Titel der Arbeit:

Breaking Bad - Essays on the Effects of the Anti-Social Side of Human Behavior

Schriftliche Promotionsleistung

zur Erlangung des akademischen Grades

Doctor rerum politicarum

vorgelegt und angenommen

an der Fakultät für Wirtschaftswissenschaft

der Otto-von-Guericke-Universität Magdeburg

Verfasserin:	Alexandra Seidel, M.Sc.
Geburtsdatum und -ort:	28.02.1991, Chemnitz
Arbeit eingereicht am:	29.11.2024
Gutachter und Gutachterin der schriftlichen Promotionsleistung:	
	Prof. Dr. Joachim Weimann
	Prof. Dr. Jeannette Brosig-Koch
Datum der Disputation:	04.02.2025

Summary

In the TV-show *Breaking Bad*, the audience witnesses the transformation of Walter White, an initially moral man with a normal life, into an increasingly anti-social and destructive figure. This dramatic shift raises fundamental questions about the conditions under which individuals choose to harm others for personal gain. This dissertation explores these questions through four essays in experimental economics. These essays investigate how subjects behave in situations that allow them to exploit others' disadvantageous status or exclude them, ultimately causing harm to others for self-enrichment. This summary briefly introduces the projects and provides the extended abstract of each of the studies.

The first two studies deal with the possibility of exploiting another one's situation and thus refusing to behave fairly and honestly. In *"Fairness under Uncertainty"*, subjects repeatedly act with a partner and have equal chances to win a period. In certain periods, one of the partners loses her possibility to actively choose a strategy, allowing the other subjects to behave unfairly to maximize their payoff. In *"Laundering the People instead of the Money - An experimental study on the effect of mental money laundering"*, there is an information asymmetry allowing the informed player to exploit the lack of knowledge of the other player.

The remaining two studies deal with the exclusion of people. In *"The Cinderella Game – finding those who will not go to the ball"*, we elicit the norm that yields to a group's exclusion of certain personas via special characteristics. *"Choosing a Victim you know – introducing communication to the Mobbing Game"* allows group members to enrich themselves by excluding one of them.

"Fairness under Uncertainty" investigates both fairness and reciprocity in uncertain situations. By using an adapted repeated matching pennies game (20 rounds), we created a situation in which both players have equal chances to win a round and thus earning a point by choosing one of their available actions (normal rounds). However, there were also rounds in which one of the players was unable to actively choose one of her actions and thus the other player could exploit this disadvantage to his own benefit (special rounds). In case the player in charge decided to play fair, both received no payoff for this round but played an additional random round in the end. In case the player chose not to play fair, there were two possible outcomes: Either, the payoff was equal to the normal rounds (one player receives a point, one player receives nothing) or the player losing the respective round was additionally hurt by losing one of her points while the unfair player received one point, leading to an efficiency loss compared to normal rounds.

We varied the degree of uncertainty in three levels, from participants having no information about quantity or timing of the special rounds and the role of players to both players being fully informed about the special rounds.

The variation of uncertainty levels was conducted in a single-blind setup. However, since we know that being observed by others highly affect human behavior, we also varied the social distance between players and the experimenter adding a double-blind, and a "visible" setup. All these treatments were played in the setting with no information at all. In the visible setup, participants were shortly introduced to their partners without being allowed to talk to each other before the experiment started.

We find that full information does not lead to fair behavior but as long as acting selfishly does not affect efficiency, neutral moves are unlikely to occur. However, the share of selfish actions decreases drastically when participants face an efficiency loss. Also, the effect of the level of information is significant. We find that the more information subjects have, the more selfish actions they take. This indicates that subjects play fair to avoid punishment instead of revealing their preference for fair behavior. However, there are significantly less selfish actions in treatments with efficiency loss yielding to the conclusion the potential loss drives fair behavior. Possible pro-social preferences fail to explain this behavior because of the high share of selfish actions in the efficiency loss-games with full information. Further, social distance does not affect the results in these games as we find no differences in the share of selfish actions between the three levels of visibility. This is different in games with the same payoff structure like normal rounds. Using the single-blind situation as a reference point, we find significantly more neutral moves in both the double-blind and the visible setup with the visible setup having a stronger increase.

In total, we find that uncertainty about future situations is a strong driver for fair or positively reciprocal behavior. The common word that "you always meet twice in life" seems to apply here. This especially holds for situations in which unfair behavior harms the other in a way that leads to efficiency loss that could, in a future situation with switched roles, lead to a loss for oneself. The fear of revenge or punishment, however, does not seem to prevail in situations with an equal outcome as in the normal rounds. In these games, the rule that if everybody takes care of herself everybody will be taken care of prevails. We see that subjects exploit their partners disadvantage to enrich themselves.

In *"Laundering the People instead of the Money"*, subjects dealt with two common rules that influence decisions in real situations. The first one is not to lie. In the first part of the experiment, subjects are matched in pairs that face information asymmetry. One partner ("sender") has full

information about the payoff structure while the other's ("receiver") only source of information is a message she receives from the informed partner. The sender can then decide to tell the truth which yields a higher payoff for the receiver and a lower payoff for himself or to tell a lie yielding a higher payoff for himself and a lower payoff for the receiver resulting in an efficiency loss for the pair. The receiver can follow the advice leading to a payoff distribution.

After this first part, subjects are informed that they now have the possibility to donate a share of their gained wealth to a local charity. In the baseline treatment, this donation was made jointly by the partners of the first stage, and thus dishonest senders acted with those who were affected by their lie. In the switch treatment, the sender from stage one was matched with a new partner who had completed a different task in a different room and had no knowledge about the history of the other player ("dummy"). They then also could donate jointly, allowing dishonest senders to "launder" the unethically earned money by acting with a new partner ("launder the partner").

After their donation, all subjects were asked about their beliefs concerning the relative donation of their partner and answered questionnaires on personality traits like the Big Five and the Dark Triad (Narcissism, Psychopathy, and Machiavellianism).

The study finds that 75% of all senders lie and thus exploit the information asymmetry. However, dishonest senders behave contrary to the expectations of mental money laundering (Imas et al. (2021)) - to contribute a high donation in order to cleanse the unethically earned money. On the one hand, dishonest senders do not donate more than the honest ones in both treatments. This suggests a potential lack of guilt or a higher utility of earning more money than the disutility of lying. On the other hand, dishonest senders donate significantly *more* in the switch treatment than in the baseline treatment. A possible explanation is that dishonest senders fear to reveal their wealth and thus their lie if they donate a high amount. Senders know the receiver's payoff and the receiver knows that senders were fully informed about the payoff structure. Thus, senders could have decided to donate less to cover their lie.

Concerning a potential correlation between the expected relative donation of the partner and the own donation, the study finds that honest senders who expected their partners not to contribute to the joint donation, donated significantly less than those who expected their partners to donate. Understanding the donation as a public good, the reason for this behavior might be the fear of being exploited. Dishonest senders do not show a similar pattern.

The study finds that besides the decision setting, personal traits influence donation behavior. High scores in extraversion, neuroticism, and Machiavellianism showed a significantly negative effect on donation behavior underlining the power of characteristics when it comes to

ethical and prosocial decisions. In total, this study shows how participants harmed their partners to enrich themselves.

In contrast to the previous studies, *"The Cinderella Game"* focuses on exclusion. The experiment started with a survey. Participants were asked to answer questions about their appearance (e.g. hair color, body type, weight) and their characteristics (e.g. trust, free time activities, and answered Big Five and Dark Triad questionnaires. This information was used to create two types of fictional person cards – 16 cards with visual attributes and 16 cards with behavioral attributes. The second stage of the experiment was a norm-elicitation experiment. Subjects were asked to rank the person cards according to their beliefs of the likelihood of being excluded by the group with rank 1 being the card with the highest likelihood to be excluded. This ranking was incentivized since the closer the subject's individual ranking was to the group's ranking, the higher the subject's payoff. After subjects finished their ranking, they received a message showing their ranking and providing the possibility to adjust it. The additional text of the message varied depending on the treatment applied. In total, there were four treatments. The *neutral* treatment message purely informed about the possibility to re-rank. The *ethical* message told subjects that most people decide based on stereotypes. The *mimic* message informed subjects about two cards that were part of the top five of a previous experiment, examining whether subjects followed this anchor and changed their ranking accordingly, and the *attention* message informed subjects about two cards that were **not** part of the top five of a previous experiment to check for attention of subjects. All treatments tested the stability of the subject's exclusion preferences when being exposed to normative influences. Conducting a rank-based conjoint analysis, we found that having a high BMI and red or colored hair were the main visual drivers to be excluded by the group. Concerning behavioral aspects, a low trust score and high negative reciprocity significantly increased the likelihood of being excluded. These exclusion patterns remained stable across different demographic groups.
We also found that the information in the *mimic* treatment led to a significantly higher rank of the cards revealed. This shows that subjects are willing to sacrifice others even if they initially spared them. The *attention* treatment showed no significant effect.
In total, this study on the one hand reveals that individuals are excluded because of immutable traits and social behaviors. On the other hand, we find that people are willing to follow previous exclusion patterns to enrich themselves and create exclusion cycles that do nearly not allow an escape.
This finding is closely linked to the last study, *"Choosing a Victim you know"*. In this project, we used the Mobbing Game by Abbink & Doğan (2019) and modified it in two ways:

communication and payoff structure. In the baseline treatment, participants were matched in groups of four, being labeled only with capital letters (M, T, P, and G). For 20 rounds, they played the same game: participants were asked to nominate a player. In case three of them successfully coordinated on one victim, they would share the victim's payoff of the respective round. It was not possible to vote oneself but it was possible to abstain. Like the authors of the original article, we were interested in nomination rates (i.e. the attempt to coordinate on someone) and mobbing rates (the successful coordination on one victim for at least three consecutive rounds).

When adding communication via chat, we chose two different timing options: groups could chat (1) before (CBI) or (2) after knowing the instructions (CAI). With this, we distinguish between social communication (1) and strategic communication (2).

Introducing incremental payoffs, we created three treatments that increased the incentive to stick to a victim once the coordination was successful. In these treatments, the excluding group members not only shared the victim's payoff but it increased for up to eight rounds and then stayed at a high level. This was added to each level of communication leading to six treatments in total.

We find that communication decreases nomination rates when comparing treatments with the same payoff structure. Further, strategic communication was more often used to agree on not mobbing anyone instead of directly coordinating on one player. Thus, participants used communication in a strategic way when possible. However, we did not observe successful coordination on a specific player. The main driver of nominations and mobbing behavior was the incentive. Comparing communication treatments with and without incremental payments, we find that nomination rates significantly increased in the latter.

In total, we find support for the prosocial effect of communication since there are lower nomination rates in CBI than in the baseline treatment. However, this effect vanishes in the incremental treatments suggesting that participants are willing to take advantage of someone who is already down if it is profitable enough.

The main findings of the projects in this dissertation are that first of all, people do exploit the disadvantageous situations of other people and harm them if the monetary gains are sufficiently high. However, there are mechanisms that counteract this behavior. On the one hand, there is the insecurity about future situations and thus potential future punishments and on the other hand there is communication. Both mitigate the harming behavior induced by respective incentives.

List of essays:

Hoffmann, S., Seidel, A., & Weimann, J. (2021). Fairness under uncertainty. *Journal of Behavioral and Experimental Economics*, 94. https://doi.org/10.1016/j.socec.2021.101746

Seidel, A. (2024) "Laundering the people instead of the money - An experimental study on the effect of mental money laundering." *working paper*, pre-registered at: *AEA RCT Registry*. https://doi.org/10.1257/rct.12032-1.1

Seidel, A. et al. 2024. "The Cinderella Game – finding the one who will not go to the ball" *working paper*, pre-registered at *AEA RCT Registry*. https://doi.org/10.1257/rct.12911-1.2

Bershadskyy, D., & Seidel, A. (2024). Choosing a victim you know. *Journal of Behavioral and Experimental Economics*, 112. https://doi.org/10.1016/j.socec.2024.102265

References

Abbink, K., & Doğan, G. (2019). How to choose your victim. Games and Economic Behavior, 113, 482–496. https://doi.org/10.1016/j.geb.2018.10.006

Imas, A., Loewenstein, G., & Morewedge, C. K. (2021). Mental Money Laundering: A Motivated Violation of Fungibility. *Journal of the European Economic Association*, *19*(4), 2209–2233. https://doi.org/10.1093/jeea/jvaa059

Fairness under uncertainty[1]

Sönke Hoffmann[2], Alexandra Seidel[3], Joachim Weimann[4]

Abstract

With the help of a repeated matching pennies game, we model a situation in which one of two players may accidentally get into an adverse situation that the other player can take advantage of. In doing so, we vary the degree of uncertainty about future rounds and especially about the frequency and role distribution of situations in which fairness and positive reciprocity are possible. It turns out that uncertainty can greatly increase the number of fair moves when unfair behavior leads to a loss of efficiency. If there is no loss of efficiency and if unfairness only leads to a redistribution between players, we observe almost exclusively selfish behavior. We also vary the social distance and find that small distance leads to a significant increase in the proportion of fair moves even in games where there is no loss of efficiency.

Keywords

Fairness; reciprocity; random interaction; communication

JEL-Classification

C91, C92, H26

[1] All data and instructions are published on X-econ, following DOI: https://dx.doi.org/10.23663/x2644

[2] University of Magdeburg, Faculty of Economics and Management, Email: soenkehoffmann@ovgu.de

[3] Corresponding author, University of Magdeburg, Faculty of Economics and Management, Magdeburg Experimental Economics Laboratory (MaXLab), Postbox 4120, D-39106 Magdeburg, Germany, alexandra.seidel@ovgu.de.

[4] Joachim Weimann, University of Magdeburg, Faculty of Economics and Management, Chair in Economic Policy, Postbox 4120, D 39106 Magdeburg, Germany, weimann@ovgu.de

1 Introduction

Fairness and fair behavior are relatively prominent topics in the economic literature. The attention this term receives is quite appropriate, because fairness plays a major role in social life. An example from sports makes this clear. In 2009, France and Ireland faced each other in a playoff match for participation in the 2010 soccer World Cup. France won the first leg in Ireland 1-0, but in the second leg the score was 1-0 for the Irish after 90 minutes, so the game went into extra time. After a free kick, the ball went to Thierry Henry, but would have gone out of bounds if Henry had not kept it in play by handling the ball twice. After the handball, he pushed the ball in front of the goal, where William Gallas easily scored for the French to make it 1-1. Everyone had seen the handball – except the referee. Brave little Ireland was eliminated and the great soccer nation France went to the World Cup (where it did not win a single game and was eliminated in the preliminary round). The reaction was huge. The Irish sports minister demanded that the match be rescheduled, Thierry Henry's career was overshadowed by the incident, and a few years after the match it came out that FIFA had transferred 5 million euros to the Irish Football Association to appease the Irish. While L'Equipe at the time wrote of "The Hand of God" in reference to a similar situation in which Diego Maradona had scored a goal with his hand, the Irish Sun read of "The Hand of Frog". Obviously, the example shows that unfair behavior can come with costs that create an inefficiency. The example also shows that there can be different forms of inefficiency due to unfair behavior. The efficiency losses occurred both in the material space (loss of income) and in the utility space (even the French fans could not really enjoy the victory). Although we will focus on material efficiency losses in this paper, the possibility of utility losses due to unfair behavior should be explicitly noted.

The unfairness of which Henry was guilty was that he broke a rule in order to gain an advantage, gambling that he would not be caught breaking the rule.[5] The example is worth noting not only because it shows that unfair behavior can cause massive negative reactions and efficiency losses, but also because it shows that chance can play a significant role. It was by chance that Henry found himself in a situation in which he could decide the game by breaking the rules and by chance that he did not get caught doing so.

In the economics literature, relatively little attention is paid to this type of uncertainty in the treatment of fairness. The term "fairness" is used there in quite different contexts with quite different meanings. Essentially, three contexts can be distinguished in which fairness plays an important role. First, fairness is used as a synonym for reciprocity. Rabin (1993) equates fair

[5] Which, by the way, he unapologetically admitted and excused with the words "I'm not the referee." Bock (2009).

behavior with being nice to someone who is nice to you and at the same time being allowed to punish people who are not nice to you. In particular, experiments on the gift-exchange game (Fehr et al. 1997; Fehr and Gächter 2000) and on the investment game (Berg et al. 1995) have made positive reciprocal behavior [6] a very important behavioral assumption and have gained much prominence[7]. However, reciprocal behavior requires that there was an initial action that is reciprocated positively or negatively. The motive of the initial action remains open, but it is relevant for the question of whether reciprocity depends on the intention of the initial actor. Falk and Fischbacher argue for "intentional reciprocity", whereas the inequality aversion theories of Bolton and Ockenfels (1999) and Fehr and Schmidt (2000) do not necessarily rely on this.

The second context in which fairness matters concerns the allocation or distribution of resources. There is widespread agreement in the literature that the benchmark for fair allocations is that people who are equal in their rights should have equal shares of the pie to be distributed. Unequal distributions must therefore be based on inequality among people, for example, on different contributions in the creation of the pie (Konow, 1986). This notion of allocative fairness is the basis not only of inequality aversion models, but also of Charness and Rabin's (2002) model.

The third context in which fairness plays a role is described by the notion of "procedural fairness." Here, the question is whether mechanisms used to resolve conflicts can be considered fair. As a rule, this involves the use of random processes to bring about an unbiased decision. For this purpose, the process must be truly random, i.e., it must grant equal chances to all participants (Elsner 1989). Bolton et al. (2005) show in a very insightful experiment that a fair random process is considered equivalent to a fair (equal) division of a given pie. As long as the process does not favor anyone because it is based on unbiased probabilities, unequal results of the process are also accepted as fair.

Overall, although the concept of fairness is quite broad in economics, the economic applications of this concept do not quite fit our example given at the beginning. Of course Thierry Henry behaved unfairly, but in what sense? Was he not reciprocal? This could only be answered if we knew that the Irish players also had the possibility to decide the game with a simple hand movement, but did not do so because they wanted to follow the rules. But even if they had had such an opportunity, and even if they had tried and had been caught, one would still say that

[6] The ultimatum game, introduced by Güth et al. (1982), shows that negative reciprocity is also a behavior that can be reliably observed experimentally. Compare Güth and Kocher (2014) for an overview.
[7] For an overview, see Johnson and Mislin (2011).

Henry's behavior was unfair. The best explanation is provided by the concept of procedural fairness. This requires that in a soccer match the chances must be equally distributed and that a victory of a team can only be achieved by its performance within the rules. However, this interpretation does not take into account the fact that the example shows that chance plays an important role. The example therefore shows that there are use cases for the notion of fairness that are not yet completely described by economic research. This is especially true for repeated interactions in which chance plays a large role.

In this paper, we want to study a situation in which repetition and chance are of central importance and we will combine this with an investigation of how inefficiency caused by unfair behavior has an impact on the behavior of players. The game we consider does not reproduce the initial example, but describes a situation which allows us to investigate these two aspects. Although it does not exactly mirror the Henry example, there is also an interpretation from soccer. In general, the situation is as follows: Two players play a repeated game, which is fair in the sense that the random outcome of the game is symmetric, i.e. both players have the same probability of winning. The game is thus procedurally fair. By chance, however, a situation may arise in which one of the two players (which one is also random) may end up in a disadvantageous position. The other player can use this to his advantage by deciding to win the game in that round. The alternative is that he cancels the round and replaces it with one in which both players again have equal chances. The soccer example is told in the following story: Imagine a player of team A injures herself without a previous foul. Team A then plays the ball out to get a break. It is an unwritten rule that fair behavior of team B in this situation is considered to be playing the ball back to team A in order not to gain any (unfair) advantage from the injury. Reciprocity comes about if team B, when one of its players is injured, can count on team A to act just as fairly and kick the ball back.

