Wilcoxon, 1945) for paired tests. While in the non-contest environment, group level data (four players) constitutes an independent observation, we use paired group data (eight players) in the contest environment.

4.2 Contributions to the Group Account

Average group level contributions over all treatments are significantly higher than our benchmark, the Nash equilibrium (as deducted in Subsection 3.1).²¹ This is also true for both environments and treatments, if analysed separately.

¹⁷Find an example of the three stages in Appendix B.

¹⁸Level of Both the White's test and the Beusch-Pagan test constantly report high evidence for inhomogeneous variance.

¹⁹Rule of thumb by Stock and Watson (2012): use lag length of $0.75 \cdot T^{1/3}$ with T being the number of rounds in the experiment. The Newey-West autocorrelation consistent covariance estimator is nicely explained in Greene (2012, pp. 960-961).

²⁰It is skewed to the right and slightly platykurtic: Shapiro-Wilk test for normality. H_0 : Contribute $\sim N(\mu, \sigma^2)$, $P = 0.00$. Skewness: 0.82, Kurtosis: 3.15

²¹Wilcoxon test: H_0 : group contr. = 25, H_1 : diff $\neq 0$, $N = 62$. $P = 0.000$. Higher rank sum than expected.

Figure 1 illustrates average contribution to the group account per player over all 15 rounds. Results for the contest environment are on the left and for the non-contest on the right. For both environments, contribution levels for the punishment treatments are considerably lower than in the reward treatments.²² Furthermore, there is even a positive trend in contributions for the reward treatment in the contest environment until the 12th round.

Most striking, however, is the result for the R&P treatment in the contest environment, which clearly peaks out over all other ones.²³ Also, the comparably high contribution rate for the baseline treatment in the contest environment is pretty remarkable.

In Figure 2, average group contribution per treatment can be compared. The error bars give the 5 % confidence interval, on the basis of independent observations, which is the average contribution per pair of groups over all 15 periods.

By and large, contribution levels averaged over all rounds are on a fairly similar level for the reward treatment in both environments. Averaged contribution levels drop in all other treatments, compared between the environments. At the same time, though, even within the treatments, there exists considerable heterogeneity between groups' contribution levels.

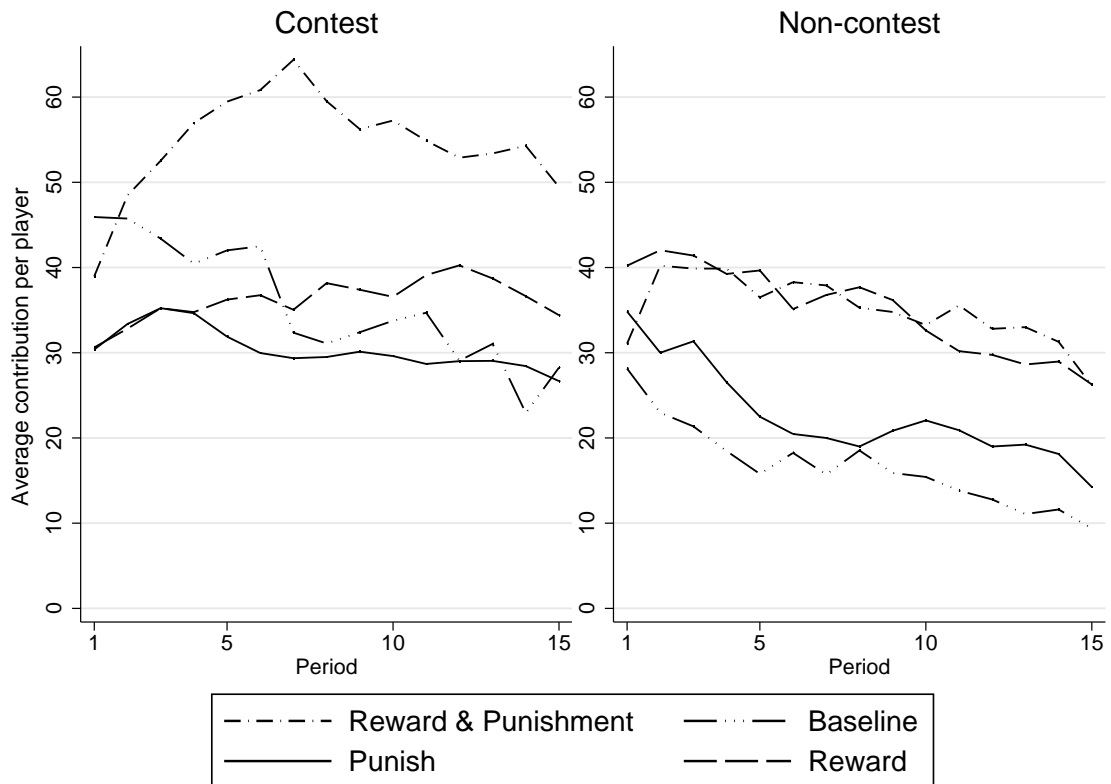


Figure 1: Individual contribution to the group account per treatment. Contest environment on the left, non-contest environment on the right.

²²Wilcoxon test: $H_0: \text{diff} = 0$, $H_1: \text{diff} \neq 0$, $N = 32$. $P = 0.024$. Higher rank sum than expected for reward treatments.

²³MWU test: $H_0: \text{diff} = 0$, $H_1: \text{diff} \neq 0$, $N = 62$. $P = 0.001$. Higher rank sum than expected for R&P treatment.

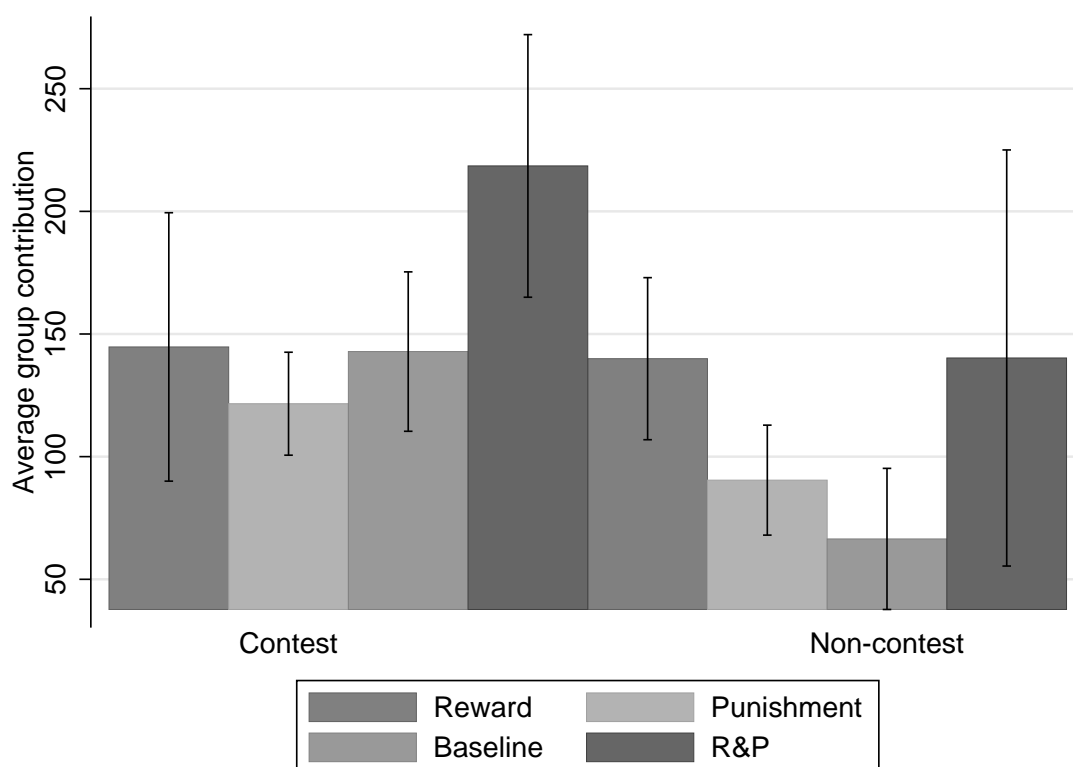


Figure 2: Average contribution to the group account per treatment. Contest environment on the left, non-contest environment on the right.

For a break down into individual group behaviour see Appendix D.

4.3 Response giving

As Figure 3 illustrates, also for the response-giving, a systematic difference can be identified between the treatments, but also within the R&P treatment. As for Figure 1, we report mean values for each treatment and period per player. The two graphs in the left column depict average spending levels for the treatments with response. On the right hand column, rewarding and punishing in the R&P treatment are compared. The first row is for the contest and the second row for the non-contest environment.

Comparing rewarding and punishing behaviour, it seems that the former is slightly more extensively used, especially in the beginning of the experiment. In the R&P treatments (analysing both environments together and separately), rewarding is significantly higher than punishing.²⁴ This relationship also holds between the reward and punishment treatments when both environments are analysed together,²⁵ but does not deliver a significant result for the separate environments.

²⁴E.g. Wilcoxon test including both environments: H0: diff = 0, H1: diff \neq 0, N = 16. P = 0.03.

²⁵MWU test: H0: diff = 0, H1: diff \neq 0, N = 32. P = 0.029. Higher rank sum than expected for reward treatments.

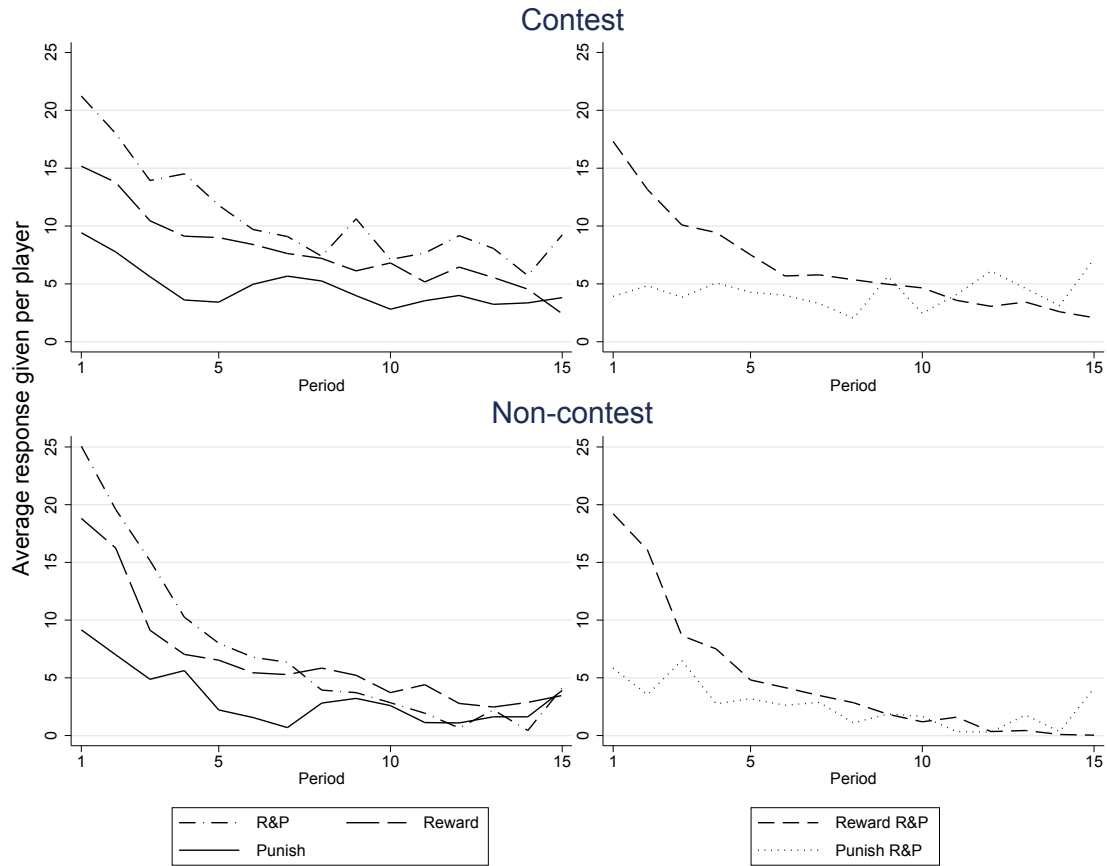


Figure 3: Response given per treatment

Especially in the beginning of the game, reward giving is used more than punishing, while the two approach each other over time, even turning punishing into the preferred form of response giving for the last period in the R&P treatment.²⁶ We conclude that players tend to behave more favourable towards other group members in the beginning, which confirms our hypothesis. In the course of the game, however, friendliness fades out.

Also, overall response giving is significantly higher in the R&P treatment than in the other treatments.²⁷ Especially in the contest environment there seems to exist a somewhat robust demand for rewarding or punishing, respectively. This means that subjects seem to not have a mental budget for overall response.²⁸ Instead, if subjects can spend resources on both reward *and* punishment, they seem to spend as much on each of the two, as they would if they could only spend the resources on one of the two. For this consider also Figure 4.

Table 1 and 2 show the propensity to engage in rewarding and / or punishing behaviour for each environment. It shows the percentage of non-negative response per treatment and the type of response being sent. While the propensity to punish is fairly similar between the

²⁶E.g. Wilcoxon test including both environments: $H_0: \text{diff} = 0$, $H_1: \text{diff} \neq 0$, $N = 16$. $P = 0.012$.

²⁷MWU test: $H_0: \text{diff} = 0$, $H_1: \text{diff} \neq 0$, $N = 48$. $P = 0.001$. Higher rank sum than expected for R&P treatments.

²⁸Like what they actual received as budget in the experiment, which was for response in general.

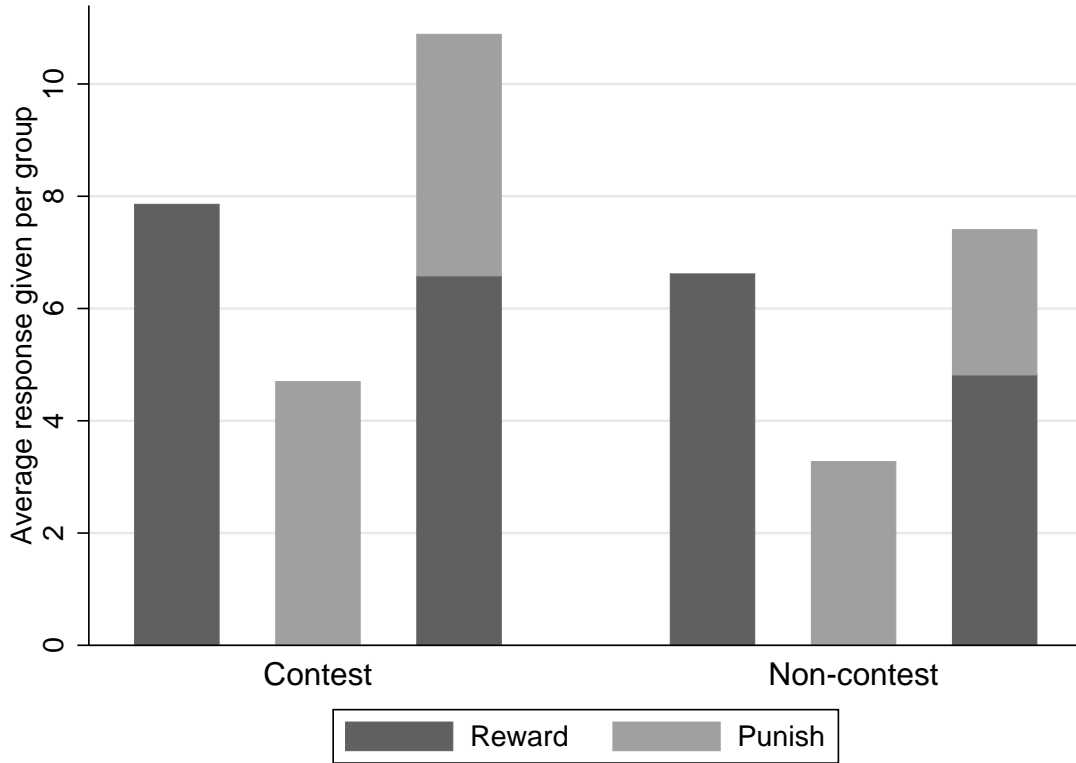


Figure 4: Response given per treatment

R&P and the punishment treatment in each environment (around 26 and 19 %, respectively), more players engage in rewarding behaviour in the reward treatment, as compared to the R&P treatment (56 % in reward, 44 % and 25 % respectively in R&P). Unlike in the R&P treatment, players in the reward treatment can only reciprocate negatively by not sending rewarding tokens. The spectrum of reciprocal actions shifts accordingly. In the punishment treatment this shift does not happen.

Between the environments, notice the close resemblance of percentages in the reward treatment. Notice also an increase in response giving behaviour for the contest environment, except for the reward treatment.

Punish				Reward		Punish	
Reward	No	Yes	Total	No	Yes	No	Yes
No	45.63	10.73	56.35	43.75	56.25	73.02	26.98
Yes	29.17	14.48	43.65	56.25	43.75	26.98	73.02
Total	74.79	25.21	100	100	100	100	100

R&P Reward Punish

Table 1: Share of response cases, contest environment (in percentages)

Punish				Reward		Punish	
Reward	No	Yes	Total	Reward		Punish	
No	62.71	11.46	74.17	No	43.13	No	81.25
Yes	17.5	8.33	25.83	Yes	56.88	Yes	18.75
Total	80.21	19.79	100	Total	100	Total	100

R&P
Reward
Punish

Table 2: Share of response cases, non-contest environment (in percentages)

4.4 Rent dissipation

Deviations from equilibrium strategy (as devised in Section 3.1) of course have payoff-relevant consequences. As mentioned before, mean income was €16.94 taken across all treatments. This translates to an amount of 2,540 tokens per player. If all players exhibited equilibrium behaviour all throughout the experiment, everyone could have earned on expectation

$$E(\pi_{tot}^{Nash}) = \left(B + p_K \cdot z - \frac{\sum_{k \in K} v_k}{4} + F \right) \cdot 15$$

or

$$(100 + 0.5 \cdot 100 - 6.25 + 50) \cdot 15 = 2,906.25$$

tokens, or €19.38, respectively. So overall, players earn almost 15 % less than what they could earn in Nash equilibrium. The social optimal expected payoff with $v_i = 0 \forall i \in K \cup M$, in comparison, amounts to $E(\pi_{tot}^{soc}) = 3,000$ tokens or €20.00.

Losses in total monetary welfare result from higher contributions, as compared to more payoff-optimal strategies.²⁹ This particularly emerges in the R&P treatment of the contest environment, where mean income was about €3.94 or 591 tokens (about 22 %) lower than in the other treatments. Also, as a result of higher spending and punishing, average income in the contest environment was roughly €4.00 lower than in the non-contest (about 601 tokens, 20%).

Table 3 gives an overview on the extent and composition of overspending per treatment. We compare the total sum of individual contest expenditures and response giving with the Nash equilibrium benchmark and report the respective spending level that exceeds this threshold. The total level of overspending in each treatment of the non-contest environment is lower as its counterpart in the contest environment. For the Baseline treatments this decline is even as high as 60 %.

4.5 Individual Level Analysis

In the following we analyse, which factors influence individual behaviour in this game. We run clustered Tobit regression, as discussed in Subsection 4.1, using treatment and group dummies. Additionally, we include data obtained from the questionnaire that each

²⁹As discussed earlier, contribution to the group account is unproductive effort which has no positive influence on the amount of tokens that can be won, as the prize at stake is fixed. Tokens put in the group account only influence the winning probability.