The soccer example may not be described exactly by the matching pennies game we will use in this paper, but it is suitable to illustrate the basic considerations that led us to use this particular game. Furthermore, there are analogies in the economic world: business relationships, employment relationships or relationships in hierarchical management structures are typically long-term. In addition, it is often uncertain *when* reciprocal behavior is required. Both can be incorporated in a repeated matching pennies game.

The repeated matching pennies game has the following important property. When the number of rounds in which a player is at a disadvantage and the distribution of roles among players are unknown, the expected payoffs are the same for strictly fair behavior (cancellation of every round in which a player is at a disadvantage) and for strictly unfair behavior. This gives us the

opportunity to check whether this property is relevant for player behavior. To this end, the game is modified in a way that causes unfair behavior to be inefficient. Because of the symmetry of the procedure, which is still present, the procedure is still fair, but the overall payoffs are lower for unfair behavior than for fair behavior. This allows us to examine the influence of possible inefficiencies on the willingness to engage in fair behavior. Our introductory example allows us to illustrate the efficiency losses that can result from unfair behavior. If the French had qualified without breaking the rules, the joy would certainly have been greater than it was after Henry's handball. More than 80% of French fans felt that the team did not deserve to qualify. For the Irish, a regular defeat would also have been easier to bear than one caused by unfairness. Unfair behavior thus causes a loss of benefit on both sides and thus a loss of efficiency compared to a game played fairly.

Another research question concerns the influence of uncertainty on fairness behavior. For this purpose, we use three levels of uncertainty, including the case of complete information. With complete information, situations arise in which it is known when retaliation can be ruled out. This gives us the opportunity not only to study the direct influence of uncertainty, but also to answer the question of whether fair (or reciprocal) behavior is motivated by altruism or is used to avoid an unfair response from the other player. Furthermore, our paper contributes to the literature on the effect of uncertainty about the roles of the players. In the investment game, Burks et al. (2003) found that the trusting behavior changed when participants played both roles with different partners and knew this before playing the first game. Iriberri and Rey-Biel (2011) investigated the role of role uncertainty in repeated games with the help of a modified dictator experiment and found that social behavior occurs significantly more frequently under uncertainty than under certainty. They summarize the significance of their findings as follows:

> *"Our results warn against the use of role uncertainty in experiments that aim to measure the prevalence of independent preferences." (p. 160).*

This warning may be justified if one is really interested in the detection of independent preferences. Nevertheless, should role uncertainty indeed be an important feature of many reciprocal relationships, it is important to investigate them specifically and uncover their effect on the different motives of non-selfish behavior. Two other questions that our investigation aims to answer is the effect of more or less social distance on the frequency of fair behavior and the role of the costs incurred by fair behavior.

Overall, we hope to learn more about how fair behavior emerges or fails to emerge in repeated interactions under uncertainty. The paper is structured as follows: In Section 2, we describe the basic games used in our experiment. In Section 3, we explain our research questions in more

detail and, in Section 4, we describe the treatments we use in order to answer them and derive the corresponding hypotheses. In Section 5, we describe the experimental procedure and, in Section 6, we report the results of the experiments and perform the statistical analysis. Section 7 summarizes and discusses the results.

2 The basic games

Our experiment is a modified and repeated matching pennies game. The experiment consists of several rounds and the subjects are matched in pairs in which they remain for the entire session. The subjects have to deal with two types of games in the experiment. In total, there are 20 rounds to play consisting of 10 rounds of Game A and 10 rounds of Game B, which were mixed randomly at the beginning.

Games

Consider the version of a simultaneous two-player matching pennies game in Figure 1a. Each player chooses a number from the set {1,2}. E denotes the player who will win if the sum of numbers chosen is even, and O is the corresponding player who will win if this sum is odd. Played repeatedly, this game will generate approximately the same average payoffs if both players randomize their mixed equilibrium probabilities of 50%. This models the regular state of play between two similar teams in which both teams have about the same chance of benefiting from a situation (e.g. winning the ball after a duel). In the following, we will refer to this game as *Game A*.

Figure 1: The three basic games of the experiment. Numbers at the endnodes represent points acquired in the game.

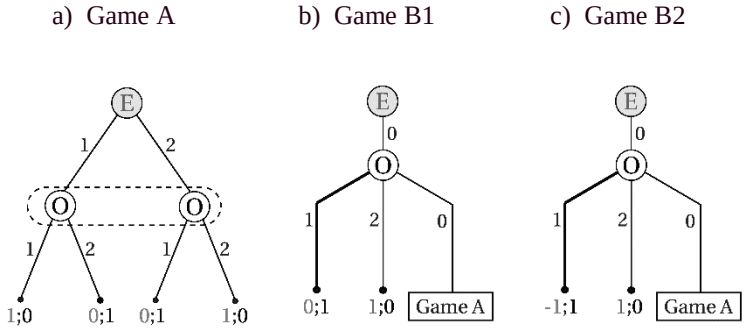


Figure 1 provides an example. In the experiment, the order in which the players decide (in the B-Games) can be as in Figure 1 or reverse. Getting back to the sports context, *Game B* represents a situation in which one team unexpectedly faces a disadvantageous situation and the other team can take advantage of that. One can think of a sudden injury that was not caused by an active foul on the part of an opponent player, but happened by accident. The team with

this injured player then usually shoots the ball out in order to get a break in the game to have the injured player treated. Consequently, the opponent team gets a throw-in. We allow for three different "throwing-in options" in this dominant position: (1) throw in the ball such that one's own team benefits (selfish choice), (2) throw in the ball such that the opponent team benefits (altruistic choice) or (0) throw the ball back to the player that had it before the injury such that the previous default state of play is restored (neutral choice). Two types of Game B incorporate this idea in our experiment.

In *Game B1* (cf. Fig. 1b), the first mover is forced to play a zero (have an injured player) whereas the second mover is free to choose one of the decisive moves as described above by choosing a number. He is thus in a position to decide who will win the corresponding point. If he plays the strategy "1" the payoff of the other player is zero. Moreover, he can also play a zero (neutral choice), which does not lead to an immediate payoff but to another round of *Game A* being played additionally at the end of the 20 rounds. In this case the expected payoff of the other player is 0.5. This choice represents a ball being *kicked back* to restore the regular state of play. In both B-Games, the first player is disadvantaged and the second advantaged. In this type of game, it holds that strictly playing selfish (1) on the part of the advantaged player (in what follows "the advantaged") in a B-Game leads to the same expected payoff for both players as strictly playing neutral (0) if it is expected that each player is in the advanced position equally often.

In *Game B2* (cf. Fig. 1c), the advantaged mover's selfish decision creates losses for the disadvantaged instead of leading to a neutral payoff of zero. In a soccer game, this may for example mean that the player throwing in not only passes the ball to his own team but also right into the penalty area (the box) of the injured team, where the risk of conceding a goal is much higher than somewhere else on the field. In this type of game, an unfair move by the advantaged creates an efficiency loss because of the negative payoff of the first mover. If we again assume that each player is in each position equally often, their expected payoff over all B-Games would be zero. The fair move (0) answered by strict reciprocal behavior would lead to an expected payoff of five points for each player.

After each round, or period, the players learn what their own and their partner's decision was and whether or not they have won a point. However, they do not receive information about the total points.

3 Research questions

The most important part of our study is devoted to the question of how uncertainty affects the willingness to engage in fair behavior. What is uncertain is the way in which situations that offer room for reciprocity and fairness arise. In our experiment, these situations are characterized by the two B-Games (B1 and B2). With regard to these B-Games, we consider three levels of uncertainty.

d0) Uncertainty about the number of B-Games, the time when they will occur and which player will be advantaged.

d1) The players know how often and when B-Games will take place, but it is not known who will be advantaged in the B-Games.

d2) Full information about the number and the timing of the B-Games and on the advantage and disadvantage in each B-Game

d0) might be the most realistic case. We use d1) and the case of perfect certainty d2) as reference points, which are necessary to estimate the effect of complete uncertainty. Before we formulate the research questions that we want to answer with the three levels of information, it is useful to first explain the function of the B2-Games.

Selfish behavior in a single game can lead to the partner being affected more or less. The decisive factor in this context is whether selfish behavior leads to a loss of efficiency. The following consideration makes this point clear: If a number of B-Games are played and players do not know in advance whether they will be advantaged in a B-Game, players will form expectations about how often they are in the two positions (advantaged or disadvantaged). It is very likely that due to a lack of information and the fact that their places in the lab were taken by chance, they will assume that both players will be in the two positions equally often. The principle of insufficient reason would also lead to this result. However, if fair or unfair decisions only affect the distribution of a given payoff, then the advantages and disadvantages of unfair or non-reciprocal behavior will balance out over all games. But if unfair behavior leads to a loss of efficiency, this is not the case. If there is no efficiency loss, strictly selfish behavior and strictly fair/reciprocal behavior will on average result in the same payoffs for both players. In this case, if everyone thinks of himself or herself, everyone is thought of. With an efficiency loss, both players still have the same expected payoff, but this payoff is smaller than without the efficiency loss.

With the help of the two B-Games and the three levels of information, it is possible to investigate questions regarding the effects of increasing uncertainty on the willingness to be

fair, on the one hand, and questions regarding the effect of inefficiency generated by unfair behavior, on the other. Of particular interest is the interaction between uncertainty and efficiency. This gives rise to three main research questions, which we attempt to answer using a 3x2 design:

(1) What is the effect of decreasing uncertainty (from d0 to d1 to d2) in a repeated procedurally fair game if it holds that choosing the self-interested strategy (1) does not lead to a loss of efficiency?

As mentioned earlier, the peculiarity of the B1-Game is that in this game the principle "If everyone thinks only of himself, everyone is thought of" applies. This means that even selfish behavior always leads to a fair payoff. Under such conditions, can fair behavior (i.e., that of the option (0 = neutral)) be expected? And does uncertainty matter at all in this context?

(2) What is the effect of decreasing uncertainty (from d0 to d1 to d2) in a repeated procedurally fair game if it holds that choosing the self-interested strategy (1) leads to a loss of efficiency?

As before, the game is procedurally fair because the expected payoffs are the same for both players. However, choosing the self-interested option (1) now generates a lower payoff to both players than that expected under fair behavior (option (0)). Under these conditions, does decreasing uncertainty have the same effects as in B1-Games?

Comparing the B1- and B2-Games helps answer the question of whether the inefficiency caused by selfish choices (option (1)) in the B2-Games leads to a change in behavior.

(3) Is there more fair gaming (option (0)) to avoid a loss of efficiency? Does the frequency of fair behavior interact with the degree of uncertainty?

A fourth research question relates to the motive for fair behavior of the advantaged player:

(4) Is it altruism or the desire to escape later punishment that leads advantaged players to play fairly?

The reason this question can be answered is that, given complete information, it is known if and when there are any opportunities for punishment. Under information conditions d2), it is revealed that the number of rounds in which a player is advantaged is asymmetrically distributed. One player is favored six times and the other only four times. This gives room for the more often advantaged player to use this information to realize an advantage over the other player when the latter no longer has an opportunity to execute a punishment. Under d2 the B-games are no longer *procedurally fair.*

The first four questions can be answered with the treatments of the 3x2 design, which results from the combination of the two B-Games with the three information levels. All these treatments are played single blind, i.e. the players cannot observe their respective partner, but

the experimenter observes them. Our fifth research question focuses on the effects of varying social distance. This is possible in both directions starting from a single-blind procedure. This leads to the question:

5. How does willingness to play fairly in B-Games change when social distance is decreased or increased?

Our final research question concerns the cost to advantaged players of playing fairly in a round:

6. If the cost of making a fair decision is halved, does this have an effect on the frequency with which a fair decision is made?

The following section explains how the six questions are transformed into corresponding treatments and hypotheses.

4 Treatments and hypotheses

In order to investigate Research Questions (1) to (3), 10 A-Games were combined with 10 B-Games, having an equal number of B1 and B2 treatments. Both variants were then played in three different information modes: no information (d0), information only on the number and timing of the B-Games (d1) and complete information including the role the players played in each B-Game (d2). This results in a 3x2 design. For all six treatments in this design, the same order of A- and B-Games and the same role assignment applies as shown in Table 1:

Table 1: Sequences of B-games played over 20 rounds. The numbers 1 and 2 in each row indicate whether Player 1 or Player 2 was the decider in the B-Game. Blank fields represent the regular state of play (A-Game).

Round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B-Game	x		x			x			x		x	x		X	x		x			x
Deciding player	2		1			1			2		2	1		2	2		1			2

The distribution of the roles is not symmetrical in the 10 B-Games. While player 2 decides six times, player 1 is in the position to decide only four times. Furthermore, player 2 decides in the last B-Game, which is of particular strategic importance because player 2 can choose the selfish action without fear of punishment. This asymmetry has been deliberately introduced because it leads to the fact that the argument "if everyone thinks of himself or herself, everyone is thought of" in the B1-treatment does not apply under full information (d2), whereas it does in the d0 and d1 treatments. Table 2 summarizes the treatments of the 3x2 design, all of them performed as single-blind treatments (sb):

Table 2: Overview of the treatments in the 3x2 design.

Treatment	Efficiency effect of unfair decisions	Level of Information
sb-B1-d2	no	full information
sb-B2-d2	yes	full information
sb-B1-d1	no	Information about timing, no information about advantage
sb-B2-d1	yes	Information about timing, no information about advantage
sb-B1-d0	no	No information about timing, no information about advantage
sb-B2-d0	yes	No information about timing, no information about advantage

With regard to these six treatments, we formulate three hypotheses. The first concerns the behavior in the B1-Games. In the following, we refer to the decision "1" in a B-Game as a *selfish* decision, "0" as *neutral* and "2" as *altruistic* (see Figure 1). We assume that under the information conditions d0 and d1, the players will expect symmetry, i.e. that both may decide equally often. In this case, the two strategies "always play 1" and "always play 0" lead to the same expected payoffs across all games. However, a player who plays "always 0" would run the risk of the other player not behaving reciprocally. In this case, "always 0" is disadvantageous for the player. In contrast, "always 1" is a safe strategy that does not harm the other player, because by choosing the same strategy, the player can secure the same payoff. For this reason, we expect that in the B1-Games under d0 and d1, only the “always 1” strategy will be chosen. Under complete information (d2), it becomes obvious that symmetry is not present because player 2 decides in six rounds and player 1 only in four rounds. Selfish play will therefore lead player 2 to have six points from the B-Games while player 1 has only four. Since we expect that "1" is always played under d0 and d1, the behavior shown under d2 can only differ if “0” is played. A reason for choosing the "0" strategy for d2 in the B1-Game could be that player 2 refrains from exploiting the asymmetry of the game to his advantage for altruistic reasons. We therefore expect that more "0" moves are observed under d2 than under d0 and d1. From the literature, Cox 2004 provides a hint that supports this expectation. He shows, in an arrangement with three experiments, that altruistic motives as well as "conditional kindness" can exist simultaneously. There is no reason for an altruistic move ("2") in the B1-Games.

In summary, the following hypotheses will cover Research Questions (1) and (3):

1. Hypotheses B1-Games:

a) Under the information conditions d0 and d1, i.e. with role uncertainty, only selfish ("1") moves will be played in the B1-Games.
b) Under complete information (d2), neutral moves ("0") will also be observed.
c) Altruistic moves will not be observed under all three information conditions.

In the B2-Games, a selfish move causes the other player to suffer a loss that could be avoided with a neutral or altruistic move. A selfish move is advantageous for the player who nevertheless decides for it because it leads to a sure win of one point, while a neutral move has an expected value of half a point. This constellation means that two motives can become effective, neither of which are expected in the B1-Games. On the one hand, a player who is faced with choosing selfishly or neutrally (we do not expect altruistic decisions in this game either) could feel sympathy for the other player and therefore forgo the "1" for reasons of altruism. On the other hand, a player who chooses the selfish move may expect that the other player will punish him with an equally selfish move in the future. Therefore, besides altruism, fear can also be a motive for a neutral move. We expect both motives to be effective in the B2-Games. The altruistic motive should not depend on how many games are left and who can still make decisions. Therefore, this motive should be effective in all 20 rounds of the game. In contrast, fear will play an increasingly minor role in the course of the experiment as the number of possibilities for retaliation decreases. In addition, the decline in the fear motive under the information condition d2 should be the fastest because player 2 recognizes that he has a stronger position than player 1. It is in the second half of the game, in particular, that fear should no longer play a major role for him.

When fully informed, player 1 knows that player 2 is in a strong position and in the first rounds, she will realize if player 2 is playing selfishly or not. If she expects player 2 to play selfishly in the second half (because he did so in the first), there is no reason for her not to be selfish as well. Neutral moves have the function of an insurance against retribution under role uncertainty. Therefore, we expect that the highest number of neutral moves will be observed in d0 and d1. Overall, uncertainty will thus lead to an increase in non-selfish moves. This hypothesis is supported by the results of Iriberri and Rey-Biel (2011).

Our second hypothesis concerns Research Questions (2) and (3):

2. Hypotheses B2-Games

a) The number of neutral moves will be less under full information than under role uncertainty.
b) The number of selfish moves in all three information orders will be lower in rounds 1-5 than in rounds 6-10.
c) Altruistic moves will not be observed in the B2-Games either.

The question arises as to how the B1- and B2-Games compare (Research Questions (2) and (4)). Our hypotheses in this regard follow from the considerations we have already made:

3. Hypotheses for the comparison of B1- and B2-Games

a) Although the proportion of selfish moves will increase in the B2-Games when players are fully informed, fewer selfish moves will nevertheless be observed in all the B2-Games than in all the B1-Games.
b) While the degree of uncertainty in the B1-Games has no influence on the number of selfish moves, in the B2-Games the number of selfish moves will increase with increasing information.

All the treatments described so far were performed as single-blind experiments. The experimenter was able to observe how the individual subjects behaved, but the subjects had no contact with each other. To investigate how fair or reciprocal behavior depends on the observability of individuals, we conducted the B1- and B2-Games under two additional conditions. In a double-blind treatment, it was ensured that the experimenter could not observe how the individual subjects behaved. In the "visible" arrangement, the players who formed the pairs had eye contact before the experiment, so that they could visually identify each other. In the following hypotheses, we use the single-blind arrangement as a reference point. We expect that the abandonment of anonymity in the treatment "visible" will lead to a reduction of the social distance between the partners. In the literature, a smaller social distance is associated with the expectation that selfish behavior is less pronounced (Hoffman et al. 1996; Brosig et al. 2003). However, this does not provide a clear prognosis for the B1-Games, where strictly selfish behavior delivers the same payoff as strictly neutral behavior. Therefore, we expect that the reduction of social distance will have only a weak effect there and will lead to a moderate increase in the number of neutral "0" moves. In the B2-treatments, on the other hand, the visibility of the partners should lead to a significant increase in the number of neutral moves compared to the single-blind arrangement.

A double-blind arrangement has the function of creating an increase in social distance. Accordingly, the opposite effects compared to "visible" would be expected: the number of selfish moves in the B1-Game should not decrease and it should increase in the B2-Game compared to the single-blind. We summarize this in our final hypothesis, which concerns Research Question (5):

4. Hypotheses on the effect of observability of individual actions:

a) In the B1-Games, there will be no difference between the single-blind and double-blind variants.
b) In the B1-Games, the proportion of neutral decisions in the treatment "visible" will increase only moderately compared to the single-blind variant.
c) In the B2-Games, the double-blind variant will lead to an increase in selfish decisions.
d) In the B2-Games, the visibility of the partners will lead to an increase in neutral moves.