Treatment		Overspending contribution	Overspending response	Total
Contest	R&P	725.73	163.27	889.00
	Reward	448.98	117.84	566.83
	Punish	362.19	70.48	432.67
	Baseline	441.91	–.–	441.91
Non-contest	R&P	432.09	111.06	543.16
	Reward	431.03	99.25	530.28
	Punish	245.34	49.13	294.47
	Baseline	155.46	–.–	155.46

Table 3: Individual overspending compared to the Nash equilibrium benchmark

participant filled in after the experiment. In appendix X we discuss the additional control factors in an explorative analysis. As in Subsection 4.3, the baseline treatment was excluded from this analysis for conceptual reasons, as players in the baseline treatment cannot give a response. Regressions (1) and (2) show individual contribution, averaged over all 15 periods, on a number of factors. In (3) and (4) we regress own response on a similar set of factors.

The contest environment factor – a binary variable for the contest / non-contest environment – has significant positive explanatory power for individual contribution, which is in line with results from Subsection 4.2: Subjects tend to spend more resources to secure the group prize if they compete against another group. However, its negative coefficient in regression (4) comes somewhat unintuitive with regard to results in Subsection 4.3.

As for the factors “Contribute” and “Own response” it can be seen that players who give more response, tend to also be those who contribute more and vice versa. This should not come as too much of a surprise, considering the argument that players who display more involvement in the project, both contribute more and give more response in order to induce free-riding group members to contribute as well.

In a nutshell, results for “Group contribution level” indicate that players in a competitive group tend to spend more themselves on the contest. At the same time, response giving is slightly reduced. The term “Group response level”, on the other hand, is considerably more erratic and differs remarkably between the regressions with and without controls. We will have a closer look at the effect of response on individual behaviour in Subsection 4.6.

Also the actions of the other group have a somewhat erratic effect on own behaviour. “Other group contribute” are the sum of contributions of the opposing group, averaged over the 15 periods. Note that although being of minuscule size (about $1.5 \cdot 10^{-4}$), it’s very small standard error (about $2.0 \cdot 10^{-5}$) results in a pretty robust confidence interval.

VARIABLES	(1) Contribute	(2) Contribute	(3) Own response	(4) Own response
Contest environment	0.201*** (0.03)	29.735*** (6.90)	0.685*** (0.13)	-22.023*** (6.14)
Contribute			0.288*** (0.05)	0.333*** (0.05)
Own response	0.686*** (0.15)	0.837*** (0.14)		
Group contribution level	0.252*** (0.00)	0.181*** (0.02)	-0.072*** (0.01)	-0.041** (0.01)
Group response level	-0.158*** (0.04)	0.342** (0.12)	0.250*** (0.00)	-0.162 (0.09)
Other group contribute	0.000*** (0.00)	-0.135*** (0.04)	-0.004*** (0.00)	0.098*** (0.03)
Constant	-0.532*** (0.07)	40.682** (13.48)	0.092*** (0.02)	-13.879 (11.12)
Controls	<i>No</i>	<i>Yes</i>	<i>No</i>	<i>Yes</i>
Pseudo R-squared	0.192	0.225	0.120	0.160
N	288	270	288	270

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parentheses. Treatment and group dummies not reported.

Table 4: Individual level analysis

4.6 Dynamics in Decision Making

For this part, we closely follow the analysis technique of [Ashley, Ball, and Eckel \(2010\)](#). We use both OLS regression with Newey-West standard errors with the contribution to the group account as dependent variable.³⁰ The explanatory variables are the contribution to the group account lagged for one and two periods, individual positive or negative deviation from the mean of own group's contribution level lagged one period and response received one period ago.³¹ We analyse the effect of receiving reward or punishment, respectively, on subsequent contribution levels. Results are presented in Table 5.

This analysis delivers three main insights: *First of all*, again there is significant autocorrelation for contribution levels. *Second*, the lagged relative comparison of own contribution to the mean group contribution reveals that students contributing relatively more to the project reduce their spending for the group account in the subsequent period, whereas the opposite does not happen. Players who contribute less than their group mean do not increase their spending, ceteris paribus. *Third*, in line with [Sefton et al. \(2007\)](#), receiving rewards induces an increase in contribution for the following round, while sanctions do not have this effect. [Sefton et al. \(2007\)](#) argue that this explains the negative trend for punishment-giving over the periods. Furthermore they say that reward-giving also declines in the course of the experiment, but at a faster pace than punishment-giving. This is the case, although reward induces higher changes in other players' behaviour and shows efficiency-enhancing features. It might seem odd that participants do not stick to a strategy

³⁰We also use multivariate Tobit regression, obtaining substantially similar results.

³¹We also include a period factor, Netherlands & Belgium dummy and treatment, environment, group and study major dummies as control variables.

that seem to be effective in inducing other players' contribution. [Sefton et al. \(2007\)](#) do not elaborate further on this point.

4.7 Who receives Response?

In *The Theory of Moral Sentiments*, Adam Smith suggests: "Actions of a beneficent tendency, which proceed from proper motives, seem alone to require reward; because such alone are the approved objects of gratitude, or excite the sympathetic gratitude of the spectator. Actions of a hurtful tendency, which proceed from improper motives, seem alone to deserve punishment; because such alone are the approved objects of resentment, or excite the sympathetic resentment of the spectator" [Smith, Raphael, and Macfie \(1976, Part II, Section II, Chapter I, Paragraph 1-2\)](#). This early characterisation of reciprocity nicely captures our hypothesis concerning who will be punished or rewarded in this game.

For the contest environment, simple OLS regressions with Newey-West standard errors indicate that in the punishment treatment, players who contribute less are more likely to receive punishment from their teammates. The opposite holds for the reward treatment. Players, who contribute more, are more likely to receive reward from their teammates (see [Tables 6 and 7](#)). This relationship is much less pronounced in the non-contest environment. While lower contributors receive more punishment in the punish treatment, there is no significant effect for other treatments of the non-contest environment.

[Figure 5](#) displays response received as a function of individual's deviation from the mean of contributions per group in the contest environment. These deviations are grouped into intervals, which illustrates the gradient nature of response-giving. The higher the own deviation from the mean towards the positive region, the more reward and less punishment does a player receive. We observe the contrary for punishment. The higher the deviations from the group's mean contribution towards the negative array, the more punishment and less reward does a player receive. We observe a small amount of asocial punishment, as well as some reward for players that contribute less than the mean of the group contribution level.

[Figure 5](#) suggests that the response methods are well in place, such that they are used to punish defectors and reward high contributors. Furthermore, the gradient of response-giving seems to be rather smooth. There seems to be an exponential relation between contributing and receiving response, such that players who contribute much more than the mean of the group contribution, receive disproportionate reward. The same holds for

VARIABLES	(5) Newey- West Contribute
Contribution to group account in previous period	0.460*** (0.04)
Contribution to group account lagged two periods	0.252*** (0.03)
Contribution exceeding group mean previous period	-0.201** (0.06)
Contribution below group mean previous period	-0.080 (0.05)
Reward received previous period	0.130* (0.06)
Punishment received previous period	0.049 (0.04)

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parentheses. Period effect, NLB, treatment, environment, group and study major dummies not reported.

Table 5: Dynamic analysis

	(6)	(7)	(8)	(9)
	Punish	Reward	Punish R&P	Reward R&P
VARIABLES	response received			
Contribute	-0.624*** (0.08)	0.239*** (0.04)	-0.265** (0.09)	-0.016 (0.05)
Squared Contribute	0.005*** (0.00)	0.000 (0.00)	0.00128 (0.00074)	0.000907* (0.00047)

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parentheses. Period effect, NLB dummy, treatment, group and study major dummies not reported.

Table 6: Contest environment.

	(10)	(11)	(12)	(13)
	Punish	Reward	Punish R&P	Reward R&P
VARIABLES	response received			
Contribute	-0.697*** (0.16)	0.094 (0.09)	-0.219 (0.12)	0.069 (0.08)
Squared Contribute	0.005*** (0.00)	0.001 (0.00)	-0.000 (0.00)	0.000 (0.00)

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parentheses. Period effect, NLB dummy, treatment, group and study major dummies not reported.

Table 7: Non-contest environment.

the punishment treatment. Players can expect to get disproportionately punished if they contribute much less than their fellow group members.

We observe a similar, gradual composition of response received over deviations from average group contributions in the non-contest environment in Figure 6. Although the data is considerably more noisy than for the contest environment, a definite trend can be observed towards higher punishment of group-mates whose contribution deviates negatively from the group average, as compared to the contest environment. At the same time, this is not true for the positive domain: positive contributions are not rewarded more than in the contest environment.

Alternatively, subjects might actually compare others' contribution levels to their own ones, when making their decision of whom to reward or to punish. Hence, it must not necessarily be the case that the mean of group contribution delivers the relevant benchmark. To capture this, we use a dyadic relation, instead of an average group contribution level. This means that clusters for relative contribution are not formed by comparing own contribution level to the group mean, but by comparing it to the contribution of the "response giver". Hence, the deviation of own contribution to the contribution of those who punish or reward. This relative contribution, $dev_l^k = v_l - v_k$, gives a term of relative contribution for each of

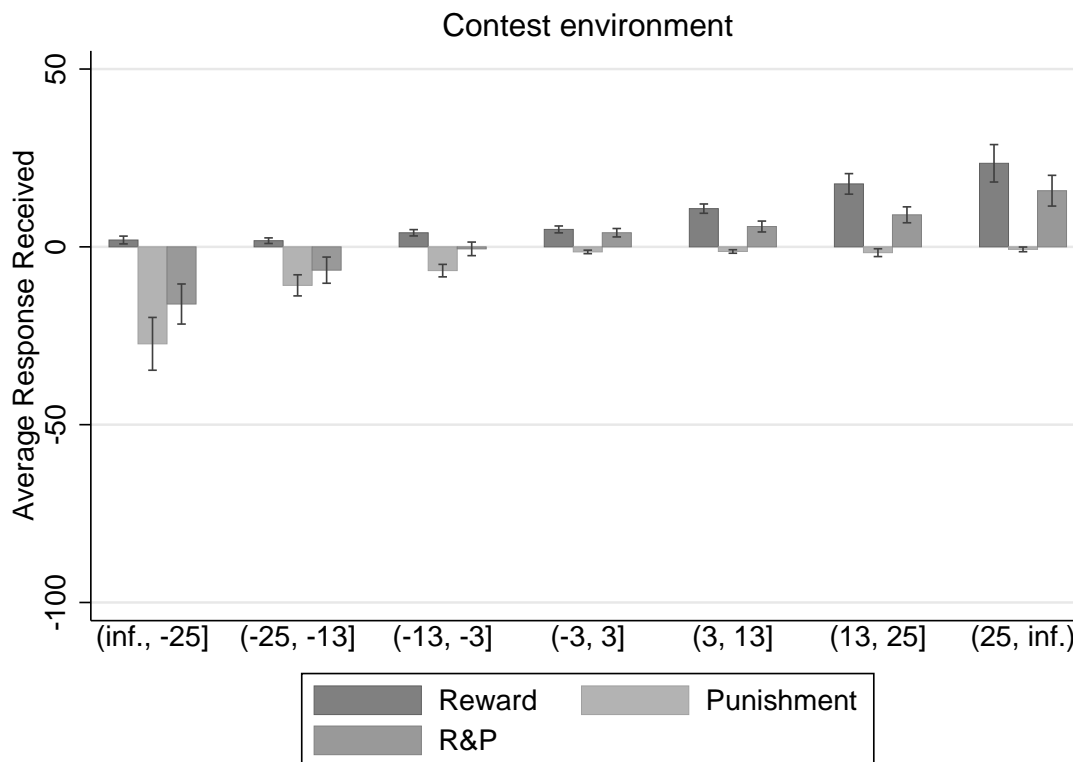


Figure 5: Response received in relation to deviation from average group contribution with 5% confidence interval.

the three co-players in the group for a given period t . dev_l^k measures the difference between contribution between player l and player k . Consequently, we get three dev_l^k -terms per player.

Results depicted in Figure 18 and 19 (in Appendix F) stay qualitatively similar. They show, however, that reward and punishment mechanisms are used in a redistributive way. Players tend to recompense team mates who contribute more to the contest by means of reward and reduce the payoff of lower contributors (those who earn more in the contest) by punishing. Comparing Figure 18 and 19 also suggests that low contributors are punished more heavily in the non-contest environment.

5 Concluding Comments

In this article we complement related work by [Sefton et al. \(2007\)](#), [Abbink et al. \(2010\)](#) and [Fehr and Gächter \(2000\)](#). The common conceptual starting point of this type of experimental research lies on the focus of reciprocity as a behavioural norm. In an experimental study, we investigated a team contest for a group prize. Furthermore, we varied the economic environment between treatments, such that the effect of competing with another group can be isolated. In this study we answer four main hypotheses, which are: 1) Contributions to the team contest are higher in the contest environment. 2) More rewarding in the contest environment. 3) Less punishing in the contest environment. 4) More total

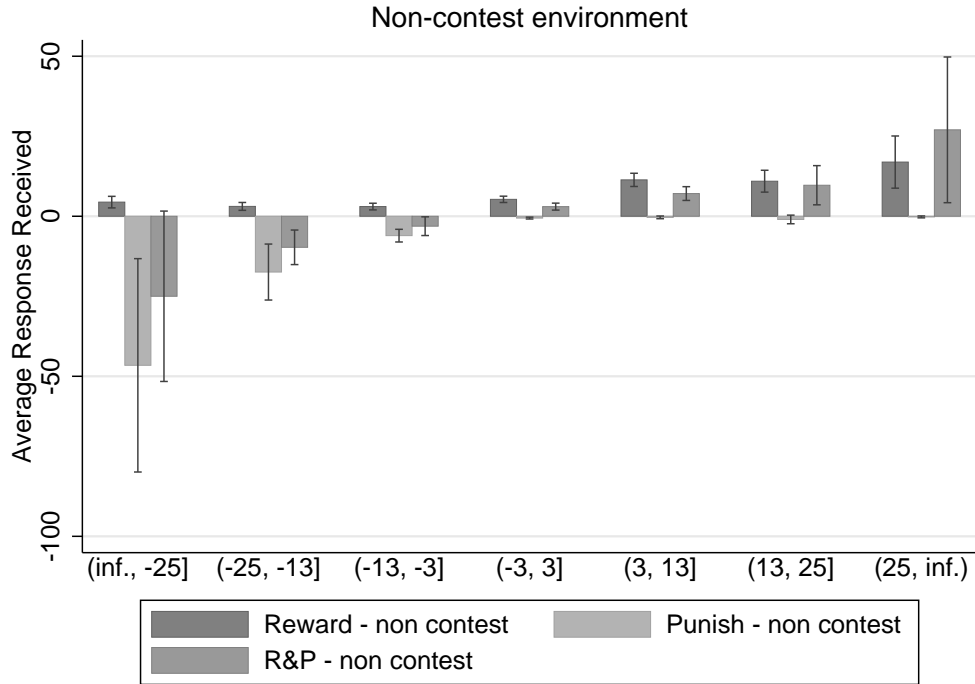


Figure 6: Response received in relation to deviation from average group contribution with 5% confidence interval.

response in the contest environment.

Our results give definite support for hypotheses 1), 2) and 4). Results for Hypothesis 3), however, are considerably less clear-cut. While it seems that low contributors are more heavily punished, overall propensity to punish and average total punishing are not higher in the non-contest environment.

Furthermore, we find that across both environments, contribution to the group project is higher in the reward treatment than in the punishment treatment. Especially during the first rounds of the game, participants distribute more rewarding than punishing response to their group mates. We also find that for response-giving, the own contribution in relation to the contribution of other group members is an important determinant of the severity of punishment, or reward, respectively. This means that in the punishment treatment, the further a player’s contribution to the group account is below the mean of the group’s contribution level, the higher punishment this player can expect to receive from fellow group members. The opposite holds for the reward treatment. Here, the higher the relative contribution, the higher the expected reward from other group mates. Also, we find two dynamic patterns of behaviour: Players who contribute more than their group mates reduce contribution in subsequent periods, while the opposite does not hold. Players who contribute comparably less do not increase their contribution significantly. Most strikingly, however, players react on being rewarded by increasing their contribution, while players who get punished do not change their contribution level.

The results of our experiment paint a gloomy picture of within-group punishment and reward systems in team contests. When confronted with a rival group, subjects show a

clear tendency to exhibit a more competitive behaviour. The opportunity to reciprocate teammates' actions (by the means of reward or punishment), is utilised to fan the flames of the intergroup conflict, even though this comes at a price. Our results contribute to explaining the high humanitarian and material costs in socio-political conflicts. When confronted with a rival party, subjects seem willing to accept a materially inefficient outcome to outrun the opponent. All the more so, subjects incite their comrades to join in arms. In the anonymous and abstract environment of this computer laboratory, we find these outcomes absent of any religious or ethnic spur. In field environments, the emotional impetus attached to these kind of conflicts will be even stronger, given a usually higher social identity within the conflict parties.

In a corporate context, consider a construction firm running for a tender. Although each unit has some incentive to not invest resources in the project of the firm, high inputs increase the probability for the firm to win the tender and to get the project awarded. Our research suggests that in this kind of situation, the company should not only point out to their employees that they are in a competitive situation in order to encourage higher effort. The implementation of a rewarding scheme should particularly trigger up individual effort in the contest. Results by [Balliet et al. \(2011\)](#); [O'Gorman et al. \(2009\)](#) suggest that it should not matter if there exists a scheme of mutual rewarding, similar to the setup in our experiment, or whether one central agent is entrusted with this task.

There are a number of conceivable extensions: Allowing for communication could formalise a sort of non-binding contract between the participants. On the one hand, subjects might form a tighter emotional bond if they can communicate with each other. On the other hand, however, communication could be perceived as a weaker response towards other group members, as it has no payoff consequence and is just cheap talk. Our results further suggest that subjects incentivise group mates to increase contribution even beyond an efficient spending level. It is not clear, if communication would lead to a further appeasement or escalation.

By randomising identification numbers of players, reputation building would be excluded. Hypothetically, this should lead to a decrease in response-giving and also to a reduction in contribution to the group account. Going one step further, groups could even be set up in a randomized way. Group cohesion is expected to be even smaller than in the previous case, resulting in less response-giving and lower contribution levels.

In our design of the contest game, the prize for winning a round is set such that the Nash equilibrium for contribution is on a rather low level of players' endowment with tokens. Throughout the game we constantly observe a massive over-investment across treatments. By varying the prize at stake, which changes the game's Nash equilibrium, can elicit subjects' elasticity towards the contest prize. This might also have an effect on response-giving, making subjects react more aggressively towards free-riders, with an increase in rewarding and punishing.

Appendix A Instructions

When entering the cubicle, each participant found a printed version of the following experiment instructions at her seat. Paragraphs headed by a treatment name in square brackets

were only given to the subjects of the corresponding treatment. The following example is from the contest treatments. Instructions for the non-contest treatments are slightly adjusted such as to not mention another group and avoid referring to teams or team mates.³²

CG Experiment

Welcome and thank you for participating in this experiment. Please read these instructions carefully. If you have any questions, please raise your hand and one of the hosts of the experiment will come to your cubicle to answer your question. Talking or using mobile phones or any other electronic devices is strictly prohibited. Mobile phones and other electronic devices should be left in the waiting room or switched off. If you are found violating these rules, you will both forfeit any earnings from this experiment, and may be excluded from future experiments as well.