The final variation we undertook occurred in a B1-Game: we reduced the payoff per point, which was otherwise €2, to the value of €1[8] . This makes it cheaper to play neutral because the expected loss is halved. We do not expect this to have a significant impact. Table 3 summarizes all the treatments.

Table 3: Treatments and participants

Treatment	Efficiency effect of unfair decisions	Level of information	Level of anonymity	Number of participants (female)
sb_B1_d2	no	full information	sb	36 (21)
sb_B2_d2	yes	full information	sb	34 (19)
sb_B1_d1	no	Information about timing, no information about advantage	sb	40 (19)
sb_B2_d1	yes	Information about timing, no information about advantage	sb	40 (27)
sb_B1_d0	no	No information about timing, no information about advantage	sb	40 (19)
sb_B2_d0	yes	No information about timing, no information about advantage	sb	40 (23)
sb_B1_1_d0	no	No information about timing, no information about advantage	sb	40 (17)
db_B1_d0	no	No information about timing, no information about advantage	db	40 (22)
db_B2_d0	yes	No information about timing, no information about advantage	db	40 (19)
vis_B1_d0	no	No information about timing, no information about advantage	vis	40 (20)
vis_B2_d0	yes	No information about timing, no information about advantage	vis	40 (17)

[8] Short form: sb_B1_1_d0

5 Experimental procedure

The experiment took place in the MaXLab at the Otto von Guericke University in Magdeburg and was organized and recruited with the software *hroot* (Bock et al. 2014). We programmed it with *z-Tree* (Fischbacher 2007). The only requirement subjects needed to fulfill was an adequate command of the German language in order to understand the instructions[9] properly. All of the 430 participants were students from all faculties of the university, with 225 of them being female.

After a short welcome, the participants drew a wooden sphere that displayed their booth numbers in the laboratory. They then went to their individual booths, where the instructions were already on the tables. After the instructions were read aloud, the participants were able to ask questions in private. When everyone understood the instructions, the session started with twelve practice periods that were not relevant for the payoff. Subsequently, the 20 periods started. A session ended with the payment procedure.

In order to establish the different levels of anonymity, we implemented different procedures immediately after the participants drew their booth number. In the double-blind treatments, to maintain anonymity, the individuals also drew a three-digit number in secret, which later on was their identification number for the payment procedure. They then proceeded to their booths, closed the doors or curtains and stayed there alone. After the experiment, the players went to another room to receive their payment and handed their ID to the experimenter, who was standing behind a screening wall. In the single-blind treatments, they simply went to their individual booths without knowing whom they were matched with, but faced the experimenter in the payment procedure. At the beginning of the visual treatments, the participants met their partner and made eye contact without communicating verbally. They then went to their booths, which they could leave open, and they met the experimenter face to face in the payment procedure.

The average payoff amounted to is €17.73. In total, there were 33 sessions with 13 participants on average. We thus had 430 participants and 17 to 20 independent observations per treatment (because of the pairwise matching).

[9] All data and instructions are published on X-econ, following DOI: https://dx.doi.org/10.23663/x2644

6 Results

Table 4 provides an overview of all aggregated data on all matches in all treatments.

Table 4: Total number of decisions over all rounds.

	Game A			Game B			
Treatment	1	2	total	alt	ego	neut	total
sb_B1_d2	196 (0.54)	164 (0.46)	360	4	159	17	180
sb_B2_d2	186 (0.55)	154 (0.45)	340	10	103	57	170
sb_B1_d1	214 (0.54)	186 (0.46)	400	0	159	41	200
sb_B2_d1	209 (0.52)	191 (0.48)	400	6	94	100	200
sb_B1_d0	221 (0.55)	179 (0.45)	400	1	184	15	200
sb_B2_d0	207 (0.52)	193 (0.48)	400	12	67	121	200
sb_B1_1_d0	209 (0.52)	191 (0.48)	400	2	174	24	200
db_B1_d0	205 (0.51)	195 (0.49)	400	13	147	40	200
db_B2_d0	221 (0.55)	179 (0.45)	400	11	79	110	200
vis_B1_d0	219 (0.55)	181 (0.45)	400	4	122	74	200
vis_B2_d0	199 (0.50)	201 (0.50)	400	14	80	106	200

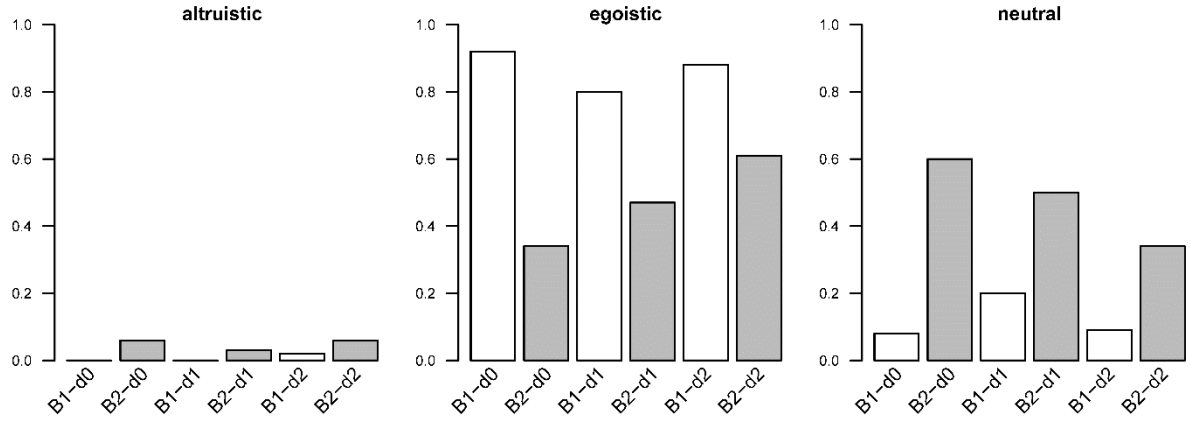
A comparison of treatments sb_B1_d0 and sb_B1_1 confirms the assumption that halving the costs of a neutral decision does not lead to significantly more neutral play in the B1-Game ($p = 0.18$, 2-sample chi-square-proportions test)[10]. We will now present the detailed results in the order of Hypotheses 1 to 4.

Figure 2 shows the aggregated data of the six treatments of our 3x2 design to investigate the effect of uncertainty and inefficiency of unfair play. The left part of the figure shows that our Hypotheses 1c and 2c that altruistic decisions will not play a major role can be confirmed. The proportion of altruistic decisions is between 0 and 7 percent, which is very low. The middle and right parts of Figure 2 show the total number of selfish and neutral moves. In the B1-Games under complete uncertainty (d0), about 90 percent of the moves are selfish. In Hypothesis 1a, 100 percent is expected, but the result is very close to the prediction. Under the information in d1, the percentage of selfish moves in B1 is very high, but with 80 percent it is significantly lower than under d0 ($p < 0.01$). With respect to d1, Hypothesis 1a is only confirmed in the tendency. In Hypothesis 1b, we express the expectation that under full information "0" moves will also be observed in the B1-Games. However, under complete information (d2), the number of *neutral* moves unexpectedly decreases to 9% compared to 20% in d1, and the number of

[10] In the following, we refer to this test without explicitly mentioning it. As a robustness test we also run a MWU tests and a t-test on the aggregated proportion of neutral moves for each pair. Only in one case the results of the proportion test was not confirmed (see footnote 12).

selfish moves in fact *increases* to 88% (both changes are only weakly significant).[11] Furthermore, there is no significant difference between sb_B1_d0 and sb_B1_d2 in the number of selfish moves ($p = 0.30$). This all means that full information does not trigger the fair behavior that we expected to take place. In summary, if selfish behavior does not lead to a loss of efficiency, as in the B1-Games, altruistically motivated neutral moves are unlikely to occur regardless of how much information the players have. The principle "if everyone only thinks of him- or herself, everyone is taken care of " seems to prevail in this game and the vast majority of players therefore obviously see no reason to deviate from a purely selfish way of playing.

Figure 2: Total shares of choices in a single-blind (sb) scenario with different levels of information about the B-Games (d0 = neither rounds (in which the B-games were played) nor roles known, d1 = rounds known, d2 = rounds and roles known) and Games B1 (blank) and B2 (grey).



The picture is completely different in the B2-Games. Although the altruistic moves "2" do not play a role there either, the percentage of neutral moves is much higher than in the B1-Games. Furthermore, the share of selfish moves in d1 compared to that in d0 increases and it does likewise in d2 compared to d1. Both increases are significant[12] ($p < 0.01$ and $p = 0.01$), thus supporting Hypotheses 2a and 3b, which state that in B2-Games more information will lead to more selfish moves. Hypotheses 3a and 3b, expressing the idea that inefficiency resulting from unfair behavior will lead to more fair moves, can also be confirmed. Although the proportion of selfish moves increases with increasing information, it always remains below that of all the B1-Games. All differences in this respect are significant ($p < 0.01$).

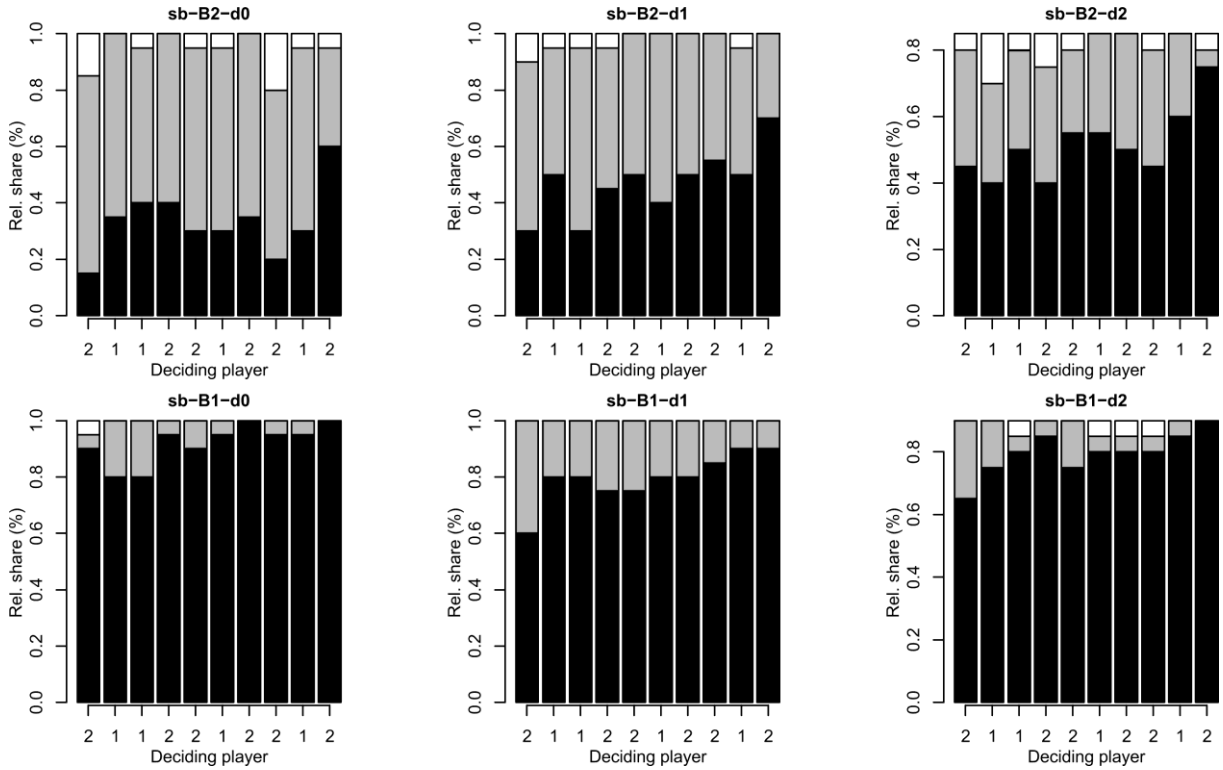
Figure 3 shows the proportions of the three possible moves during the 10 rounds of the B-Games. It can be seen that the proportion of selfish moves in the B2-Games is on average

[11] Here and in the following, we use a two-sided single proportion test.
[12] This holds for the proportion test. Using the MWU and the t-test, only the second increase is significant.

somewhat smaller in the first five rounds than in the last five rounds, as stated in Hypothesis 2b. A Wilcoxon signed rank test shows that the difference is significant[13]. The same holds for the B1-Games. This observation is an indication that, in the course of the game, the players learn that there is an equivalence between "always neutral" and "always selfish".

It is striking that there is a strong final round effect in the B2-Games in all information scenarios. This is possible because the participants were informed before the last B-Game in all treatments that this was the last B-Game. Note that, in this final round, the advantaged player can be sure that the disadvantaged has no opportunity to punish her. Obviously, the neutral move ("0") is also driven by the fear of being punished. Nevertheless, even in the last B-Game, the number of selfish moves is always lower in the B2-Games than in the B1-Games, although this difference becomes smaller with increasing information levels d0 to d2.

Figure 3: Selfish (black), neutral (grey) and altruistic (white) shares in B-Games over time (without A-Game periods) for the 3x2 factorial design with factor levels "no information" (d0), "known B-Game periods" (d1), "known B-Game periods and roles" (d2) and "no efficiency loss" (B1), "efficiency loss" (B2).



To gain further insights, we disaggregate the data in Figure 4:

[13] The same holds for a matched sample t-test.

Figure 4: Sequences of Game B decisions. The factorial design, color coding and notation are the same as in Figure 3. Each field stands for the decision of the deciding partner indicated on the lower axis.

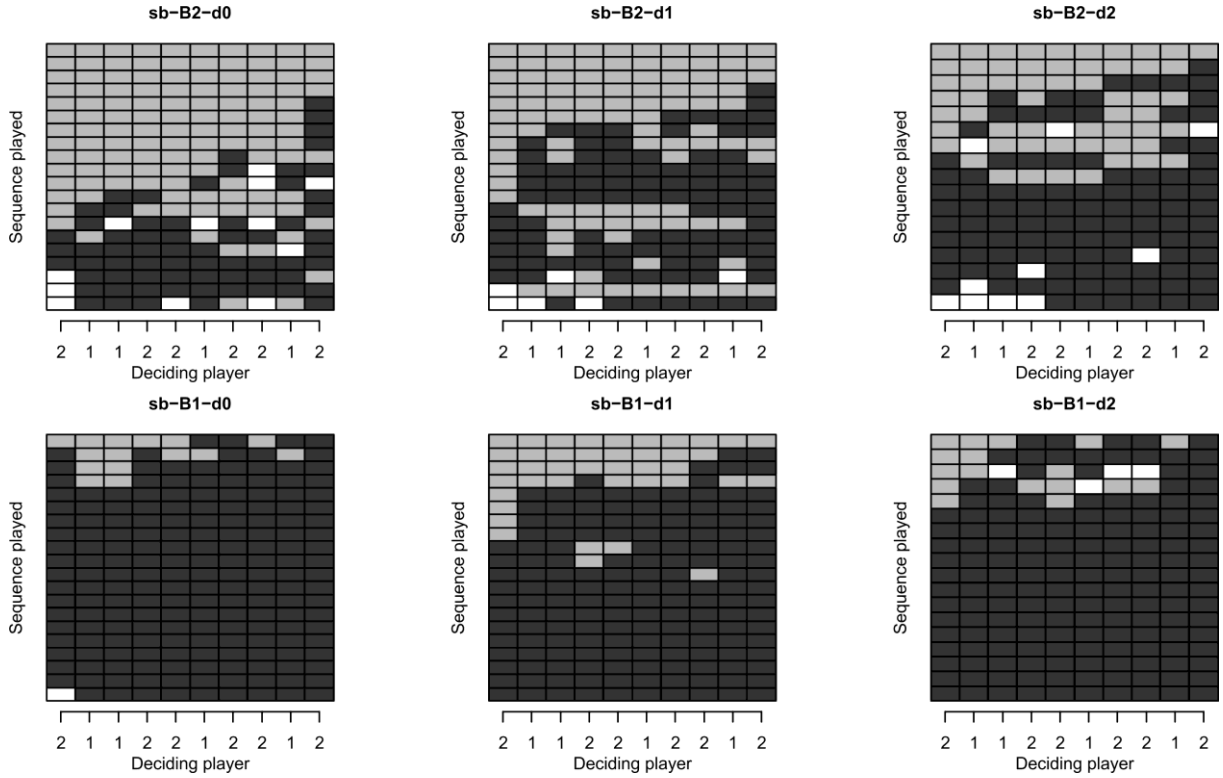


Figure 4 shows that, in the B1-Games, the relatively large number of neutral moves in the d1 variant is due to the fact that three pairs tried to play "always 0" instead of "always 1". For one pair, this was also successful until the last round. In two pairs, this "always 0" strategy collapsed in the last two or three rounds probably because the players noticed that the distribution of roles is not symmetrical in these rounds. Otherwise, the "always selfish" strategy is very popular with the vast majority of pairs in the B1-Games.

The picture is much more heterogeneous in the B2-Games. First, it is noticeable that the continuous grey area in the upper left of the graphs, which shows the pairs that always, or at least for a certain time, only made neutral moves, becomes smaller as the information increases. Although altruistic “2” moves play only a minor role overall, Figures 4 and 5 show that most of the few "2" moves occur in the B2-Games. A possible explanation for this could be that the realization of the fact that a player can suffer damage in the B2-Games awakens altruistic motives. Figure 4 shows that these altruistic motives do not occur systematically and do not depend on the amount of information available to the players.

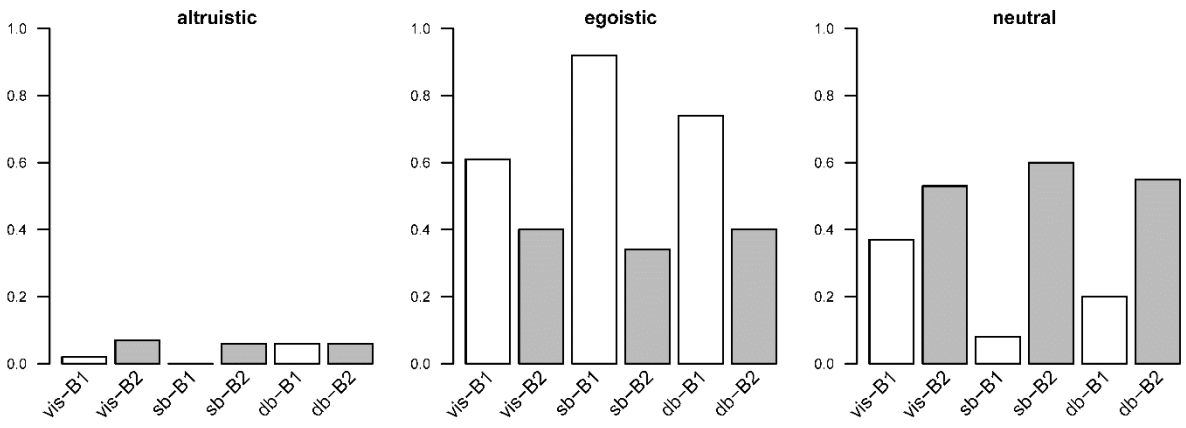
In summary, the results on the influence of information and inefficiency on the fairness and reciprocity of players (research questions 1 to 3) are the following:

- In the B1-Games, uncertainty about the distribution of roles and the number of B-Games does not have a noticeable effect on behavior. Since selfish play does not generate a loss of efficiency, the simplest solution is to always play selfishly, regardless of the information situation.
- In the B2-Games, the willingness to play neutrally depends on the players' information. The better informed they are about how many B-Games are still to come and what role they will play in them, the less inclined they are to play neutrally. This indicates that strategic considerations (avoidance of punishment) play an important role in the decisions of the advantaged players.
- On the other hand, even with complete information, the number of selfish moves is clearly and significantly smaller in the B2-Games than in the B1-Games, which suggests that, in addition to strategic reciprocity, the efficiency loss of unfairness also leads to neutral moves in the B2-Games.