This is an experiment about decision making. The instructions are simple and if you follow them carefully you might earn a considerable amount of money which will be paid to you privately and in cash at the end of today's session. The amount of money you earn depends on your decisions, on other participants' decisions and on random events. You will never be asked to reveal your identity to anyone during the course of the experiment. Your name will never be associated with any of your decisions. In order to keep your decisions private, do not reveal your choices to any other participant.

During the experiment you will have the chance to earn points, which will be converted into cash at the end of today's session, using an exchange rate of

$$15 \text{ points} = 0,1 \text{ €}.$$

Thus, the more points you earn, the more cash you will receive at the end of the session.

You and three other participants are assigned to a team. Four other participants are assigned to another team. All participants will remain in their teams for the entire experiment. None of you will learn who the own team members or the other team members are.

The experiment will consist of 15 periods, and in each period your team and the other team are competing for a prize in the following way:

At the beginning of each period you will receive 100 points. Then you can use these points to buy lottery tickets for your team. Any point you invest gives one lottery ticket for the team. Any point you do not invest in lottery tickets will be accumulated in your private point balance. Likewise, your team members can buy tickets for your team and the members of the other team can buy tickets for their team in exactly the same way.

As soon as everybody has chosen how many tickets to buy, a lottery will determine whether your team or the other team wins a prize of 400 points (100 for each team member). One of the sold tickets is randomly assigned the winning ticket. Each ticket has the same chance. Hence, the more tickets your team buys, the higher is your chance of winning the prize.

³²Details available upon request.

Examples: If your team and the other team buy the same amount of tickets then the chance of winning the prize is 50:50. If your team buys three times as many tickets as the other team, then also your team's chances are three times as high as those of the other team. If only one of the teams buys tickets then this team wins the prize with certainty.

If neither your team nor the other team buys a ticket, then the prize is randomly allocated to one of the teams with equal chances.

After the winning team is determined the prize of 400 points is equally shared and added to the private point balances of the winning team (100 points for each team member).

[Reward treatment:] At the end you get 50 feedback points. You can put these points to your private point balance or you can assign them to one or more of your team members (**not** to members of the other team). Each feedback point assigned to a team member increases this team member's private balance by one point.

[Punishment treatment:] At the end you get 50 feedback points. You can put them to your private point balance or you can assign the feedback points to one or more of your team members (**not** to members of the other team). Each feedback point assigned to a team member decreases this team member's private balance by one point. A team member's balance for one particular round cannot turn negative. If this happens then the balance for this particular round is set to zero and all excess feedback points are invalidated.

[Baseline treatment:] At the end you get 50 extra points which will add to your account.

[R&P treatment:] At the end you get 50 feedback points. You can put them to your private point balance or you can assign the feedback points to one or more of your team members (**not** to members of the other team). You can decide whether you assign the points to be added to the respective team member's account or to be subtracted from that account. Each feedback point assigned to a team member increases or decreases this team member's private balance by one point – depending on what you choose. A team member's balance for one particular round cannot turn negative. If this happens then the balance for this particular round is set to zero and all excess feedback points are invalidated.

The points you earn in each period will be added together. At the end of the session you will be paid based on your total point earning from all 15 periods.

The experiment starts with a trial period in which you will be asked to fill in some questions in order to check your understanding of the experiment and to give you the opportunity to get acquainted with the setup. Points earned in this trial period will not be paid off.

Appendix B Stages

Figure 7 to Figure 9 show screenshots from the punishment treatment. Other treatments look similar, adjusting only references to the payoff consequences of receiving response and leaving out mentioning of another group in the treatments without competing group.

This is period #: 1.
In this stage you decide, how many Lottery tickets you want to buy.

	Initial Endowment	Lottery tickets bought	Income from the Lottery	Earnings in this stage
You	100	--	--	--
Member 1	100	--	--	--
Member 2	100	--	--	--
Member 3	100	--	--	--

How many Lottery tickets do you want to buy for your team?

Figure 7: First stage

This is period #: 1.
Now you see if your team has won the price and how many Lottery tickets each of your team members bought.

	Initial Endowment	Lottery tickets bought	Income from the Lottery	Earnings in this stage
You	100	6	0.0	94.0
Member 1	100	6	0.0	94.0
Member 2	100	6	0.0	94.0
Member 3	100	50	0.0	50.0

Your team has lost!

The other team bought in total 56 Lottery tickets.
Your winning probability was 55%.

How many feedback points do you want to assign to the other members of your team?
You can assign up to 50 feedback points:

Member 1: Member 2: Member 3:

Figure 8: Second stage

<p>In total, other team members have allocated to you 11 deduction points.</p> <p>This decreases your earnings by 11 tokens.</p>	
<p>Your earnings for Period 1 are:</p>	
Your initial endowment:	100
- Lottery tickets bought by you:	-6
+ prize won:	0
= your first stage earnings:	94
+ feedback points:	50
- 1 point for every feedback point you assigned to others:	-2
- 1 points for every feedback point assigned to you by others:	-11
= your earnings for this period:	131.0
Total earnings up to this period:	131.0
<div style="text-align: right; margin-top: 10px;"> <input type="button" value="Ready"/> </div>	

Figure 9: Third stage

Appendix C Mathematical Appendix

Player l maximises expected profit π_l by setting own contribution $v_l = (v_i)_{i \in K \cup M}$. Investment of players $k \in K$ of own group K is labelled $\sum_{k \in K} v_k$, while players $m \in M$ in the other disjoint group M invest $\sum_{m \in M} v_m$. Individual prize for winning a round is z . Solve individual optimisation problem for any period t , time indices are omitted.

$$\pi_l((v_i)_{i \in K \cup M}) = \frac{v_l + \sum_{k \in K \setminus \{l\}} v_k}{v_l + \sum_{k \in K \setminus \{l\}} v_k + \sum_{m \in M} v_m} \cdot z - v_l$$

Deriving delivers the best response function for any player l of group K :

$$\frac{\partial \pi_l((v_i)_{i \in K \cup M})}{\partial v_l} = 0 \quad \Leftrightarrow \quad v_l = \sqrt{\sum_{m \in M} v_m \cdot z} - \sum_{m \in M} v_m - \sum_{k \in K \setminus \{l\}} v_k$$

Checking the second order condition confirms that we find a maximum:

$$-\frac{2 \sum_{m \in M} v_m \cdot z}{\left(\sum_{k \in K} v_k + \sum_{m \in M} v_m \right)^3} < 0$$

Using the first order condition of group M , we find a multiplicity of equilibria, characterised by $\sum_{m \in M} v_m = \frac{z}{4}$ and $\sum_{k \in K} v_k = \frac{z}{4}$. If we assume symmetry in own group: $v_l = \frac{z}{16}$.

Appendix D Group wise Analysis of Contribution

Figures 10 – 17 depict average contribution per player across each of the 15 periods for each group. In Figures 10 – 13, odd-numbered group always plays against a group with a number that is one higher than the own one. The opposite is the case for even-numbered groups. This means that, for example, group 1 plays against group 2 or group 12 against group 11. As mentioned before, there has been no session of a group 3 & 4 in the baseline contest and group 6 in the baseline non-contest treatments, for why they are left blank.

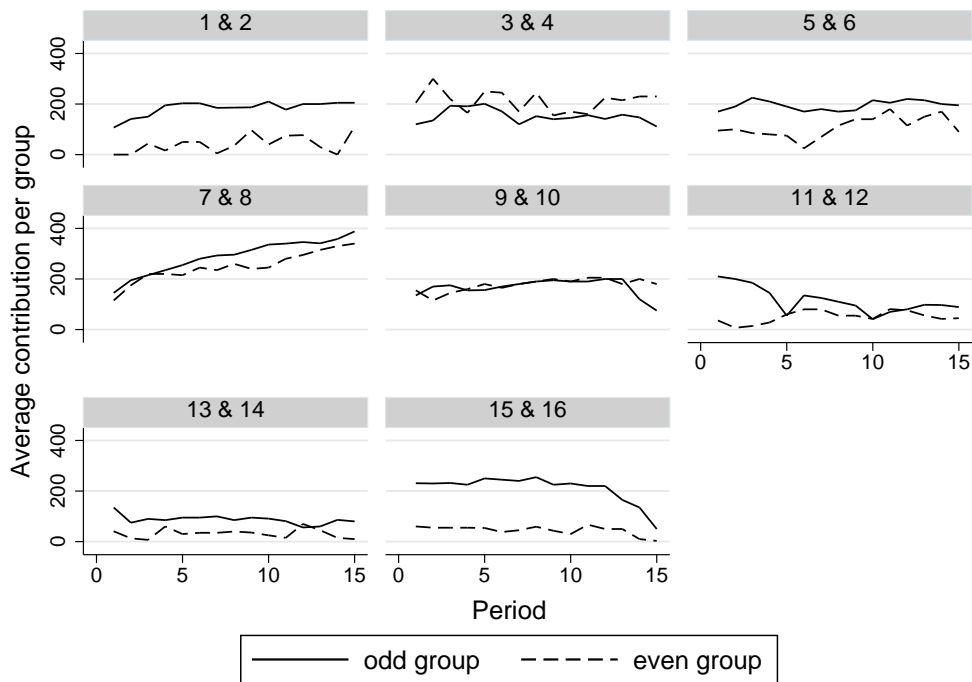


Figure 10: Average contribution per group, reward treatment, contest environment. Paired groups are displayed together.