One may argue that the last finding can be explained by some kind of social preferences. For example, inequality aversion may be a candidate. But note that this kind of preferences concerns the way people care about inequality. In our experiment, the expected payoffs in the B1-Games as well as in the B2-Games are *the same for both players*. Inequality aversion could nevertheless play a role if players are aware that the opportunity for unfair behavior is not symmetrically distributed. Under the corresponding information condition (d2), however, we do not observe an increase in social behavior, but more self-interest. Thus, social preferences are not able to explain our findings.

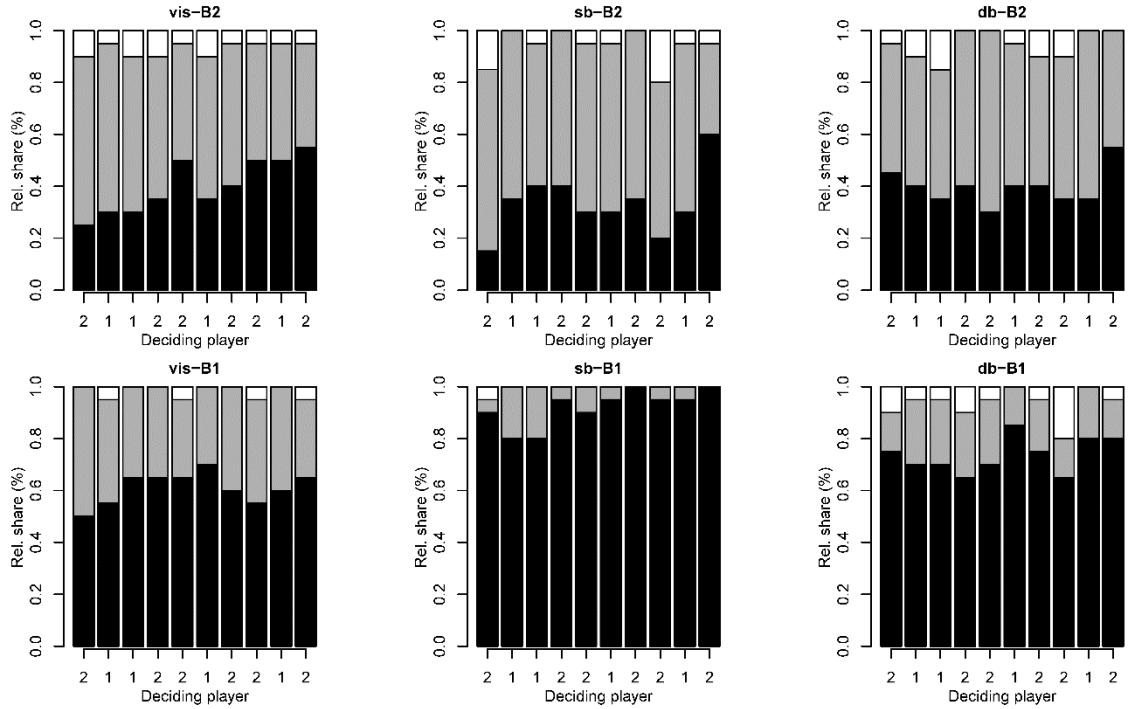
In addition to the effect of uncertainty and inefficiency on the willingness to behave fairly or reciprocally, we investigated if and how the visibility of decisions, i.e. social distance, determines reciprocal or fair behavior. In the following figures, we will re-list the single-blind arrangements of the B1- and B2-Games we have already reported on in order to present the full range of "visible", "single-blind" and "double-blind" arrangements. Figure 5 first shows the proportions of the three possible moves, altruistic ("2"), selfish ("1") and neutral ("0"). Note that all the reported treatments were conducted under the information level d0 (no information about timing and advantage).

Figure 5: Relative frequencies of Game B choices aggregated over rounds in the 3x2 factorial design with anonymity levels visible (*vis*), single-blind (*sb*), double-blind (*db*) and the efficiency effect of fair behavior (*B1* = no, white), (*B2* = yes, grey) for information level d0.



Again, altruistic moves occur only rarely (a maximum of 7%). The selfish and neutral moves show that the B2-Games do not react to the variation of the social distance. Although the number of neutral moves under the single-blind arrangement is slightly higher than under the other two arrangements, the differences are not significant ($p = 0.16$ and $p = 0.31$). This observation contrasts with the expectation we express in Hypotheses 4c and 4d that lower social distance will increase, and higher social distance will decrease, the number of fair ("0") moves. With the single-blind arrangement as a reference point, there is an increase in neither the number of neutral moves when the partners can identify each other nor the number of selfish moves when the players are protected from observation by other players and the experimenter. The situation is different in the B1-Game. Compared to the single-blind variant, the proportion of neutral moves increases significantly, both in the "visible" arrangement ($p < 0.01$) and under double-blind conditions ($p < 0.01$). However, the increase is stronger when the subjects have eye contact. This is consistent with Hypothesis 4b, even though the increase is greater than we expressed in this hypothesis. Hypothesis 4a, stating that there will be no difference, on the other hand, must be clearly rejected. Figure 6 shows the proportion of the three possible moves for each round separately.

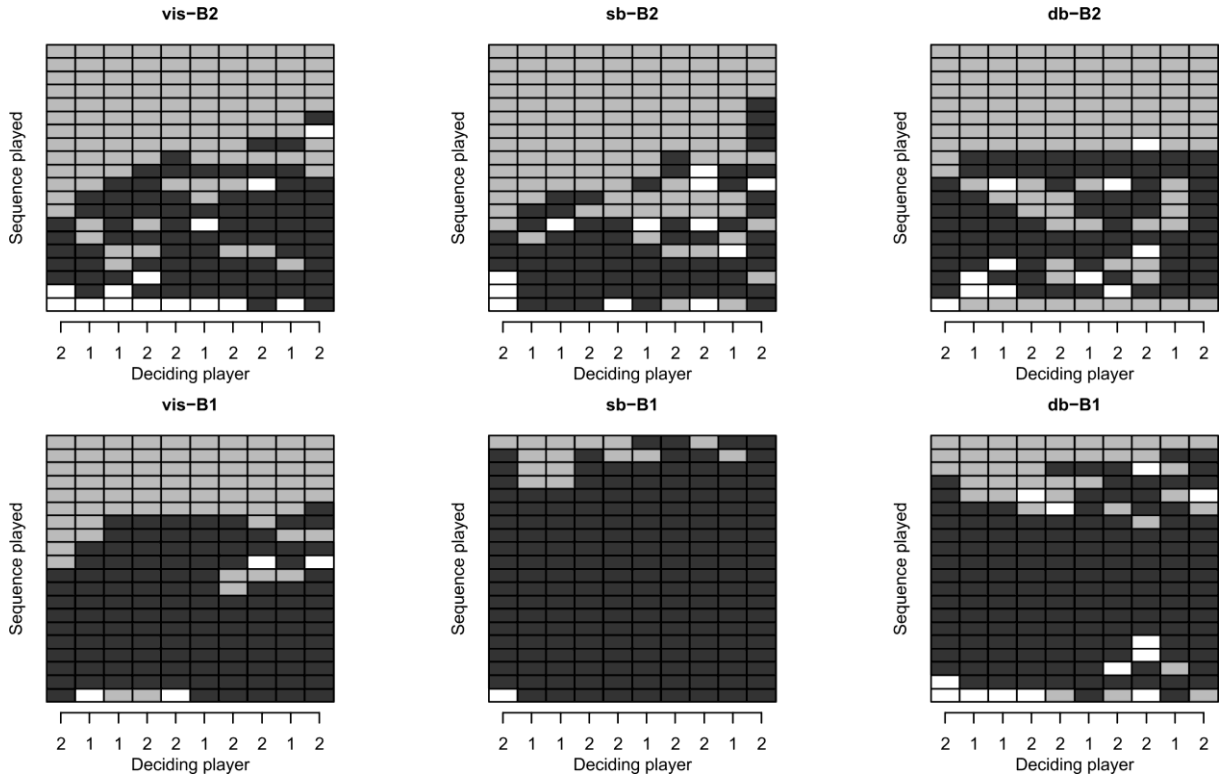
Figure 6: Selfish (black), neutral (grey) and altruistic (white) shares in B-Games over time (without A-Game periods) for the 3x2 factorial design with factor levels "visible" (vis), "single-blind" (sb), "double-blind" (db) and the efficiency effect of fair behavior (*B1* = no), (*B2* = yes). The information level for all is d0.



Again, there is a clear final-round effect in all the B2-Games, which is not observed in the B1-Games. Two further observations are worth mentioning. Only in vis-B2 is there a clear increase in selfish decisions over the 10 B-Game rounds. There is no trend to be observed in the other five treatments. Secondly, it is noticeable that altruistic moves again occur mainly in the B2-Games – with the exception of the double-blind arrangement of the B1-Game. We will come back to this point later. Figure 7 shows the disaggregated data of the six treatments.

If the players can identify each other in the B1-Game, there is a clear division into two "camps". 25% of the pairs (5 out of 20) successfully decide to always play "neutral". All others play predominantly "selfish", although only seven pairs succeed in maintaining this strategy in all rounds. Seven pairs deviate from this one to three times and one pair four times. In the B2-Games, it is noticeable that eight to ten pairs (almost half of them) played almost continuously neutral (a maximum of one deviation). In contrast, the remaining pairs show a very heterogeneous behavior and change strategy relatively often. Therefore, in 14 pairs (in the three B2-treatments) all three moves are played. This finding suggests that the players in the B2-Games were very unsure what the appropriate behavior in this game was. This in turn could be an indication of how to explain the, at first glance, paradoxical observation that more selfish behavior cannot be observed under double-blind conditions than under single-blind conditions.

Figure 7: Sequences of Game B decisions within a pair in the 3x2 factorial design. Color coding and notation are the same as in Figure 4.



It is known from the literature (Barmettler et al. 2012) that a double-blind arrangement can trigger an experimenter demand effect if the double-blind procedure is explicitly referred to in the instructions (as it was in our instructions). In this case, subjects may get the impression that the experimenter is deliberately trying to make selfish behavior particularly easy. This in turn can lead to the expectation that the experiment is about showing that non-selfish behavior can be observed even when it is especially easy to behave selfishly. If subjects should pursue the goal of doing what is "expected" in the experiment, this can lead to their consciously not behaving in a selfish way in a double-blind arrangement. This explanation fits to the observation that in the B2-Games, many subjects apparently did not quite know what the appropriate or correct behavior in this experiment was.

In summary, it can be stated that the visibility of the behavior, or the social distance, has only a very limited influence on the behavior in the B2-Games. Only the arrangement with the smallest social distance (visible) shows a rounding effect, which we cannot observe in the other B2-arrangements. After the personal interaction, the willingness to behave neutrally is initially very pronounced, but decreases in the following rounds. This could indicate that the social bond created by the face-to-face contact is gradually masked by the experiences of the A- and B-Games and therefore loses influence.

In the B1-Games, on the other hand, the influence of less social distance is clearly noticeable. Since the two strategies "always selfish" and "always neutral" lead to the same results, it is easy to decide for the neutral variant and the lower social distance might have been the reason for this. Finally, there is much to suggest that the observations in the double-blind treatment must be attributed to an experimenter demand effect.

7 Conclusion

Our experimental results show that uncertainty about the structure of future interactions is a strong motive for fair or positively reciprocal behavior. This suggests that part of the reciprocity observed in the real world is due to the conviction that "you always meet twice in life". When interacting with a person, it cannot be ruled out that you will interact with the same person again in the future, possibly in different roles. This is particularly important if unfair behavior is to the detriment of the partner, with the result that a loss of efficiency arises in relation to a fair solution. In more symmetrical situations, on the other hand, the rule that everyone is taken care of if everyone takes care of him- or herself has prevailed in our experiment.

The observability of behavior only plays a role if fair and unfair behavior in stochastic interaction (i.e. randomly changing roles) leads to very similar results. Returning to our example from soccer, it cannot be ruled out that being observed by the spectators contributes significantly to the fact that the players almost always pass the ball back. On the other hand, our results show that if unfair behavior is destructive insofar as it creates an efficiency loss, the observability of behavior is not needed to incentivize fairness and positive reciprocity. Nevertheless, even in the case of an efficiency loss, roughly half of the subjects behave strictly selfishly – and even more, the safer they are from punishment. Overall, our results show that reciprocity and selfishness are both deeply rooted human traits. Given this, the fact that we live in a non-deterministic environment is quite helpful. For the fear of meeting one another again, perhaps in reversed roles, is an important driver of fairness and positive reciprocity.

References

Barmettler, F., E. Fehr, and Ch. Zehnder: "Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory" *Games and Economic Behavior*, 75 (2012), pp. 17-34

Berg, J., Dickhaut, J., and K. McCabe: "Trust, reciprocity, and social history", *Games Econ. Behav*. 10, (1995), pp. 122–142.

Bock, O., I. Baetge, and A. Nicklisch: "hroot: Hamburg Registration and Organization Online Tool", *European Economic Review* Vol. 71, issue C (2014), pp. 117-120

Brosig, J., J. Weimann, and A. Ockenfels: "The Effect of Communication Media on Cooperation". *German Economic Review* 4, (2003) pp. 217–241. https://doi.org/10.1111/1468-0475.00080

Burks, S. V., J. P. Carpenter and E. Verhoogen: "Playing both roles in the trust game", *Journal of Economic Behavior & Organization, 51(2003),* pp. 195–216

Charness G., and M. Rabin: "Understanding Social Preferences with Simple Tests", *The Quarterly Journal of Economics*, Volume 117, Issue 3, (August 2002), pp. 817–869.

Cox, J. C.: "How to identify trust and reciprocity", *Games and Economic Behavior*, 46, (2004) pp. 260–281.

Dhami, S.: The Foundations of Behavioral Economic Analysis, Oxford (2016)

Falk A., and U. Fischbacher: "A theory of reciprocity", *Games and Economic Behavior*, 54, 2, (2006), pp. 293-315.

Fehr, E., and S. Gächter: "Fairness and retaliation: the economics of reciprocity" *Journal of Economic Perspectives* 14, (2000), pp. 159–181.

Fehr, E., S. Gächter, and G. Kirchsteiger: "Reciprocity as a contract enforcement device: experimental evidence." *Econometrica* 65, (1997), pp. 833–860.

Fischbacher, U.: "z-Tree: Zurich Toolbox for Ready-made Economic Experiments", *Experimental Economics* 10(2), (2007), pp. 171–178

Geanakoplos, J., D. Pearce, and D. Stacchetti: "Psychological games and sequential rationality", *Games and Economic Behavior*, Volume 1, Issue 1, (March 1989), pp. 60-79.

Güth, W., and M. G. Kocher: "More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature", *Journal of Economic Behavior & Organization*, Vol. 108, (2014) pp. 396-409

Güth, W., R. Schmittberger, and B. Schwarze: "An experimental analysis of ultimatum bargaining," *Journal of Economic Behavior & Organization,* 3 (1982), pp. 367-388.

Hoffman, E., K. McCabe, and V. Smith: "Social Distance and Other-Regarding Behavior in Dictator Games", *American Economic Review* Vol. 86, No.3 (1996), pp. 653-660

Iriberri, N., and P. Rey-Biel: "The role of role uncertainty in modified dictator games", *Experimental Economics*, 14, (2011), pp. 160–180. DOI 10.1007/s10683-010-9261-5

Johnson, N. D., and A. A. Mislin: "Trust Games: A meta-analysis", *Journal of Economic Psychology* No. 32 (2011), pp. 865-889

Plott, C.R.: "Rational Individual Behavior in Markets and Social Choice Processes" *Social Science Working Paper*, 862, (1993) California Institute of Technology , Pasadena, CA.

Rabin, M.: Incorporating fairness into game theory and economics. *American Economic Review*, 83, (1993), pp. 1281–1302.

Smith, V.: "Rational Choice: The Contrast between Economics and Psychology," *Journal of Political Economy*, 99, no. 4, (1991), pp. 877-897.

Appendix

Instructions (treatment sb_B1_d0)

By participating in this experiment, you have the opportunity to earn money. Please be aware that throughout the whole session, you are neither allowed to communicate with other participants, nor to leave your seat. We ask you to read the instructions carefully. If you have any questions, please raise your hand and we will come to you in order to answer your questions. When all the participants have understood the instructions, we will start the experiment.

Experimental Procedure

You are matched randomly with a partner with whom you will play the same game for 20 periods. In each of these 20 rounds, both of you have to choose either a "1", or a "2". Your choices will be added up. If the sum of your choices is even (this is the case if both choose the same number), the player type "Even" will receive a point. If the sum is odd (this is the case if you choose differently), the player type "Odd" will win this round and gets a point. After each period, we will inform you about whether you have won.

Which type of player you are is random and will be displayed on your screen. The types do not change throughout the whole session.

In randomly chosen periods, one of the two players (who is also randomly chosen) will be forced to play a "Zero" and cannot make a choice. We inform the other player about this restriction and that she can therefore decide alone in this period. She can win the round by choosing the number accordingly (i.e. choosing a "2" if she is player type "Even", or choosing a "1" if she is player type "Odd"). Furthermore, there is also the possibility to voluntarily choose a "Zero" in these special periods. If the player chooses "Zero", this period will not be taken into account, but there will be an additional "normal" period in which both players can choose a number. We will not tell you in advance how many of these "special" periods will occur, but we will inform you about the last one to come.

To practice the procedure, there will first be 12 periods without monetary payoff, then follow the 20 periods described above. You will keep your partner throughout the whole session.

Payment Procedure

For each round you win, you will get a point. You will not lose a point when losing a period. The first part of the session, the 12 practice periods, will be played without receiving monetary payoff. In the second part, the 20 "real" periods, you will receive € 2.00 for each point you get. The money earned will be paid out at the end of the session. We ask you to confirm the receipt of the money by signing our form. After that, you may leave the laboratory.

Instructions (treatment db_B1_d0)

By participating in this experiment, you have the opportunity to earn money. Please be aware that throughout the whole session, you are neither allowed to communicate with other participants, nor to leave your seat. We ask you to read the instructions carefully. If you have any questions, please raise your hand and we will come to you in order to answer your questions. When all participants have understood the instructions, we will start the experiment.

Experimental Procedure

You are matched randomly with a partner with whom you will play the same game for 20 periods. In the beginning, you will see your partner but you are not allowed to talk to each other. In each of the 20 rounds, the both of you have to choose either a "1", or a "2". Your choices will be added up. If the sum of your choices is even (this is the case if both choose the same number), the player type "Even" will receive a point. If the sum is odd (this is the case if you choose differently), the player type "Odd" will win this round and gets a point. After each period, we will inform you about whether you have won.

Which type of player you are, is random and will be displayed on your screen. The types do not change throughout the whole session.

In randomly chosen periods, one of the two players (that is also randomly chosen) will be forced to play a "Zero" and cannot make a choice. We inform the other player about this restriction and that they can therefore decide alone in this period. She can win the round by choosing the number accordingly (i.e. choosing a "2" when being player type "Even", or choosing "1" being player type "Odd" respectively). Furthermore, there is also the possibility to voluntarily also choose a "Zero" in those special periods. If the player chooses the "Zero", this period will not be taken into account but there will be an additional "normal" period, where both players can choose a number. We will not tell you in advance, how many of those "special" periods will occur, but we inform you about the last of those periods to come.

For practicing the procedure, there will first be 12 periods without monetary payoff, then follow the 20 periods described above. You will keep your partners throughout the whole session.

Payment Procedure

For each round you win, you will get a point. You will not lose a point when losing a period. The first part of the session, the 12 practice periods, will be played without receiving monetary payoff. In the second part, the 20 "real" periods", you will receive € 2.00 for each point you get. The money earned will be paid out in the room next to lab at the end of the session. We ask you to bring your three-digit ID card for identification. This way, you stay completely anonymous throughout the whole session. No one can retrace your decisions.

Laundering the people instead of the money

An experimental study on the effect of mental money laundering

Alexandra Seidel *

Abstract

Since the introduction of mental accounting, research has consistently shown that the ethical source of money influences its use. People tend to donate unethically earned money to charitable causes instead of using it for personal consumption. Thus, to weaken the guilt associated with unethical gains, individuals may engage in "mental money laundering" to justify spending decisions. This study uses a laboratory experiment to test whether participants donate a higher share of their money to charity when donating alongside the person they deceived in a game, compared to donating with a new partner ("people laundering" effect). It also examines whether participants try to justify their unethical behavior by reporting the assumption that their partner would not donate anything of her wealth. Additionally, it examines the role of personality traits, including the Big Five and the Dark Triad, in influencing donation behavior. Results show the opposite effect of the hypothesized people laundering effect with individuals spending significantly more with the new partner. Also, there are significant differences in donation behavior based on personality traits, with high extraversion, neuroticism, and Machiavellianism scores correlating with lower donations.

Keywords

mental accounting, antisocial, deception, donation, experiment

JEL codes

C72, C91, D91

1. Motivation and hypotheses

"I have spent the best years of my life giving people the lighter pleasures, helping them have a good time, and all I get is abuse, the existence of a hunted man."
Al Capone (cited in Carnegie, 1936)

Al Capone is one of the most iconic mobsters of the 1920s and 1930s. He earned money with illegally sold alcohol in the speakeasies, prostitution, and gambling. He was also known for his violence and unscrupulousness mounting in the St. Valentine's massacre in 1929, although no violent crime could be proven against him (Bair, 2016; Berggreen, 1994; Kobler, 1971). While reading about Capone, positive stories often emerge alongside his mafia reputation. During the Great Depression in the 1930s, he opened soup kitchens in Chicago, feeding the poor but also donated money to charities, and supported local businesses (Berggreen, 1994; Kobler, 1971).
These two sides in the personality and work of Al Capone seem to be diametral. How could they coexist? Some authors argue that Capone, who enjoyed media attention, engaged in charitable acts to improve his public image and to conceal his crimes (Bair, 2016; Berggreen, 1994; Kobler, 1971). Undoubtedly, he spent some of his money on charitable purposes.
Starting with Thaler in 1985 and 1999, there is a vast amount of articles examining mental accounting and thus the violation of the principle of fungibility. Numerous studies found that indeed there is a utility difference of money obtained in a legal way and money obtained in an unethical/illegal way. The source of money matters (Fogel, 1997; Muehlbacher & Kirchler, 2019; Tversky & Kahneman, 1981). People tend to spend unethically earned money rather on other-regarding than on hedonic expenditures, thus punishing themselves (Levav & McGraw, 2009; Shalvi et al., 2015).
In 2021, Imas et al. published a paper on (Motivated) Mental Money Laundering (hereinafter referred to as MML) examining donation behavior of participants that obtained money in an unethical way. They found that these participants donated a higher share of their money than participants who have obtained their wealth via ethical behavior. One of the most surprising results in their first study was that the pure exchange of dollar bills (money that originated in unethical decisions) decouples the origin of money from the money itself, leading to significantly reduced donation amounts.
To extend this research, the goals of this article are (i) to examine whether also the exchange of a partner, participants can donate with will disentangle the money and its origin and (ii) to examine whether besides the external factors also personality traits influence donation behavior.
In their experiment, Imas et al. (2021) implemented the deception game by Erat and Gneezy (2012) where one player, the sender, can choose to lie to another player, the receiver, and thus

earn more money than by telling the truth. Senders were then given the possibility to donate an amount of their wealth. In this study, I extended this donation procedure: In the basis treatment, senders and receivers donated together as a group. In the switch treatment, senders were matched with a new partner that had not played the deception game before the donation stage but had completed an unrelated task in the meantime. Senders were then given the possibility to donate together with this new partner.

Thus, the baseline treatment is similar to previous studies (Gneezy et al., 2014; Imas et al., 2021). Thus, I argue that people who made an unethical decision beforehand will be more likely to donate more money to a charity (Levav & McGraw, 2009; Shalvi et al., 2015). This leads to my first hypothesis:

Hypothesis 1: The Guilt Hypothesis

Senders who have been lying successfully in the deception game donate relatively more than senders who told the truth.

In MML, Imas et al. created a "laundry" by implementing a lottery after the deception game. All participants were paid in cash directly after the deception game while still staying in the cabins. Then, there were two types of lottery the senders took part in: the *laundered* and the *unlaundered* one. With a high probability (p=0.83), the lottery returned the earnings of the deception game[1]. In the unlaundered lottery condition, the bills remained on the participants' desks until the lottery was finished. In the laundered lottery condition, the bills were removed and participants received other bills after the lottery ended. They found that participants in the laundered lottery condition donated significantly less. Thus, solely exposing the earned money to a risk had no influence on the senders' donation but the pure exchange of bills had.

My study extends these findings by giving participants the opportunity to change the *partner* they are donating with, fueling my second hypothesis:

Hypothesis 2: The People Laundering Hypothesis

Senders who have been lying successfully in the deception game donate less if being re-matched with a new person.

After the deception game and their own donation, all participants were asked to report their beliefs concerning the relative donation of their partner. Since most people strive to be morally good and belong to a community (Baumeister & Leary, 1995) lying disrupts this self-image, sparking ethical dissonance. To ease this tension, people often justify their actions, distancing

[1] With p=0.085 each, the earnings could also be doubled or halved.

themselves from immorality and demonizing others (Cialdini & Goldstein, 2004; Shalvi et al., 2015). This drives my third hypothesis regarding self-justification:

> ***Hypothesis 3: The Self-Justification Hypothesis***
>
> *Senders who have been lying successfully in the deception game perceive their partners as less willing to donate compared to senders who have told the truth.*

In the next section, I will explain the experimental design and treatments. Section three presents the results of the study followed by a discussion in section four and limitations in section five. The article closes with a conclusion in section six.

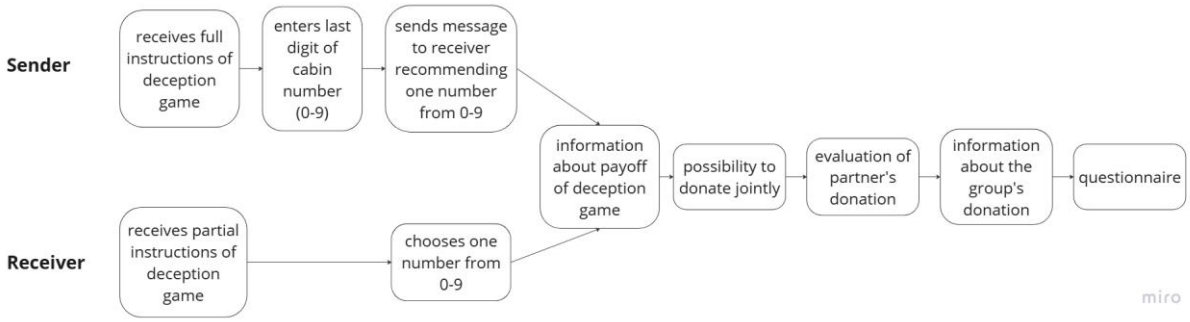
2. Experimental Design

In my experiment, I keep close to the first study in MML and will report the changes that were necessary to conduct this study.

2.1. Baseline

Figure 1 depicts the design of the baseline treatment.

Figure 1: Baseline treatment



Like in MML, participants were randomly matched in pairs with one sender and one receiver. They then played the deception game (Erat & Gneezy, 2012). This game has two possible outcomes: either, the receiver chooses the sender's secret number – she then receives a high payoff (option A) – or she chooses any other number and receives a lower payoff (option B). Thus, the receiver has an incentive to choose the sender's secret number. However, there is a steep information asymmetry as the sender only has full information about the payoff structure and his true secret number. Table 1 illustrates the payoff structure of both options.

Table 1: payoff scheme deception game

	Receiver chooses the sender's secret number (Option A)	**Receiver chooses any other number (Option B)**
Sender	€ 10	€ 25
Receiver	€ 25	€ 5

In the beginning of the experiment, the sender received this secret number by chance when drawing his cabin number. He was then asked to insert the last digit of his cabin number and received full information about the payoff structure and also learned that the receiver's only source of information will be the message the sender provided. When starting the deception game, the sender would choose 1 of 10 possible messages to send to the receiver. The messages would read: "If you choose 0 *(1,2,…,9)*, you will earn more money than with any other number." Thus, exactly one of ten possible messages was correct. The receiver then had to decide whether to follow this advice while she only had the sender's advice and no further information about the payoff structure. After the receiver chose one of the messages (and thus numbers), both players were informed about their personal payoff. Sticking to the deception game procedure of Erat & Gneezy (2012), receivers remained uninformed about the sender's payoff or other payoff options.

The players were subsequently informed that they could now donate a share of their earned money to the children's hospice in Magdeburg. However, in contrast to MML, *both* players could donate and they were told that their donation would be considered jointly, as a group. They could choose in 1€-steps from 0 to their complete wealth (payoff from the deception game and show-up fee of € 7.00)[2].

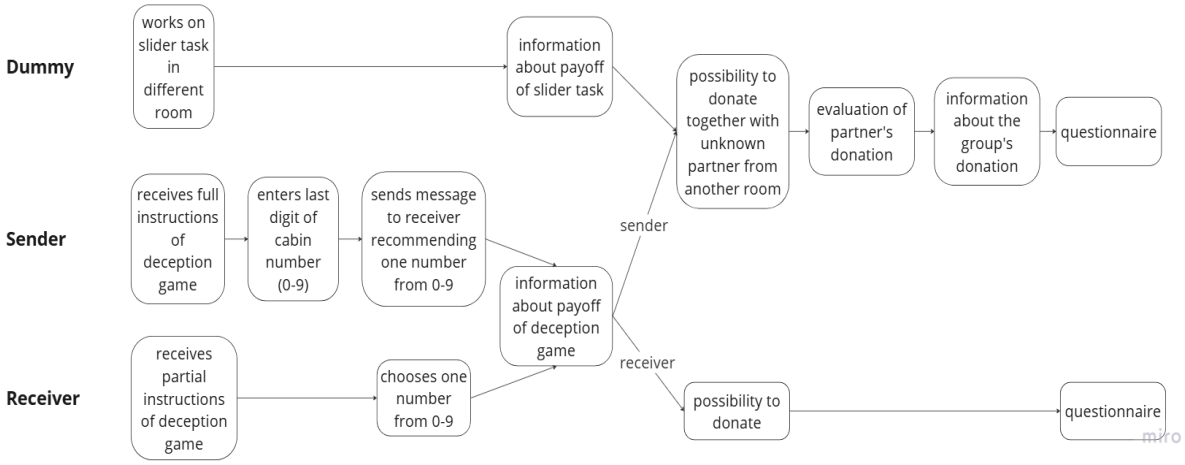
After that, both players were asked to estimate the relative amount their partner donated and were then informed about the group's joint donation. In the next step, participants were asked whether they usually volunteer or donate and to what extent they think the children's hospice in Magdeburg is an institution that should be financially supported. In the end, participants were asked to answer two personal traits questionnaires: the big five (Rammstedt et al., 2017) and the dark triad (Jonason & Webster, 2010).

[2] This was in contrast to MML who allowed donations up to $10. When asking Alex Imas, he told me that was done to restrict potential variance in the responses and decrease the range of potential noise and had no conceptually reason. (September 2023)

2.2. Switch

To allow the examination of the potential people laundering effect, the switch treatment deviates in one main aspect: there are three players. In the beginning, the participants were split into two groups and two rooms. In room 1, the deception game took place as described in the baseline while participants in room 2 (dummies) worked on an unrelated slider task. Their maximum payoff was €5.00 for completing this task. Neither the sender nor the dummy were informed about their partner's wealth. After the deception game and slider task were finished, the senders were matched with dummies from the other room before deciding on their donation. They were informed about the new matching and the group donation. The receivers could donate on their own. Figure 2 depicts the procedure of the switch treatment.

Figure 2: Switch treatment



2.3. Experimental Procedure

The experiment was programmed using *oTree* (Chen et al., 2016) and took place in the MaXLab, the Magdeburg Laboratory for economic experiments, in November 2023. All participants were recruited via *hroot* (Bock et al., 2014). Participants randomly drew a wooden sphere that depicted their cabin number and were asked to immediately go to their cabins. In the *switch*-treatment, the group was split up and went to different rooms. Before the sphere drawing process, participants were informed that the numbers led to different rooms[3]. Thus, they observed people entering both rooms proving that there were people in different rooms, decreasing skepticism towards this fact (Frohlich et al., 2001). They then participated in the experiment. In the end, participants were paid by a person who was not involved in the experiment and arrived only for the payoff procedure. All participants were informed about the

[3] I took care to ensure that the first room was never completely full to make it clear that this group separation did not happen due to a lack of space.

blinded payment procedure in the beginning of the session to avoid moral concerns towards the experimenter.

The average duration of a session was about 20 minutes with an average payoff of EUR 18.00, including a show-up fee of EUR 7.00.

3. Results

For a short overview, Table 2 shows the descriptive statistics. All participants were students of the Otto-von-Guericke University Magdeburg or the University of Applied Sciences Magdeburg-Stendal.

Table 2: descriptive statistics of the sample

	Baseline	**Switch**
number of subjects	98	150
number of groups	49	50
male/female/diverse	42/56/0	72/76/2
average age	24.6	23.9
volunteer	73	77

On the whole, 75 % of all senders lied and 89 % of all receivers followed the sender's advice in the deception game stage. Thus, there are four types of senders: those, who lied (1), those, who told the truth (2), those, who lied and failed (3), and those, who told the truth and failed (4). Table 3 shows the distribution of these types among the 99 senders.

Table 3: sender types

	N
lied	73
truth	23
lied *and failed*	1
truth *and failed*	2

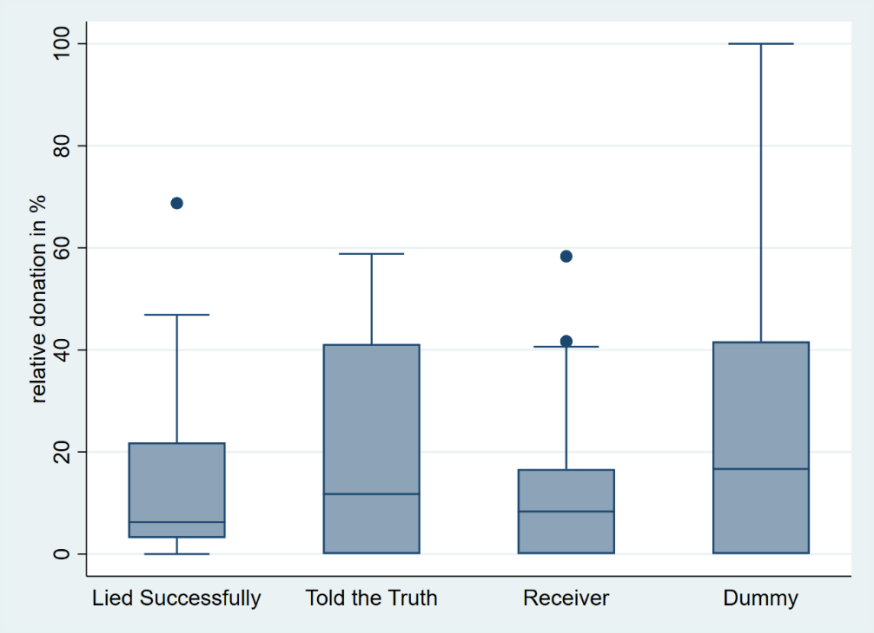
Since the number of observations for types (3) and (4) are very low, I will not take them into consideration and focus the results on senders' types (1) and (2).

84% of all lying senders and 74% of the truthful senders donated a share of their wealth (61% receivers, 74% dummies). In total, 71% of all participants donated to the children's hospice. The average donation was €2.69.

Like in MML, my variable of main interest is the participants' donation behavior. However, since I also compare donations of player roles with different wealth situations, my variable of

interest is the participant's *relative donation*. Figure 3 provides a first overview of the donation behavior of all player types in all treatments.

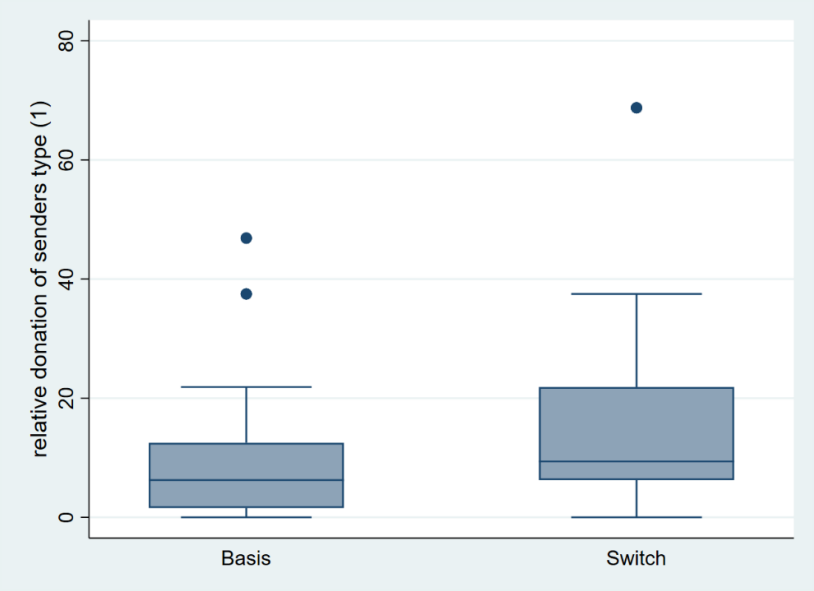
Figure 3: box and whiskers relative donation over all player types



First of all, when checking for the relative donation of both sender types conducting Mann-Whitney tests, there is no significant difference in donation behavior (p=0.111 basis and p=0.204 switch). Hence, there is no support for *Hypothesis 1*. Using the same test, I then also examined the relative donation of lying senders and receivers and also found no significant difference in the baseline treatment (p=0.449) and lying senders and dummies in the switch treatment (p=0.300).

Checking the people laundering hypothesis, I compared donation behavior of successfully lying senders in both treatments and find that they donate significantly (p=0.01, Mann-Whitney) *more* after being matched with a new person. Figure 4 depicts the relative donation by treatment.

Figure 4: relative donation of senders who lied successfully by treatment



The observed effect is statistically significant; however, it occurs in the direction opposite to that predicted by *Hypothesis 2.* Further investigating the result by conducting multiple tobit regressions (models (1) to (3)), I found that the treatment effect is robust. Including the BigFive (Rammstedt et al., 2017) in the regression, I see that the more extraverted ($p<0.05$) and neurotic ($p<0.1$) participants are, the fewer they donate. Adding the dark triad (Jonason & Webster, 2010), results indicate that also participants with Machiavellian preferences donate less ($p<0.1$). The effect of extraversion remains valid also when checking the relative donation for all player types (models (4) to (6)). However, when considering all player types collectively, neuroticism no longer shows a significant effect on donation behavior, whereas openness does. Specifically, individuals with high openness levels contribute significantly more ($p < 0.1$). The reference group for these models is the type 1 sender-group.