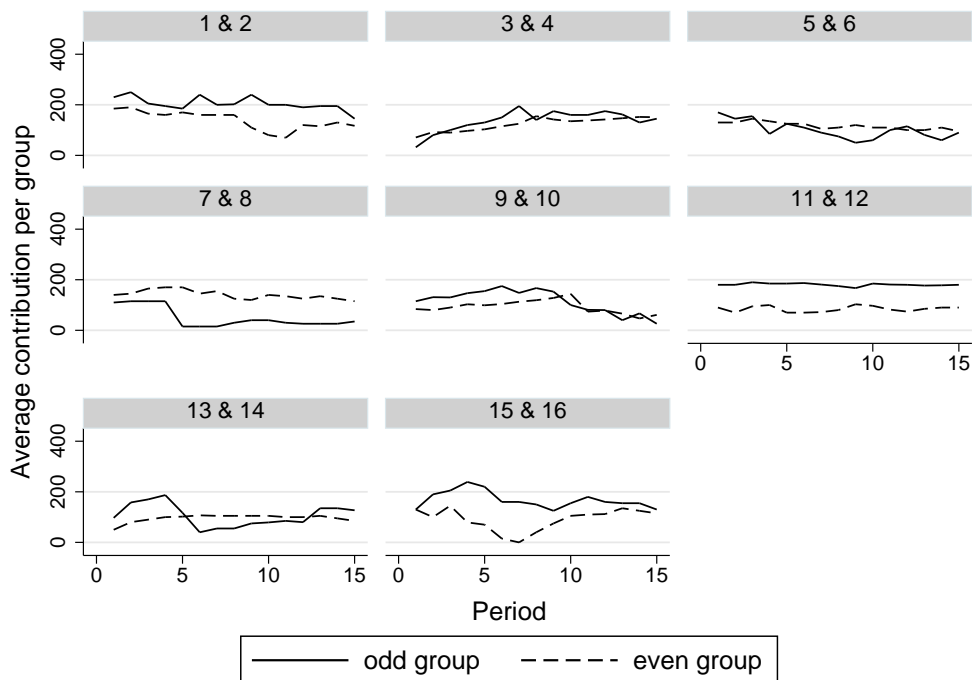


Figure 11: Average contribution per group, punishment treatment, contest environment. Paired groups are displayed together.

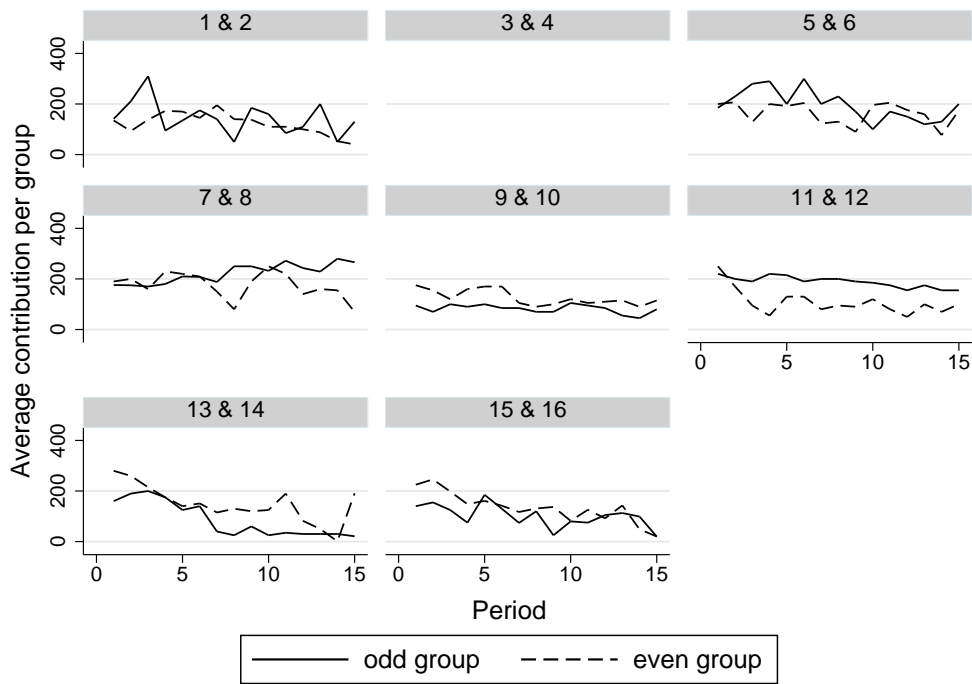


Figure 12: Average contribution per group, baseline treatment, contest environment. Paired groups are displayed together. Session for groups 3 & 4 did not take place due to no-shows.

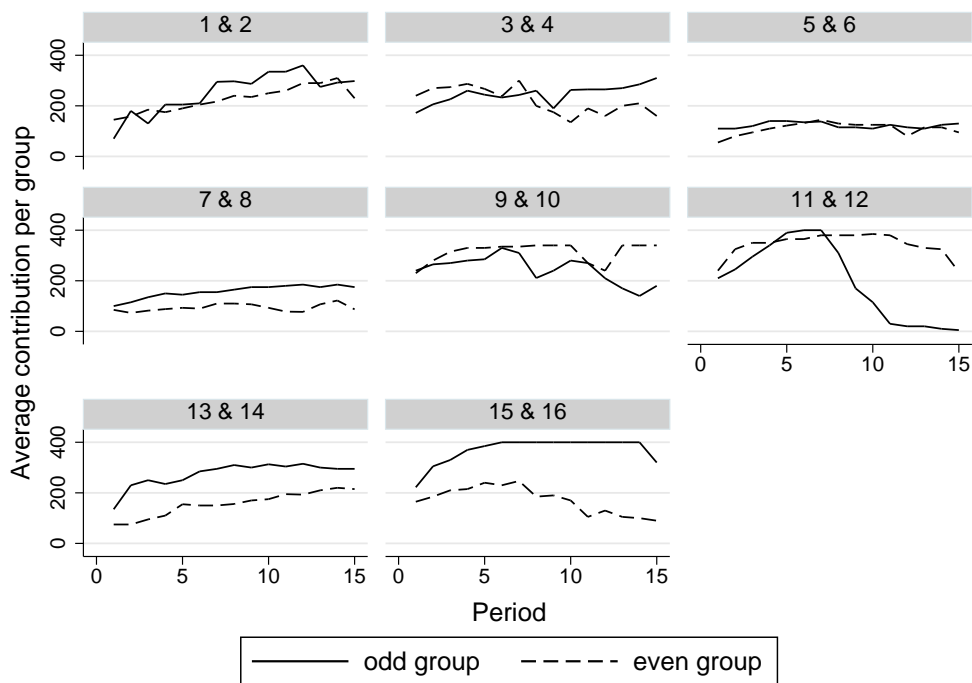


Figure 13: Average contribution per group, R&P treatment, contest environment. Paired groups are displayed together.

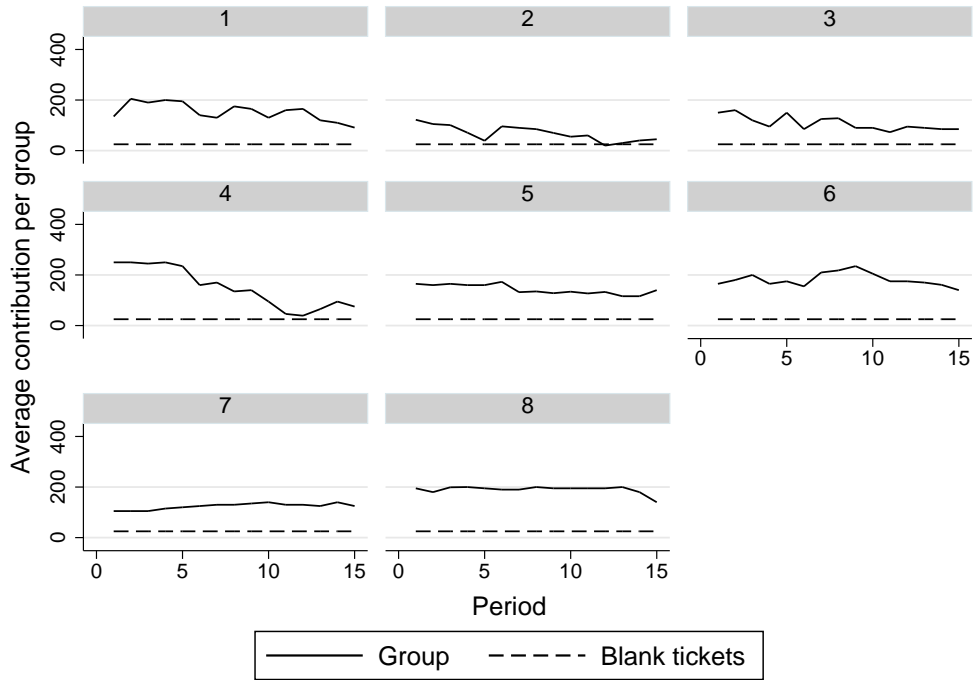


Figure 14: Average contribution per group, reward treatment, non-contest environment.

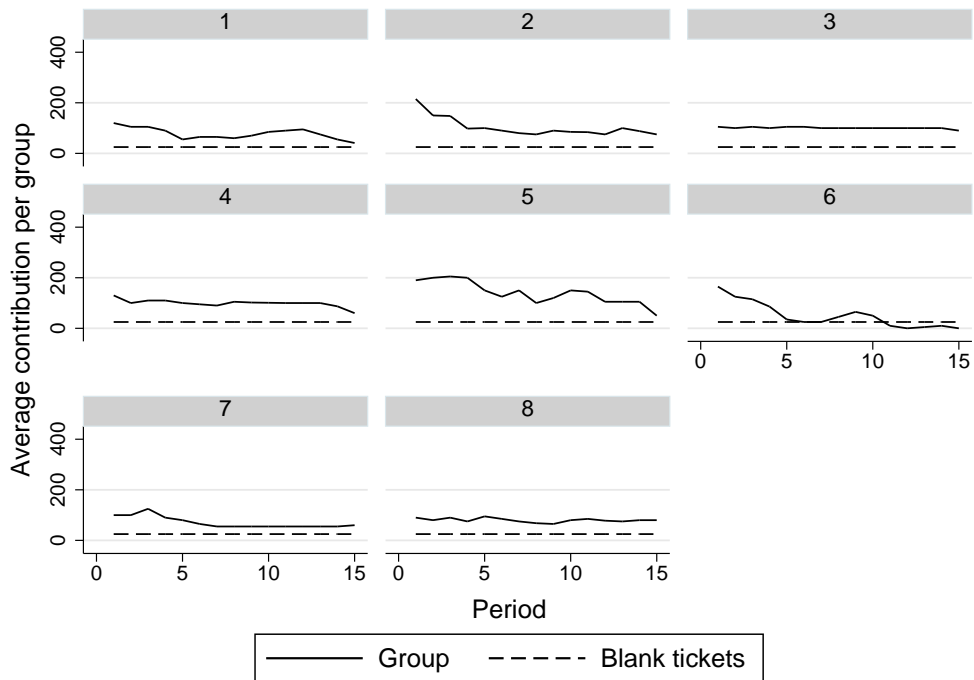


Figure 15: Average contribution per group, punishment treatment, non-contest environment.