Being female and the participants' age had no significant influence. For volunteering and the hospice's worthiness to receive donations, a positive trend in relative donation is to be observed but the effect is not significant. Table 4 summarizes the regression results.

Table 4: summary of the tobit regression results

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	relative donation sender (1)	relative donation sender (1)	relative donation sender (1)	relative donation all playertypes	relative donation all playertypes	relative donation all playertypes
Switch	6.564**	5.500*	6.670**	2.062	1.412	1.870
	(3.195) [0.044]	*(3.168) [0.087]*	*(3.243) [0.044]*	*(3.047) [0.499]*	*(3.031) [0.642]*	*(3.041) [0.539]*
Sender type (2)				3.974	2.142	2.468
				(5.041) [0.431]	*(5.037) [0.671]*	*(5.032) [0.624]*
Receiver				-4.748	-5.106	-4.942
				(3.280) [0.149]	*(3.259) [0.119]*	*(3.247) [0.129]*
Dummy				7.952*	7.780*	7.538
				(4.122) [0.055]	*(4.073) [0.057]*	*(4.066) [0.065]*
Extraversion	-4.935**	-4.680**	-4.272**	-3.765**	-2.628*	-2.584
	(1.867) [0.010]	*(1.869) [0.015]*	*(1.895) [0.028]*	*(1.520) [0.014]*	*(1.551) [0.092]*	*(1.573) [0.102]*
Neuroticism	-3.376*	-2.905*	-3.181*	-2.350	-1.376	-1.747
	(1.705) [0.052]	*(1.704) [0.093]*	*(1.875) [0.095]*	*(1.469) [0.111]*	*(1.497) [0.359]*	*(1.616) [0.281]*
Openness	-0.277	-0.024	-0.138	2.138	2.363*	2.250*
	(1.458) [0.850]	*(1.448) [0.987]*	*(1.461) [0.925]*	*(1.346) [0.114]*	*(1.341) [0.079]*	*(1.340) [0.094]*
Conscientiousness	2.750	2.518	2.195	2.623	2.033	1.964
	(1.894) [0.151]	*(1.874) [0.184]*	*(1.876) [0.247]*	*(1.650) [0.113]*	*(1.673) [0.225]*	*(1.685) [0.245]*
Agreeableness	2.517	3.076	2.836	2.150	1.256	1.051
	(2.167) [0.250]	*(2.422) [0.209]*	*(2.436) [0.249]*	*(1.815) [0.237]*	*(1.949) [0.520]*	*(1.956) [0.592]*
Machiavellianism		-1.693	-1.997*		-1.9257*	-1.875*
		(1.048) [0.111]	*(1.058) [0.064]*		*(1.033) [0.059]*	*(1.031) [0.070]*
Psychopathy		1.602	1.590		1.122	1.264
		(1.364) [0.244]	*(1.361) [0.247]*		*(1.205) [0.353]*	*(1.230) [0.305]*
Narcissism		0.212	0.584		-1.173	-1.002
		(1.241) [0.865]	*(1.260) [0.645]*		*(1.002) [0.243]*	*(1.008) [0.321]*
Female			1.970			1.977
			(3.735) [0.600]			*(3.055) [0.518]*
Age			0.293			0.188
			(0.405) [0.472]			*(0.335) [0.575]*
Volunteer			0.744			-0.955
			(3.180) [0.816]			*(2.744) [0.728]*
Hospice			2.094			1.630
			(1.920) [0.280]			*(1.630) [0.318]*
Constant	16.662	13.847	-2.97	4.394	11.138	-0.716
	(12.684)	*(15.172)*	*(19.398)*	*(11.670)*	*(13.846)*	*(17.680)*
	[0.193]	*[0.365]*	*[0.879]*	*[0.708]*	*[0.422]*	*[0.968]*
Observations	73	73	73	245	245	245

Models (1) – (3) show tobit regression taking only lying senders into account, Models (4)-(6) show tobit regression regarding all playertypes with lying senders as a reference group
Robust standard errors in parentheses, p-values are in brackets.
*** p<0.01, ** p<0.05, * p<0.1

Testing the *self-justification hypothesis*, I find that 11% of the lying senders expected their partner not to donate anything. However, I find no significant difference in the donation behavior of these senders compared to the lying senders who expected that their partner donated something (p=0.186) and thus no support for *Hypothesis 3*. When checking the truthful senders,

I find that the 22% who expected their partner not to donate anything donated significantly less (p=0.000, Mann-Whitney) than those truthful senders that expected their partners to donate something.

Conducting a Mann-Whitney-test, I find no significant difference (p=0.978) in the expectations of the partner's donation behavior between these sender types. However, lying senders have a significant lower correlation (0.5419) between their own relative donation and their expectation about their partner's donation than truthful senders (0.8068) (p=0.044, Fisher-Z-test).

The regression results depicted in Table already showed an influence of personality traits on relative donation behavior. To examine a potential influence of personality traits on the probability of lying, I tested whether certain character traits are more distinct in one group of senders. Conducting a logistic regression, I found that the higher the agreeableness attribute, the greater the probability of being a sender that told the truth (p=0.012). Senders with a strong Machiavellian attribute were more likely to be senders that lied (p=0.011). Surprisingly, participants with a strong psychopathic attribute were more likely to tell the truth (p=0.082). Table 5 summarizes the results.

Table 5: logistic regression results for belonging to sendertype 2

VARIABLES	sendertype 2
Extraversion	-0.174 *(0.340) [0.608]*
Neuroticism	-0.165 *(0.305) [0.588]*
Openness	-0.025 *(0.233) [0.915]*
Conscientiousness	-0.253 *(0.351) [0.471]*
Agreeableness	1.087** *(0.431) [0.012]*
Machiavellianism	-0.547** *(0.215) [0.011]*
Constant	-2.391 *(2.385) [0.316]*
Observations	96

Robust standard errors in parentheses,
p-values are in brackets.
*** p<0.01, ** p<0.05, * p<0.1

4. Discussion

The aim of this study was (i) to extend the findings in MML by introducing joint donation and a third player that allowed a distance between the possibly unethical behavior and the donation and (ii) to examine whether besides the external factors also personality traits influence donation behavior.

Although having a smaller sample than Imas et al. (2021) my results of the deception game are quite similar. 75% of all senders lied (61% in MML), and 89% of all receivers followed the sender's advice (76% in MML).

However, in contrast to MML, I found no significant difference in the donation behavior between lying senders and senders who told the truth. One possible explanation for this finding is that lying senders might simply not have felt guilty and thus did not need to spend money selflessly to reduce their feeling of guilt (Levav & McGraw, 2009). Another is that senders gained a higher utility through the money earned than they experienced disutility because of lying (Bazerman & Gino, 2012).

In line with mental accounting research (Thaler, 1999) this study shows that participants who lied to enrich themselves behaved in a different way than those who didn't. However, it turned out to be the opposite way than expected. There is a significant treatment effect that shows that lying senders donate *less* if donating with the person they lied to in the deception game than with a new person. One explanation is that participants have not felt guilty and thus did not feel the need to compensate for their lies. It is also possible that – knowing that the receiver will be informed about her own, relatively small, revenue – lying senders refuse to donate more since a higher donation might signal a greater wealth and thus a possible lie. In the switch treatment, however, both participants had no information about the task the other player accomplished before donating so this fear might have vanished.

Although I found no support for the self-justification hypothesis concerning lying senders, I was surprised to find truthful senders donated significantly less when expecting their partners to contribute nothing. Given full knowledge of the payoff options, these senders knew about the distribution of payoffs. If the joint donation is understood as a public good (Samuelson, 1954), these senders might not want to contribute a substantial share of their wealth particularly if they believe their partners will not contribute at all, as this could evoke concerns about being exploited.

Besides the external factors that can influence donation behavior, internal factors, such as stable personality traits, can have impact. My first finding with respect to personality traits was that the more extraverted and neurotic participants are, the fewer they donate. This is not in line

with other studies, as e.g. Lim et al. (2021) reported that extraversion and conscientiousness were more likely to result in actual donations. Also, Yarkoni et al. (2015) reported a positive influence of extraversion, openness, and conscientiousness on perceived responsibility and likeability resulting in monetary donations.

Another finding indicates that participants with a high Machiavellian score donate less. According to Christie and Geis (1970) Machiavellian individuals are defined by their tendency for interpersonal manipulation, employing tactics like flattery and deception. They hold cynical, emotionally detached views and prioritize personal gain over traditional moral standards, often choosing strategies purely to advance their self-interest (Bereczkei et al., 2010; Christie & Geis, 1970; Fehr & Paulhus, 1992; Wilson et al., 1998). Bereczkei et al. (2010) found that Machiavellian individuals act altruistically in public but prioritize self-interest privately, unaffected by charity cost. Given that donations in my study were anonymous, this aligns well with the observed behaviors.

Regarding openness as personality trait, I found that compared to the reference group (lying senders), other playertypes with a high degree of openness donate significantly more ($p<0.1$). This is supported by other studies. Kline et al. (2019) examined the link between personality traits and prosocial actions, finding that agreeableness and openness were significantly and positively related to prosocial behavior, while other Big Five traits showed no such association. Similarly, Zhang et al. (2023) observed a positive effect of openness on household donations.

Further research (e.g. Carlo et al., 2005; Claxton-Oldfield & Banzen, 2010; Jensen-Campbell et al., 2002) reported that high conscientiousness and agreeableness correlate with increased volunteering and monetary donations. In my study, I observed positive relationships were not statistically significant.

Furthermore, I investigated a possible interaction of personality traits and the probability of senders to lie or tell the truth. I found that the higher the level of agreeableness, the greater the probability of being a truthful sender. One possible explanation for this finding is that agreeableness is described as being forgiving, nonjudgmental, altruistic, trusting, cooperative, and willing to compromise (Lee & Ashton, 2004; McCrae & Costa, 1999). Furthermore, this is supported by Sarzyńska et al. (2017) who found that people being low on agreeableness were most likely to lie.

Typically, Machiavellianism is linked to strategic lies for self-gain or social manipulation, narcissism to lies for self-image enhancement, and psychopathy to impulsive, purposeless lies (Jonason et al., 2014; Kashy & DePaulo, 1996; McLeod & Genereux, 2008). It is not surprising, that I found that senders with Machiavellian tendencies were more likely to lie, aligning with

prior studies showing that individuals with high scores in "Dark Triad" traits often engage in deception (Paulhus & Williams, 2002). Notably, however, this study found that participants with strong psychopathic traits were more likely to tell the truth, which contrasts with the usual tendency for people having a high level of psychopathy to lie both to themselves and others (Gudjonsson & Sigurdsson, 2004).

5. Conclusion

This study extends the mental money laundering framework by examining the effect of the exchange of the partner to donate with and exploring the influence of personality traits on donation behavior. While I found no support for the hypothesis that lying senders donated more than honest ones, the switch treatment revealed a significant increase in donations when lying senders were paired with a new partner, indicating that there might be other drivers for this decision than the pure exchange of partners such as a high donation might signal dishonest behavior in the deception game. Results on the self-justification aspect shed light on possible concerns of being exploited when contributing to the joint donation. Furthermore, personality traits like extraversion, neuroticism, and Machiavellian tendencies demonstrated significant negative influence on donation behavior, underscoring the role of stable personality factors in shaping ethical and prosocial actions. These findings highlight that both social and internal factors play critical roles in shaping how individuals reconcile unethical behavior through monetary choices. With this experiment, donations of 851 EUR were collected and given to the children’s hospice Magdeburg in February 2024.

Acknowledgements

I thank Joachim Weimann, Abdolkarim Sadrieh, Ronnie Schöb, and Dmitri Bershadskyy for helpful comments, Jannik Greif for his assistance in conducting the experiment, and Franziska Rumpel for numerous discussions on this project. I am also grateful that the Faculty of Management and Economics of the Otto-von-Guericke University Magdeburg funded his research study.

References

Bair, D. (2016). *Al Capone: His life, legacy, and legend*. Knopf Doubleday Publishing Group.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bazerman, M. H., & Gino, F. (2012). Behavioral Ethics: Toward a Deeper Understanding of Moral Judgment and Dishonesty. *Annual Review of Law and Social Science*, *8*(1), 85–104. https://doi.org/10.1146/annurev-lawsocsci-102811-173815

Bereczkei, T., Birkas, B., & Kerekes, Z. (2010). The Presence of Others, Prosocial Traits, Machiavellianism. *Social Psychology*, *41*(4), 238–245. https://doi.org/10.1027/1864-9335/a000032

Berggreen, L. (1994). *Capone: The man and the era*. Simon & Schuster.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review*, *71*, 117–120. https://doi.org/10.1016/j.euroecorev.2014.07.003

Carlo, G., Okun, M. A., Knight, G. P., & Guzman, M. R. T. de (2005). The interplay of traits and motives on volunteering: agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences*, *38*(6), 1293–1305. https://doi.org/10.1016/j.paid.2004.08.012

Carnegie, D. (1936). *How to win friends and influence people*. Simon & Schuster.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Christie, R., & Geis, F. L. (1970). *Studies in Machiavellianism*. Academic Press.

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591–621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Claxton-Oldfield, S., & Banzen, Y. (2010). Personality characteristics of hospice palliative care volunteers: The "big five" and empathy. *The American Journal of Hospice & Palliative Care*, *27*(6), 407–412. https://doi.org/10.1177/1049909110364017

Erat, S., & Gneezy, U. (2012). White Lies. *Management Science*, *58*(4), 723–733. https://doi.org/10.1287/mnsc.1110.1449

Fehr, B., & Paulhus, D. L. (1992). The Construct of Machiavellianism: Twenty Years Later. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in Personality Assessment* (pp. 77–116). Hillsdale.

Fogel, S. (1997). Income Source Effects. *Unpublished Manuscript*.

Frohlich, N., Oppenheimer, J., & Moore, J. B. (2001). Some Doubts about Measuring Self-Interest Using Dictator Experiments: The Cost of Anonymity. *Journal of Economic Behavior and Organization*, *46*(3), 271–290.

Gneezy, U., Imas, A., & Madarász, K. (2014). Conscience Accounting: Emotion Dynamics and Social Behavior. *Management Science*, *60*(11), 2645–2658. https://doi.org/10.1287/mnsc.2014.1942

Gudjonsson, G. H., & Sigurdsson, J. F. (2004). The relationship of suggestibility and compliance with self-deception and other-deception. *Psychology, Crime & Law*, *10*(4), 447–453. https://doi.org/10.1080/10683160310001634278

Imas, A., Loewenstein, G., & Morewedge, C. K. (2021). Mental Money Laundering: A Motivated Violation of Fungibility. *Journal of the European Economic Association*, *19*(4), 2209–2233. https://doi.org/10.1093/jeea/jvaa059

Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furdella, J. Q., & Egan, S. K. (2002). Agreeableness, Extraversion, and Peer Relations in Early Adolescence: Winning Friends and Deflecting Aggression. *Journal of Research in Personality*, *36*(3), 224–251. https://doi.org/10.1006/jrpe.2002.2348

Jonason, P. K., Lyons, M., Baughman, H. M., & Vernon, P. A. (2014). What a tangled web we weave: The Dark Triad traits and deception. *Personality and Individual Differences*, *70*, 117–119. https://doi.org/10.1016/j.paid.2014.06.038

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/a0019265

Kashy, D. A., & DePaulo, B. M. (1996). Who lies? *Journal of Personality and Social Psychology*, *70*(5), 1037–1051.

Kline, R., Bankert, A., Levitan, L., & Kraft, P. (2019). Personality and Prosocial Behavior: A Multilevel Meta-Analysis. *Political Science Research and Methods*, *7*(1), 125–142. https://doi.org/10.1017/psrm.2017.14

Kobler, J. (1971). *Capone: The life and world of Al Capone*. Hachette Books.

Lee, K., & Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, *39*(2), 329–358. https://doi.org/10.1207/s15327906mbr3902_8

Levav, J., & McGraw, A. P. (2009). Emotional Accounting: How Feelings About Money Influence Consumer Choice. *Journal of Marketing Research*, *46*(1), 66–80. https://doi.org/10.1509/jmkr.46.1.66

Lim, H. S., Bouchacourt, L., & Brown-Devlin, N. (2021). Nonprofit organization advertising on social media: The role of personality, advertising appeals, and bandwagon effects. *Journal of Consumer Behaviour*, *20*(4), 849–861. https://doi.org/10.1002/cb.1898

McCrae, R. R., & Costa, P. T. (1999). A Five-Factor Theory of Personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality: Theory and research.* Guilford.

McLeod, B. A., & Genereux, R. L. (2008). Predicting the acceptability and likelihood of lying: The interaction of personality with type of lie. *Personality and Individual Differences*, *45*(7), 591–596. https://doi.org/10.1016/j.paid.2008.06.015

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*(1), 33–43. https://doi.org/10.1037//0022-3514.81.1.33

Muehlbacher, S., & Kirchler, E. (2019). Individual Differences in Mental Accounting. *Frontiers in Psychology*, *10*, 2866. https://doi.org/10.3389/fpsyg.2019.02866

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2017). A Short Scale for Assessing the Big Five Dimensions of Personality: 10 Item Big Five Inventory (BFI-10). Advance online publication. https://doi.org/10.12758/mda.2013.013 (17 Pages / methods, data, analyses, Vol 7, No 2 (2013)).

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, *20*(4), 523–528. https://doi.org/10.1111/j.1467-9280.2009.02326.x

Samuelson, P. A. (1954). The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*, *36*(4), 387. https://doi.org/10.2307/1925895

Sarzyńska, J., Falkiewicz, M., Riegel, M., Babula, J., Margulies, D. S., Nęcka, E., Grabowska, A., & Szatkowska, I. (2017). More intelligent extraverts are more likely to deceive. *PloS One*, *12*(4), e0176591. https://doi.org/10.1371/journal.pone.0176591

Shalvi, S., Gino, F., Barkan, R., & Ayal, S. (2015). Self-Serving Justifications: Doing Wrong and Feeling Moral. *Current Directions in Psychological Science*, *24*(2), 125–130. https://doi.org/10.1177/0963721414553264

Thaler, R. H. (1985). Mental Accounting and Consumer Choice. *Marketing Science*, *4*(3), 199–214. https://doi.org/10.1287/mksc.4.3.199

Thaler, R. H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making*, *12*(3), 183–206. https://doi.org/10.1002/(SICI)1099-0771(199909)12:3<183::AID-BDM318>3.0.CO;2-F

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science (New York, N.Y.)*, *211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Wilson, D., Near, D. C., & Miller, R. R. (1998). Individual Differences in Machiavellianism as a Mix of Cooperative and Exploitative Strategies. *Evolution and Human Behavior*, *19*(3), 203–212. https://doi.org/10.1016/S1090-5138(98)00011-7

Yarkoni, T., Ashar, Y. K., & Wager, T. D. (2015). Interactions between donor Agreeableness and recipient characteristics in predicting charitable donation and positive social evaluation. *PeerJ*, *3*, e1089. https://doi.org/10.7717/peerj.1089

Zhang, K., Cao, B., Zhang, Y., & Han, Y. (2023). A study on the influence of personality characteristics on household charitable donation behavior in China. *PloS One*, *18*(5), e0284798. https://doi.org/10.1371/journal.pone.0284798

Appendix

Instruktionen Sender: Basis

Herzlich Willkommen zu unserem Experiment.

Durch die Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen.

Ihr Auszahlungsbetrag hängt von Ihren und den Entscheidungen der Person ab, der Sie zufällig zugeordnet werden. Sie werden nicht erfahren, mit wem sie spielen und auch Ihr Partner wird Ihre Identität nicht erfahren.

Sie werden das gesamte Experiment mit **derselben Person** durchführen.

Bitte lesen Sie sich die Instruktionen aufmerksam durch.

Wenn Sie Fragen haben, melden Sie sich bitte. Die Experimentatorin wird dann zu Ihnen kommen und die Fragen beantworten.

Das Experiment beginnt erst, wenn alle die Instruktionen verstanden haben.

Im Kontext dieses Experiments sind Sie in der Rolle des **Senders** und die andere Person in der Rolle des **Empfängers**.

Bitte tragen Sie hier die letzte Ziffer Ihrer Kabinennummer ein (sie steht auf der Holzkugel, die Sie zu Beginn gezogen haben): _____

Das ist Ihre geheime Ziffer.

Der Empfänger kennt weder Sie noch Ihre geheime Ziffer, sondern erhält ausschließlich die Information, die aus Ihrer Nachricht hervorgeht.

Die Nachricht, die Sie senden, wird eine Ziffer zwischen 0 und 9 enthalten. Sie wird dem Empfänger gezeigt und dieser wählt daraufhin wiederum eine Zahl zwischen 0 und 9.

Diese Entscheidung des Empfängers wird die Höhe Ihrer beider Auszahlungen bestimmen.

> Option A: Wählt der Empfänger Ihre geheime Ziffer, erhalten Sie **10 Euro** und die Person erhält **25 Euro.**
>
> Option B: Wählt der Empfänger irgendeine andere Ziffer (NICHT Ihre geheime Ziffer), erhalten Sie **25 Euro** und die Person erhält **5 Euro**.

Obacht! Der Empfänger hat keinerlei Informationen über die möglichen Auszahlungen. Er erhält ausschließlich Ihre übermittelte Nachricht und entscheidet daraufhin.

Der Empfänger wird auch im Nachgang nicht über Ihre Auszahlung informiert.

<<Button>>: Ich habe die Instruktionen verstanden

DECEPTION GAME

Bitte wählen Sie Ihre Nachricht.

Ihre geheime Ziffer lautet: {{ **player.instruction_id** }}

Bedenken Sie: Wenn der Empfänger diese Zahl wählt, erhalten Sie **10 Euro** und die Person erhält **25 Euro**.

Wählt der Empfänger eine andere Ziffer als die Ihre, erhalten Sie **25 Euro** und die Person erhält **5 Euro**.
Bitte wählen Sie aus den folgenden Nachrichten die aus, die an den Empfänger geschickt werden soll:

- Wenn Du Nummer **0** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **1** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **2** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **3** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **4** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **5** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **6** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **7** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **8** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **9** wählst, verdienst Du mehr als mit jeder anderen Zahl.

<<Button>>: Abschicken

ERGEBNIS

Ihre Auszahlung beträgt **XX EUR**. Zusätzlich erhalten Sie **7,00 EUR** für Ihre Teilnahme am Experiment.

<<Button>>: Weiter

SPENDE

Sie haben die Möglichkeit, einen beliebigen Teil der Ihnen zur Verfügung stehenden Summe (Ihre Auszahlung und die 7 Euro für Ihr Erscheinen) an das Kinderhospiz Magdeburg zu spenden.
Dabei entscheiden der Empfänger und Sie individuell über die Höhe der Spende, der Spendenbeitrag wird aber summiert und als Ganzes gespendet.

Wie viel möchten Sie an das Kinderhospiz Magdeburg spenden? _____

<<Button>>: Weiter

FRAGE

Wie schätzen Sie Ihren Partner ein? Wie viel Prozent seiner Auszahlung hat er gespendet? (drop down menu)

<<Button>>: Weiter

Instruktionen Receiver: Basis

Herzlich Willkommen zu unserem Experiment.

Durch die Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen.

Ihr Auszahlungsbetrag hängt von Ihren und den Entscheidungen der Person ab, der Sie zufällig zugeordnet werden. Sie werden nicht erfahren, mit wem sie spielen und auch Ihr Partner wird Ihre

Sie werden das gesamte Experiment mit **derselben Person** durchführen.

Bitte lesen Sie sich die Instruktionen **aufmerksam** durch.

Wenn Sie Fragen haben, melden Sie sich bitte.

Die Experimentatorin wird dann zu Ihnen kommen und die Fragen beantworten.

Das Experiment beginnt erst, wenn alle die Instruktionen verstanden haben.

Im Kontext dieses Experiments sind Sie in der Rolle des **Empfängers** und die andere Person in der Rolle des **Senders.**

Es gibt zwei Auszahlungsoptionen, die nur dem Sender bekannt sind. Mit Ihrer Auswahl aber entscheiden Sie, welche der beiden Optionen gewählt wird. Dabei haben Sie als Information nur die Nachricht, die der Sender Ihnen im nächsten Schritt schicken wird. So läuft es ab:

Zu Beginn des Experiments hat der Sender eine nur ihm bekannte Ziffer erhalten. Er wird Ihnen eine Nachricht schicken, die wiederum eine Ziffer zwischen 0 und 9 enthält.

Schlussendlich wählen Sie eine Ziffer zwischen 0 und 9 aus, wobei Sie entscheiden können, ob Sie dem Vorschlag des Senders folgen oder nicht.

Es gilt: Wenn die Ziffer, die Sie auswählen, der Ziffer entspricht, die der Sender zu Beginn erhalten hat, werden Sie beide entsprechend der einen Auszahlungsoption ausgezahlt, sonst erhalten Sie Ihre Auszahlung entsprechend der anderen Option.

Obacht! Sie werden nicht erfahren, wie die Auszahlungssummen verteilt sind oder welche Optionen zur Verfügung standen.

<<Button>>: Ich habe die Instruktionen verstanden

SIE HABEN EINE NACHRICHT ERHALTEN

Der Sender hat Ihnen folgende Nachricht zukommen lassen:

{{ message_sender }}

Bitte wählen Sie nun eine Ziffer zwischen 0 und 9. (drop down menu)

<<Button>>: Abschicken

ERGEBNIS

Ihre Auszahlung beträgt **XX EUR**. Zusätzlich erhalten Sie **7,00 EUR** für Ihre Teilnahme am Experiment.

<<Button>>: Weiter

SPENDE

Sie haben die Möglichkeit, einen beliebigen Teil der Ihnen zur Verfügung stehenden Summe (Ihre Auszahlung und die 7 Euro für Ihr Erscheinen) an das Kinderhospiz Magdeburg zu spenden.

Dabei entscheiden der Sender und Sie individuell über die Höhe der Spende, der Spendenbeitrag wird aber summiert und als Ganzes gespendet.

Wie viel möchten Sie an das Kinderhospiz Magdeburg spenden? _____

<<Button>>: Weiter

FRAGE

Wie schätzen Sie Ihren Partner ein? Wie viel Prozent seiner Auszahlung hat er gespendet? (drop down menu)

<<Button>>: Weiter

Instruktionen Sender: Switch

Herzlich Willkommen zu unserem Experiment.

Durch die Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen.

Ihr Auszahlungsbetrag hängt von Ihren und den Entscheidungen der Person ab, der Sie zufällig zugeordnet werden. Sie werden nicht erfahren, mit wem sie spielen und auch Ihr Partner wird Ihre Identität nicht erfahren.

Bitte lesen Sie sich die Instruktionen aufmerksam durch.

Wenn Sie Fragen haben, melden Sie sich bitte. Die Experimentatorin wird dann zu Ihnen kommen und die Fragen beantworten.

Das Experiment beginnt erst, wenn alle die Instruktionen verstanden haben.

Im Kontext dieses Experiments sind Sie in der Rolle des **Senders** und die andere Person in der Rolle des **Empfängers**.

Bitte tragen Sie hier die letzte Ziffer Ihrer Kabinennummer ein (sie steht auf der Holzkugel, die Sie zu Beginn gezogen haben): _____

Das ist Ihre geheime Ziffer.

Der Empfänger kennt weder Sie noch Ihre geheime Ziffer, sondern erhält ausschließlich die Information, die aus Ihrer Nachricht hervorgeht.

Die Nachricht, die Sie senden, wird eine Ziffer zwischen 0 und 9 enthalten. Sie wird dem Empfänger gezeigt und dieser wählt daraufhin wiederum eine Zahl zwischen 0 und 9.

Diese Entscheidung des Empfängers wird die Höhe Ihrer beider Auszahlungen bestimmen.

> Option A: Wählt der Empfänger Ihre geheime Ziffer, erhalten Sie **10 Euro** und die Person erhält **25 Euro.**
>
> Option B: Wählt der Empfänger irgendeine andere Ziffer (NICHT Ihre geheime Ziffer), erhalten Sie **25 Euro** und die Person erhält **5 Euro**.

Obacht! Der Empfänger hat keinerlei Informationen über die möglichen Auszahlungen. Er erhält ausschließlich Ihre übermittelte Nachricht und entscheidet daraufhin.

Der Empfänger wird auch im Nachgang nicht über Ihre Auszahlung informiert.

<<Button>>: Ich habe die Instruktionen verstanden

DECEPTION GAME

Bitte wählen Sie Ihre Nachricht.

Ihre geheime Ziffer lautet: {{ **player.instruction_id** }}

Bedenken Sie: Wenn der Empfänger diese Zahl wählt, erhalten Sie **10 Euro** und die Person erhält **25 Euro**.

Wählt der Empfänger eine andere Ziffer als die Ihre, erhalten Sie **25 Euro** und die Person erhält **5 Euro**.

Bitte wählen Sie aus den folgenden Nachrichten die aus, die an den Empfänger geschickt werden soll:

- Wenn Du Nummer **0** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **1** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **2** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **3** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **4** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **5** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **6** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **7** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **8** wählst, verdienst Du mehr als mit jeder anderen Zahl.
- Wenn Du Nummer **9** wählst, verdienst Du mehr als mit jeder anderen Zahl.

<<Button>>: Abschicken

ERGEBNIS

Ihre Auszahlung beträgt **XX EUR**. Zusätzlich erhalten Sie **7,00 EUR** für Ihre Teilnahme am Experiment.

<<Button>>: Weiter

SPENDE

Für den nachfolgenden Teil des Experiments sind Sie mit einem anderen Spieler verbunden als bisher. Dieser Spieler sitzt im Nebenraum und hat bisher eine völlig andere Aufgabe gelöst. Sie haben die Möglichkeit, einen beliebigen Teil der Ihnen zur Verfügung stehenden Summe (Ihre Auszahlung und die 7 Euro für Ihr Erscheinen) an das Kinderhospiz Magdeburg zu spenden.

Dabei entscheiden der Empfänger und Sie individuell über die Höhe der Spende, der Spendenbeitrag wird aber summiert und als Ganzes gespendet.

Wie viel möchten Sie an das Kinderhospiz Magdeburg spenden? _____

<<Button>>: Weiter

FRAGE

Wie schätzen Sie Ihren Partner ein? Wie viel Prozent seiner Auszahlung hat er gespendet? (drop down menu)

<<Button>>: Weiter

Instruktionen Receiver: Switch

Herzlich Willkommen zu unserem Experiment.

Durch die Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen.

Ihr Auszahlungsbetrag hängt von Ihren und den Entscheidungen der Person ab, der Sie zufällig zugeordnet werden. Sie werden nicht erfahren, mit wem sie spielen und auch Ihr Partner wird Ihre

Bitte lesen Sie sich die Instruktionen **aufmerksam** durch.

Wenn Sie Fragen haben, melden Sie sich bitte.

Die Experimentatorin wird dann zu Ihnen kommen und die Fragen beantworten.

Das Experiment beginnt erst, wenn alle die Instruktionen verstanden haben.

Im Kontext dieses Experiments sind Sie in der Rolle des **Empfängers** und die andere Person in der Rolle des **Senders.**

Es gibt zwei Auszahlungsoptionen, die nur dem Sender bekannt sind. Mit Ihrer Auswahl aber entscheiden Sie, welche der beiden Optionen gewählt wird. Dabei haben Sie als Information nur die Nachricht, die der Sender Ihnen im nächsten Schritt schicken wird. So läuft es ab:

Zu Beginn des Experiments hat der Sender eine nur ihm bekannte Ziffer erhalten. Er wird Ihnen eine Nachricht schicken, die wiederum eine Ziffer zwischen 0 und 9 enthält.

Schlussendlich wählen Sie eine Ziffer zwischen 0 und 9 aus, wobei Sie entscheiden können, ob Sie dem Vorschlag des Senders folgen oder nicht.

Es gilt: Wenn die Ziffer, die Sie auswählen, der Ziffer entspricht, die der Sender zu Beginn erhalten hat, werden Sie beide entsprechend der einen Auszahlungsoption ausgezahlt, sonst erhalten Sie Ihre Auszahlung entsprechend der anderen Option.

Obacht! Sie werden nicht erfahren, wie die Auszahlungssummen verteilt sind oder welche Optionen zur Verfügung standen.

<<Button>>: Ich habe die Instruktionen verstanden

SIE HABEN EINE NACHRICHT ERHALTEN

Der Sender hat Ihnen folgende Nachricht zukommen lassen:

{{ message_sender }}

Bitte wählen Sie nun eine Ziffer zwischen 0 und 9. (drop down menu)

<<Button>>: Abschicken

ERGEBNIS

Ihre Auszahlung beträgt **XX EUR**. Zusätzlich erhalten Sie **7,00 EUR** für Ihre Teilnahme am Experiment.

<<Button>>: Weiter

SPENDE

Sie haben die Möglichkeit, einen beliebigen Teil der Ihnen zur Verfügung stehenden Summe (Ihre Auszahlung und die 7 Euro für Ihr Erscheinen) an das Kinderhospiz Magdeburg zu spenden.

Wie viel möchten Sie an das Kinderhospiz Magdeburg spenden? _____

<<Button>>: Weiter

FRAGE

Wie schätzen Sie Ihren Partner ein? Wie viel Prozent seiner Auszahlung hat er gespendet? (drop down menu)

<<Button>>: Weiter

Instruktionen Dummy

Herzlich Willkommen zu unserem Experiment.

Durch die Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen.

Wir bitten Sie, Schieberegler der Aufgabe entsprechend zu positionieren. Für jeden korrekt positionierten Schieberegler erhalten Sie 0,50 EUR.

<<Button>>: Ich habe die Instruktionen verstanden

ERGEBNIS

Sie haben _____ Schieberegler richtig positioniert. Ihre Auszahlung beträgt **XX,XX** EUR.

<<Button>>: Weiter

SPENDE

Sie haben die Möglichkeit, einen beliebigen Teil der Ihnen zur Verfügung stehenden Summe (Ihre Auszahlung und die 7 Euro für Ihr Erscheinen) an das Kinderhospiz Magdeburg zu spenden.

Dabei entscheiden der Sender und Sie individuell über die Höhe der Spende, der Spendenbeitrag wird aber summiert und als Ganzes gespendet.

Wie viel möchten Sie an das Kinderhospiz Magdeburg spenden? _____

<<Button>>: Weiter

FRAGE

Wie schätzen Sie Ihren Partner ein? Wie viel Prozent seiner Auszahlung hat er gespendet? (drop down menu)

<<Button>>: Weiter

The Cinderella Game

- Finding the ones who will not go to the ball – an experimental study –

Alexandra Seidel, Jannik Greif, Franziska Rumpel, Abdolkarim Sadrieh

Abstract

The exclusion of an individual from group benefits can be damaging in numerous ways, including physiological and psychological dimensions. The central objective of the study is to understand the selection criteria used for exclusion decisions. We also investigate whether the information on being excluded by others amplifies or abates the likelihood of being selected for exclusion. In our online study, we invited a set of subjects about whom we collected information in multiple dimensions, including behavioral characteristics and appearance. Using the distributions of the attributes, we created fictional person cards and asked the subjects to rank them according to the expected likelihood that they will be excluded by their group. We incentivized the true ranking elicitation by paying subjects payoffs that increased with the similarity to the group's overall ranking. In the treatment, we provided information about some cards being among the top 5 of an earlier experiment's overall ranking in order to examine whether subjects follow this anchor.

Keywords

anti-social preferences, organizational behavior, social preferences, norm elicitation

JEL codes

C91, C92, C83, D91

1. Motivation and hypotheses

Once upon a time, Cinderella experienced exclusion by her stepmother and stepsisters. She was the one to take care of the household, slept in the ashes, sorted peas but was not allowed to have fancy dresses or go to the prince's ball. In the end of the fairy tale by the brothers Grimm, the prince and Cinderella lived happily ever after – however this article's main interest is not the happy end but to focus in the exclusion phase.

In general, humans are sociable beings that have a need to belong somewhere or to somebody (Baumeister & Leary, 1995). Despite efforts to gain acceptance, social rejection remains a common part of life (Kurzban & Leary, 2001) and can be found in everyday life in schools, work places and e.g. on social media, people experience exclusion that may cause psychological and physical damage (Z. Chen et al., 2008; DeWall & Baumeister, 2006; Eisenberger & Lieberman, 2004; Williams et al., 2000). Through stigmatization, certain people are systematically excluded from specific social interactions due to a particular characteristic they possess or group they belong to (Kurzban & Leary, 2001). Goffman (1963) describes stigmatization as a process where an individual is globally devalued for possessing a trait seen as deviant. Stigma arises in social interactions when a person's actual attributes (their "real" social identity) conflict with society's expectations of the characteristics they "should" possess (their "virtual" social identity). Crocker, Major, and Steele (1998) pointed out the challenge in pinpointing a single core aspect of stigma, proposing instead that stigmatized individuals are perceived as having "some attribute, or characteristic, that conveys a social identity that is devalued in a particular social context" (p. 505). Furthermore, various scientific disciplines generally agree that social exclusion has been a fundamental aspect of human societies across cultures and throughout history (Boehm, 1986; Gruter & Masters, 1986; Williams, 1997; Zippelius, 1986). In economic research, there are studies focusing on potential drivers of exclusion behavior like e.g. focal points, incentives, communication, or group affiliation (e.g. Abbink & Doğan, 2019; Bershadskyy & Seidel, 2024; Goette et al., 2006).

In this article, we want to discover potential focal points that lead to excluding behavior. Why are people excluded? A first glance at this question raises two possible answers: people might be excluded because of (i) *visual*[1] attributes, they aren't initially responsible for like gender, hair color, face shape, or body shape (that might result of illnesses etc.) and (ii) the *behavior* they show like being vengeful, intraverted, a couch potato, or little trusting, i.e. attributes, they *are* initially responsible for. Using typical characteristics, we asked participants to rank fictional

[1] In this paper, we use the words "visual" and "optical" synonymously.

personas according to their likelihood to be excluded by a group. All participants received a message after the first ranking and could change it once. The closer the individual ranking was to the group's ranking, the higher her payoff. Hence, our paper has two dimensions. On the one hand, we examine potential patterns of exclusive characteristics. On the other hand, we test what will happen when exposing the excluded ones.

Research outside of economic fields shows that certain traits increase the risk of social exclusion. Specifically, there is substantial evidence linking overweight status with higher chances of being excluded (Cahnman, 1968; Crocker et al., 1993; DeJong, 1980; Heatherton et al., 2000; Neumark-Sztainer et al., 1998; Puhl & Heuer, 2009; Rudolph et al., 2009) or negative behavioral traits (Brauer & Chekroun, 2005; Cuddy et al., 2007; Neuberg et al., 2000; Wesselmann et al., 2013). Using a conjoint analysis, we aim to find patterns driving excluding behavior in our sample.