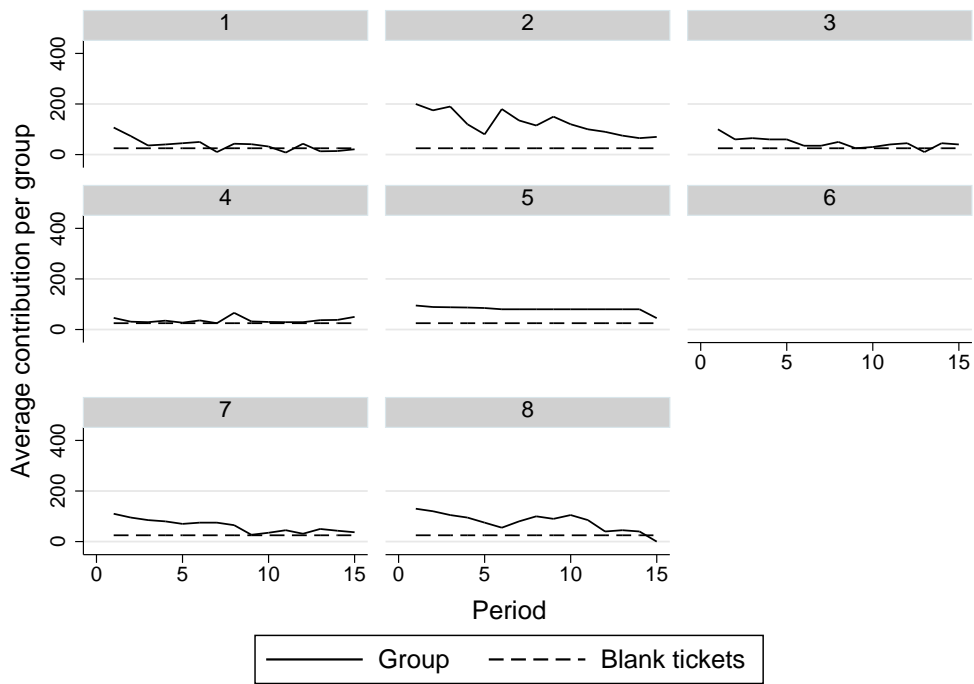


Figure 16: Average contribution per group, baseline treatment, non-contest environment. Session for group 6 did not take place due to no-shows.

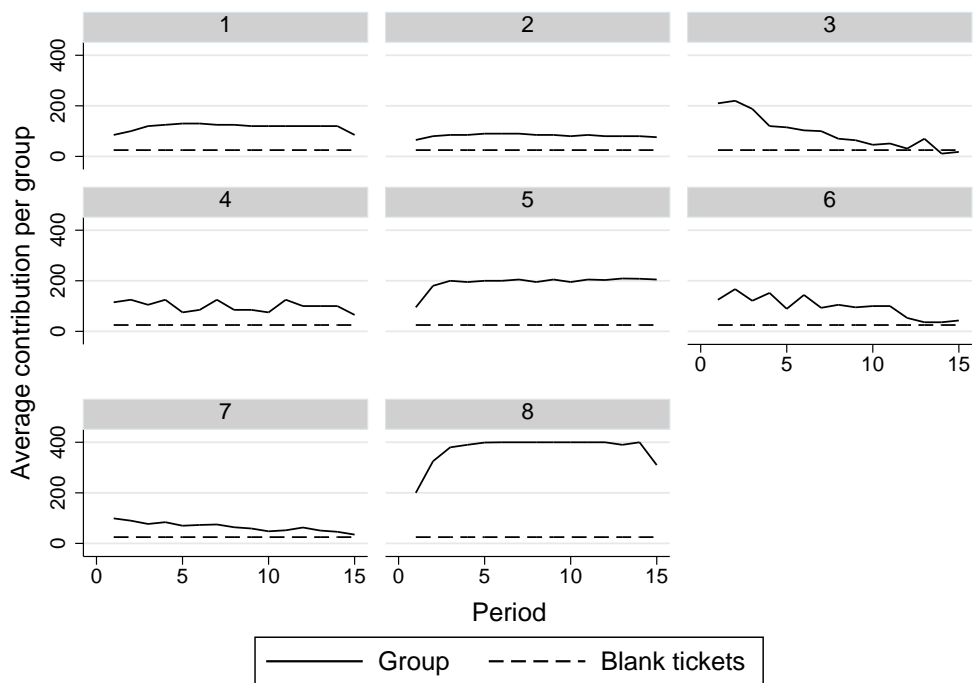


Figure 17: Average contribution per group, R&P treatment, non-contest environment.

Appendix E Personal attributes

Table 8 corresponds to regressions (2) and (4) from Table 4 with explicit control parameters. They were generated by subjects' answers in a questionnaire, in which we ask about a few personal features. In this section we examine the control factors in more detail.

Most notably, domestic students from the Netherlands and Belgium contribute substantially less to the contest than their colleagues from other countries do. This is the only significant country effect and its magnitude is somewhat impressive.

Players who state to place more importance on individual responsibility (as opposed to governmental responsibility) contribute less to the contest, but give more response. This factor was created on a scale from one to seven where individuals stated their proximity to which of the two statements they feel closer.³³ Players with a higher preference for individual responsibility hence tend to be reluctant towards spending their money for the group project, but are willing to reciprocate by the means of response. "Preference for working alone" was also created on a scale of one to seven with the statement that the individual prefers working in groups or alone at the extremes, to which she expresses proximity.³⁴ Players who prefer to work alone give less response, which means they act less reciprocal.

Further significant negative effects on contribution are smoking, "Family and friends important"³⁵ and age, while the Trust parameter³⁶ has a positive effect. For own response we observe a lowering effect of the study phase³⁷ and an increasing effect from subjects who find work important, smoking and age. Other factors showed no significant effect on subjects' behaviour.

³³The two extremes were "The government should take more responsibility to ensure that everyone is provided for." versus "People should take more responsibility to provide for themselves."

³⁴The two extremes were "I like working in groups" versus "I like working on my own"

³⁵Participants' stated preference on a scale from 1 – 7.

³⁶Proximity to the two statements: Most people can be trusted – Need to be very careful.

³⁷i.e. Bachelor, Master, Postgraduate.

VARIABLES	(14) Contribute	(15) Own response
Contest environment	29.735*** (6.90)	-22.023*** (6.14)
Contribute		0.333*** (0.05)
Own response	0.837*** (0.14)	
Group contribution level	0.181*** (0.02)	-0.041** (0.01)
Group response level	1.370** (0.48)	-0.649 (0.38)
Other group contribute	-0.135*** (0.04)	0.098*** (0.03)
Netherlands and Belgium	-7.473*** (1.76)	1.842 (1.61)
Preference for working alone	0.501 (0.49)	-0.967** (0.35)
Individual Responsibility	-1.394** (0.44)	0.771* (0.31)
Work important	0.308 (0.92)	1.524* (0.71)
Smoking	-4.573* (2.17)	4.369** (1.62)
Trust parameter	2.395* (1.19)	-0.498 (0.92)
Family and friends important	-2.872* (1.16)	-0.523 (0.99)
Politics important	-1.475 (0.85)	0.391 (0.45)
Current happiness	0.878 (1.08)	-0.759 (0.61)
Ever practised a team sport	-1.517 (1.15)	1.544 (1.11)
Siblings	1.286 (0.72)	-0.491 (0.41)
Study phase	3.194 (1.66)	-3.330** (1.17)
Age	-0.812** (0.30)	0.535* (0.24)
Constant	40.682** (13.48)	-13.879 (11.12)
Pseudo R-squared	0.225	0.160
Observations	270	270

* p<0.05, ** p<0.01, *** p<0.001

Standard errors in parentheses. Treatment, group and study major dummies not reported.

Table 8: Individual level analysis

Appendix F Response received, dyadic analysis

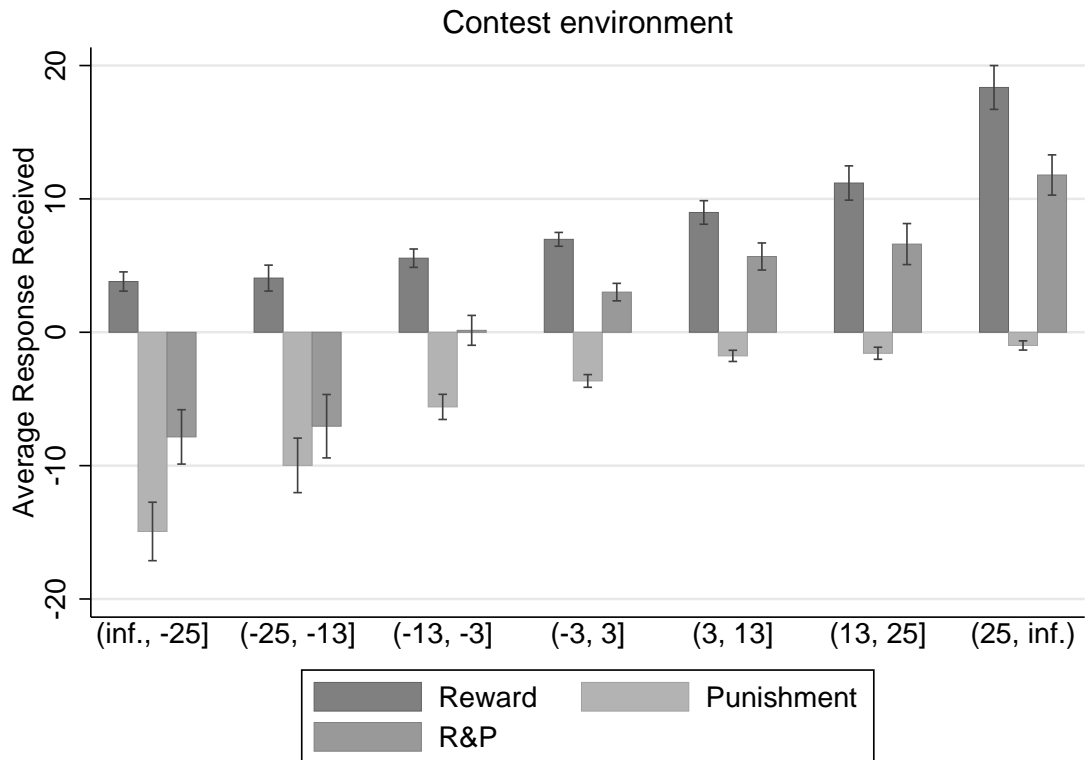


Figure 18: Response received in relation to deviation from sender's contribution with 5% confidence interval.

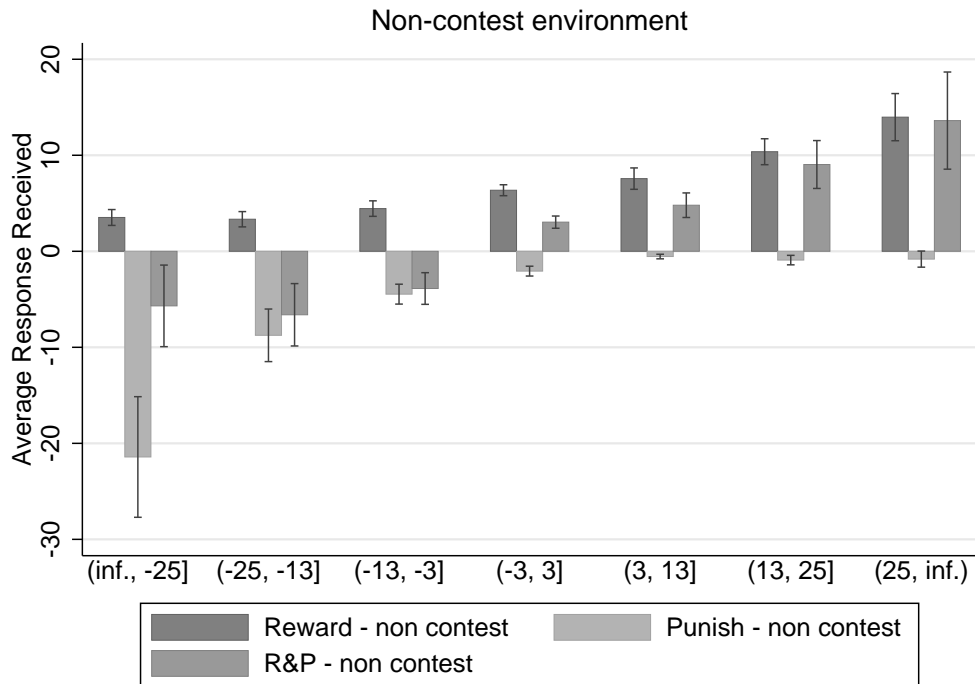


Figure 19: Response received in relation to deviation from sender's contribution with 5% confidence interval.

References

- K. Abbink, J. Brandts, B. Herrmann, and H. Orzen. Intergroup conflict and intra-group punishment in an experimental contest game. *American Economic Review*, 100(1):420–447, 2010.
- J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics an empiricist's companion*. Princeton University Press, Princeton, NJ [etc.], 2009.
- R. Ashley, S. Ball, and C. Eckel. Motives for giving: A reanalysis of two classic public goods experiments. *Southern Economic Journal*, 77(1):15–26, 2010.
- D. Balliet, L. B. Mulder, and P. A. M. Van Lange. Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, 137(4):594 – 615, 2011. ISSN 0033-2909. URL <http://search.ebscohost.com/login.aspx?direct=true&db=pdh&AN=2011-09704-001&site=ehost-live&scope=site>.
- G. Bornstein and M. Ben-Yossef. Cooperation in intergroup and single-group social dilemmas. *Journal of Experimental Social Psychology*, 30(1):52 – 67, 1994. ISSN 0022-1031. doi: <http://dx.doi.org/10.1006/jesp.1994.1003>. URL <http://www.sciencedirect.com/science/article/pii/S0022103184710031>.
- M. B. Brewer. The role of ethnocentrism in intergroup conflict. In W. G. Austin and S. Worchel, editors, *The social psychology of intergroup relations*, volume 2. Nelson-Hall Chicago, 1979.
- E. Dechenaux, D. Kovenock, and R. Sheremeta. A survey of experimental research on contests, all-pay auctions and tournaments. *All-Pay Auctions and Tournaments (September 28, 2012)*, 2012.
- W. Doise. *L'articulation psychosociologique et les relations entre groupes*. Ed. A. de Boeck, 1976.
- E. Fehr and S. Gächter. Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994, 2000.
- U. Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.
- S. Gächter and B. Herrmann. Reciprocity, culture and human cooperation: previous insights and a new cross-cultural experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1518):791–806, 2009. doi: 10.1098/rstb.2008.0275. URL <http://rstb.royalsocietypublishing.org/content/364/1518/791.abstract>.
- L. Goette, D. Huffman, S. Meier, and M. Sutter. Competition between organizational groups: Its impact on altruistic and antisocial motivations. *Management Science*, 58(5): 948–960, 2012.
- W. H. Greene. *Econometric analysis*. Pearson, Boston, Mass. [u.a.], 2012.
- B. Greiner. The online recruitment system orsee 2.0. *A guide for the organization of experiments in economics*. University of Cologne, 2004.

- E. Katz, S. Nitzan, and J. Rosenberg. Rent-seeking for pure public goods. *Public Choice*, 65(1):49–60, 1990. ISSN 0048-5829. doi: 10.1007/BF00139290. URL <http://dx.doi.org/10.1007/BF00139290>.
- K. A. Konrad. Strategy in contests: an introduction. Technical report, Social Science Research Center Berlin (WZB), 2007. Discussion Papers, Research Unit: Market Processes and Governance.
- K. A. Konrad. *Strategy and dynamics in contests*. Oxford University Press, 2009.
- K. A. Konrad and F. Morath. Evolutionarily stable in-group favoritism and out-group spite in intergroup conflict. *Journal of Theoretical Biology*, 306(0):61–67, 2012.
- H. B. Mann and D. R. Whitney. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1):50–60, 1947.
- W. K. Newey and K. D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.
- R. O’Gorman, J. Henrich, and M. Van Vugt. Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 276(1655):323–329, 2009.
- A. Öncüler and R. Croson. Rent-seeking for a risky rent: A model and experimental investigation. *Journal of Theoretical Politics*, 17(4):403–429, 2005. doi: 10.1177/0951629805056895. URL <http://jtp.sagepub.com/content/17/4/403.abstract>.
- T. F. Pettigrew. The ultimate attribution error: Extending allport’s cognitive analysis of prejudice. *Personality and Social Psychology Bulletin*, 5(4):461–476, 1979.
- D. G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, and M. A. Nowak. Positive interactions promote public cooperation. *Science*, 325(5945):1272–1275, 2009.
- T. Sasaki, R. Brännström, U. Dieckmann, and K. Sigmund. The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences*, 109(4):1165–1169, 2012.
- M. Sefton, R. Shupp, and J. M. Walker. The effect of rewards and sanctions in provision of public goods. *Economic Inquiry*, 45(4):671–690, 2007.
- M. Sherif. *Intergroup conflict and cooperation the Robbers cave experiment*, volume Rev. ed. University Book Exchange, Norman, OK, 1961.
- K. Sigmund, C. Hauert, and M. A. Nowak. Reward and punishment. *Proceedings of the National Academy of Sciences*, 98(19):10757–10762, 2001.
- A. Smith, D. D. Raphael, and A. L. Macfie. *The Theory of Moral Sentiments*, volume 1. Clarendon Press, Oxford [etc.], 1976. ISBN 9780198281894, 0198281897.
- W. G. Stephan. Intergroup relations. In L. Gardner and A. Elliot, editors, *Handbook of social psychology*. New York : Random House, 3rd ed edition, 1985.
- J. H. Stock and M. W. Watson. *Introduction to econometrics*, volume 3rd ed., global ed. of *The Pearson series in economics*. Pearson Education, Boston, MA [etc.], 2012.

- H. E. Tajfel. *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Academic Press, 1978.
- J. H. Tan and F. Bolle. Team competition and the public goods game. *Economics Letters*, 96(1):133 – 139, 2007. ISSN 0165-1765. doi: 10.1016/j.econlet.2006.12.031. URL <http://www.sciencedirect.com/science/article/pii/S0165176506004447>.
- G. Tullock. Efficient rent seeking. In J. Buchanan, R. Tollison, and G. Tullock, editors, *Toward a Theory of the Rent-Seeking Society*, pages 97 – 112. Texas A & M University Press, 1980.
- F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):pp. 80–83, 1945. ISSN 00994987. URL <http://www.jstor.org/stable/3001968>.