Furthermore, various studies showed, that people generally feel a stronger connection to in-group members, favoring them over out-group members (Hewstone et al., 2002). Out-group members are offered less help (Levine et al., 2002), are less trusted (Voci, 2006), and receive fewer resources (Tajfel et al., 1971). Hence, group membership influences social exclusion dynamics. In addition, people tend to exclude others that are outside their group (Killen et al., 2013), i.e. possibly having other characteristics than themselves or just not being part of the in-group (Lelieveld et al., 2020). However, in our setting, participants are asked to create a ranking depicting their beliefs about their *group's* behavior, creating a social distance towards their own decisions and incentivizing the disclosure of their true beliefs (Krupka & Weber, 2013). Following this, our first hypothesis is the following:

Hypothesis 1:

Participants will also exclude cards with characteristics that meet their own ones.

There is literature showing that once an individual has been excluded by the group, the likelihood to be further excluded rises (Feinberg et al., 2012; Felps et al., 2006; Salmivalli et al., 1996; Wesselmann et al., 2013). Thus, it is possible that participants will mimic others exclusion behavior when being informed about it. This leads to our second hypothesis.

Hypothesis 2:

When being informed about the top 5 of a previous ranking, participants will follow this anchor and – if necessary – change their own ranking accordingly.

In section 2, we explain the experimental design and the methods used. In section 3, we show the results of the study that will be discussed in section 4. We discuss limitations of our study in section 5 and section 6, the conclusion, will close this paper.

2. Design and Method

The study was conducted online as a two stage-survey, programmed via *oTree* (D. L. Chen et al., 2016). In the first stage, participants were asked to fill out a questionnaire that contained on the one hand questions about their outward appearance (eye color, body shape, face shape, hair color, height, weight) and on the other hand questions about behavioral aspects and habits (book lover vs. movie lover, couch potato vs. sporty person, negative reciprocity, and trust (following the global preference survey of (Falk et al., 2016) and (Falk et al., 2018)). Also, they answered the Big Five (Rammstedt et al., 2013) and the Dark Triad questionnaire (Jones & Paulhus, 2014). Before starting the questionnaire, participants were informed that the experiment consisted of two stages and they will additionally receive a fixed payment of EUR 16 for finishing the current questionnaire *after* completing the second part of the experiment. The data of stage one was used to derive person cards. For this, we selected characteristics to be relevant for preference differentiation. The initial study allowed us to identify traits in individuals within the sample that most distinctly influenced behavior along the dimensions of trust and negative reciprocity, as determined by correspondence analysis. The selected traits were chosen to be independent and capable of co-occurring in real-life scenarios. Since we derived the characteristics from an empirical sample in the previous experiment, this condition is adequately met.

We used the profile method to define the stimuli as multiple attributes with various levels were included.

Based on the initial questionnaire in stage one, we first analyzed all personal attributes descriptively. The first attribute in the visual design was the calculation of the respondents' Body Mass Index (BMI), derived from reported height and weight, and categorized into three levels (low, medium, and high) according to the mean values. The attribute “body type” was excluded, as it is captured by BMI, and we avoided verbal descriptions to reduce social desirability bias, favoring more objective measures.

Behavioral components such as trust and negative reciprocity were reduced into three balanced groups, each representing approximately one-third of the respondents. We excluded the attribute “positive reciprocity” due to insufficient variance in the data. The remaining attributes were subjected to a correspondence analysis to assess their discriminatory power, which further informed the reduction of both the number of attributes and the attribute levels to those most relevant for the exclusion task.

Consequently, we defined three levels for the attribute "eyes" (green/grey, dark, blue), "face shape" (special, representing oblong or square; heart-like, encompassing heart, diamond, and triangle; and oval-round), and "hair color" (dark, light, red or colored).

For the remaining attributes, two distinct designs were employed: one with an visual focus and one with a behavioral focus. The behavioral design included the attributes gender (male, female), leisure preference (book lover, movie lover), activity level (sporty, sedentary), trust (low, medium, high), and negative reciprocity (low, medium, high). The visual design included BMI (low, medium, high), face shape (special, heart-like, oval-round), eye color (green/grey, dark, blue), hair color (dark, light, red or colored), and gender (male, female).

As the attributes exhibit different levels of variability and importance, we used an asymmetrical design for both the visual (3x3x3x2) and behavioral (2x2x2x3x3) attributes. The full factorial design would result in 54 visual and 72 behavioral possible personas, overwhelming participants. Therefore, we generated a reduced factorial design using the Orthoplan procedure in SPSS, reducing the number of profiles to 16 for both visual and behavioral designs. No hold-out cards were included for validation due to research economy, and we used only design cards. No simulation cards were included as well, as we do not aim at simulate more than the personas in the scenarios. In total, 307 participants completely answered the questionnaire. Figure 1 displays the created person cards.

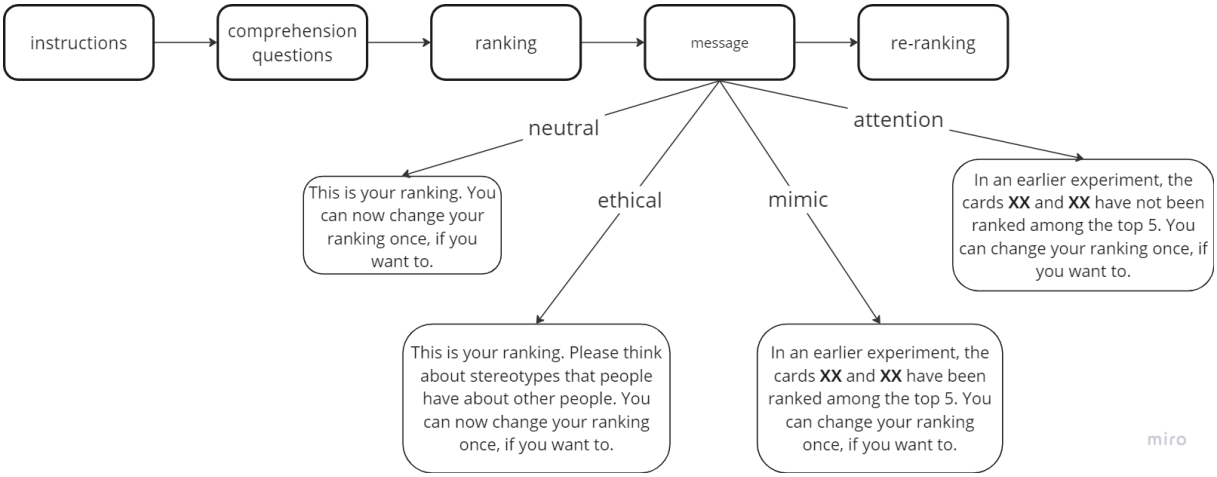
Figure 1: Person cards

BMI	Eyes	Hair	Face	Gender	Mode	No
high weight	blue	dark	oval - round	female	Design	1
normal weight	dark	dark	heart like *(heart, triangle, diamond)*	male	Design	2
low weight	blue	red or colored	heart like *(heart, triangle, diamond)*	female	Design	3
normal weight	grey-green	light	oval - round	female	Design	4
low weight	grey-green	dark	heart like *(heart, triangle, diamond)*	female	Design	5
low weight	grey-green	dark	oval - round	male	Design	6
normal weight	grey-green	red or colored	special *(oblong, square)*	female	Design	7
low weight	dark	red or colored	oval - round	male	Design	8
high weight	dark	dark	special *(oblong, square)*	female	Design	9
high weight	grey-green	light	heart like *(heart, triangle, diamond)*	male	Design	10
low weight	grey-green	dark	special *(oblong, square)*	male	Design	11
normal weight	blue	dark	special *(oblong, square)*	male	Design	12
low weight	grey-green	dark	special *(oblong, square)*	female	Design	13
high weight	grey-green	red or colored	special *(oblong, square)*	male	Design	14
low weight	dark	light	special *(oblong, square)*	female	Design	15
low weight	blue	light	special *(oblong, square)*	male	Design	16

Gender	Leisure	Activity	Trust	Neg Reciprocity	Mode	No
male	book lover	sporty	low trust	low negRec	Design	1
male	movie lover	couch potatoe	medium trust	low negRec	Design	2
male	movie lover	sporty	high trust	high negRec	Design	3
female	book lover	couch potatoe	high trust	medium negRec	Design	4
female	movie lover	couch potatoe	low trust	low negRec	Design	5
male	movie lover	couch potatoe	high trust	low negRec	Design	6
male	book lover	sporty	low trust	low negRec	Design	7
female	book lover	sporty	high trust	low negRec	Design	8
male	book lover	couch potatoe	low trust	high negRec	Design	9
female	movie lover	sporty	low trust	high negRec	Design	10
male	book lover	couch potatoe	low trust	medium negRec	Design	11
female	movie lover	sporty	low trust	medium negRec	Design	12
female	book lover	couch potatoe	medium trust	high negRec	Design	13
female	movie lover	couch potatoe	low trust	low negRec	Design	14
female	book lover	sporty	medium trust	low negRec	Design	15
male	movie lover	sporty	medium trust	medium negRec	Design	16

For variable coding, we made no assumptions regarding the correlation between attributes and the respondents' preference judgments; therefore, all attributes were treated as discrete.
Thus, we derived 32 person cards that were used in stage two. Figure 2 depicts the procedure.

Figure 2: Procedure of stage 2



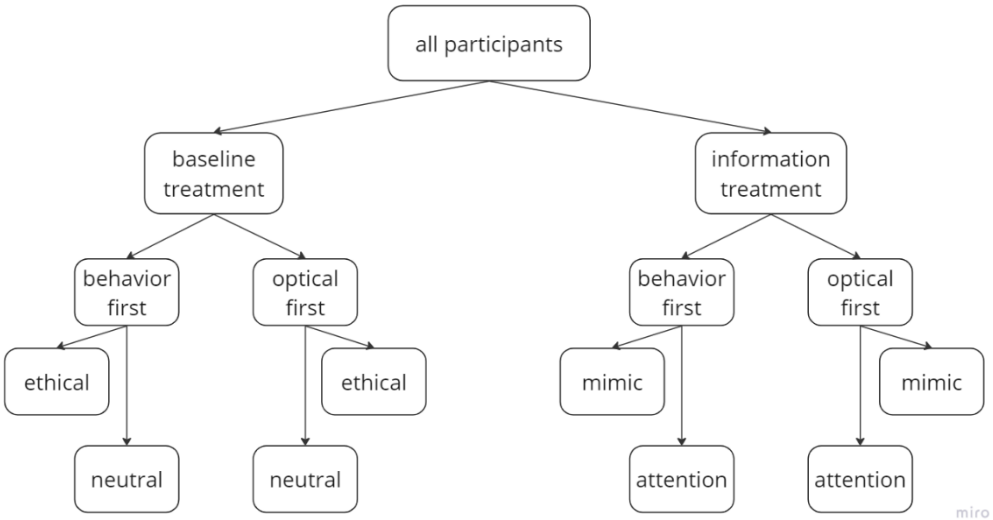
In the **baseline treatment** of the second part, participants were told to imagine a group of people in which every group member – except one person – could receive additional payoff. The group was to select the person to be excluded from receiving the additional payoff in an anonymous and covered voting. Participants' task was to estimate on how likely it is for individuals, each with a specific set of characteristics, to be *excluded* by the group. For this, they faced two sets of 16 person cards each. A person card depicted various characteristics and attributes, which were derived from the group's answers in the first part of the experiment and were of varying degrees. Each person card was unique, there was no duplication. Participants were asked to rank the cards according to their likelihood to be excluded by the group with rank number 1 being the most likely to be excluded card. One of the sets focused on the outward appearances (visual) and the other on the behavioral aspects (behavior). The two sets were depicted after one another. Participants were randomly divided into two subgroups starting either with the outward appearances (**visual first**) or the behavioral aspects (**behavior first**). It was not possible to exclude a card or to abstain.

After this initial ranking, each participant received a message telling her that she could review her ranking and decide whether she wanted to change it. Again, there were two subgroups: in the **neutral** group, the request for reevaluation was formulated *"This is your ranking. You can now change your ranking once, if you want to."* whilst the **ethical** group's request was *"This is*

your ranking. Please think about stereotypes that people have about other people. You can now change your ranking once, if you want to."[2]

In general, the **information setting** of this second part followed the same procedure. However, the messages after the first ranking differed. Using the first group's final ranking, we provided information about the top 5 most likely excluded cards of a previous session. In the **mimic treatment**, participants received the following message: *"In an earlier experiment, the cards **XX** and **XX** have been ranked among the top 5. You can change your ranking once, if you want to."* In the **attention treatment**, the message was *"In an earlier experiment, the cards **XX** and **XX** have not been ranked among the top 5. You can change your ranking once, if you want to."* The cards differed depending on the behavior/visual subgroup, individuals were part of. Figure 3 provides an overview of the treatments and subgroups.

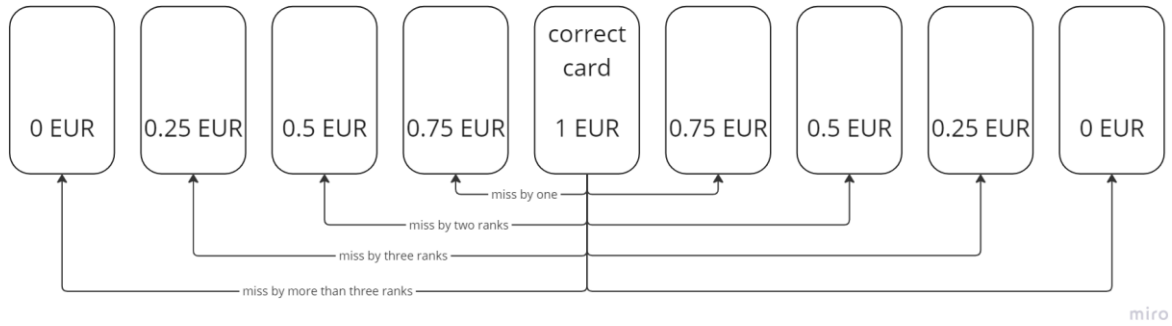
Figure 3: Treatments and Subgroups



The payoff depended on the accuracy of the participants estimate compared to the group's total ranking. Thus, participants had an incentive to reveal their true ranking (Krupka & Weber, 2013). After the whole group finished this part of the experiment, we calculated the group's total ranking. We have chosen a Borda-scheme (Levin & Nalebuff, 1995). Each time a participant ranked a certain card on the first place (highest likelihood to be excluded), this card received 16 points, for each second place 15 points etc. All points were summed up and lead to the final group ranking. Each individual ranking was then compared to the group ranking resulting to the payoff. Figure 4 shows the payoff scheme. Additionally, participants received the fixed amount of EUR 16 for completing the experiment.

[2] Both subgroups were distributed between the *visual first* and the *behavior first* group.

Figure 4: Payoff Scheme



In total, participants could earn up to EUR 48. To make sure that everyone understood the instructions, there were two comprehension questions to be solved before the experiment started. Participants could give a wrong answer to every comprehension question and retry. However, if they failed twice, they were excluded from taking part in the experiment.

All participants were students of the University Magdeburg or University of Applied Sciences Magdeburg-Stendal, recruited via *hroot* (Bock et al., 2014). The average payoff was EUR 24.

3. Results

In stage two, 175 participants finished the experiment (102 female). For a first overview, we depicted information about the sample we worked with in Table 1.

Table 1: the sample

	Neutral	**Ethical**	**Mimic**	**Attention**
Number of subjects	31	15	72	57
Male/Female/Other	14/17/0	7/7/1	29/43/0	22/35/0

Our first research question was whether there were patterns for participants to choose the cards that have the highest likelihood to be excluded.

In total, we conducted four conjoint analyses across different conjoint tasks (visual, behavioral ranking, and visual as well as behavioral re-ranking). The ranking data were analyzed using a ranking-based conjoint analysis to assess respondents' preferences. We employed an OLS regression to estimate the part-worth utilities, with the rank values serving as the dependent variable and the attribute levels as independent variables.

Importance of Attributes

The analysis revealed significant differences in the importance of the attributes across both the visual and behavioral tasks. For the behavioral conjoint task trust revealed the highest importance (55.692), followed by negative reciprocity (23.202.), activity level (15.468), and

gender (5.561). The behavioral attribute leisure (.077) contributed the least to the respondents' decisions. For the visual conjoint task BMI revealed the highest importance for respondents' decision (65.776), followed by hair color (14.862), gender (9.125), and eye color (5.540). The visual attribute face type (4.697) contributed the least to the respondents' decisions. There are no differences regarding the initial importance ranks (behavioral, visual) and the re-ranked importance ranks. All importance values and correlation between observed and estimated observations including the two re-ranked treatments are summarized in Table 2.

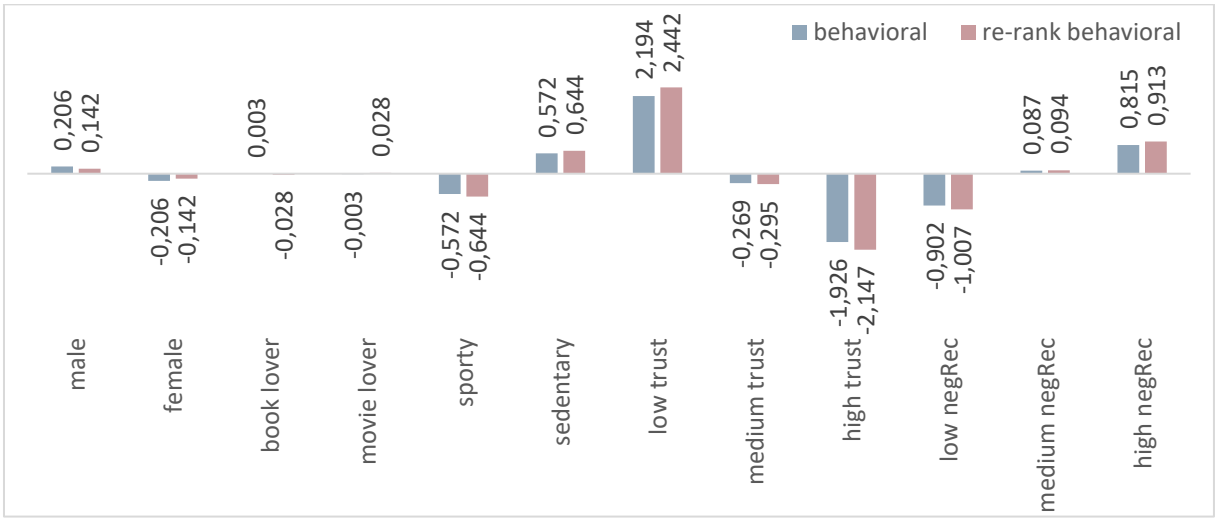
Table 2: importance values and correlations

attributes	**conjoint tasks**	
	behavioral	**behavioral re-rank**
gender	5.561	3.494
geisure	.077	.685
activity level	15.468	15.837
trust	55.692	56.395
neg. recicprocity	23.202	23.589
	Pearson r .998, $p < .001$	*Pearson r* .998, $p < .001$
attributes	**visual**	**visual re-rank**
BMI	65.776	56.240
eyes	5.540	8.562
hair	14.862	20.916
face	4.697	6.193
gender	9.125	8.088
	Pearson r .9968, $p < .001$	*Pearson r* .995, $p < .001$

Part worth utilities of attribute levels

The results indicate that certain levels of the attributes were consistently preferred by the respondents. In the behavioral conjoint task, we identified a low trust and a high negative reciprocity as the most preferred levels excluding personas from the group. Similarly, in the visual conjoint task, a high BMI and red or colored hair exhibited significantly higher preferences to exclude someone from the group. All part worth utilities for the different conjoint tasks are shown in figures 5 and 6.

Figure 5: Part worth utilities behavioral and behavioral re-rank



Thus, it turns out that a male person, preferring books and being rather sedentary showing low trust and a high negative reciprocity has the highest probability to be excluded from the group.

Figure 6: Part worth utilities visual and visual re-rank

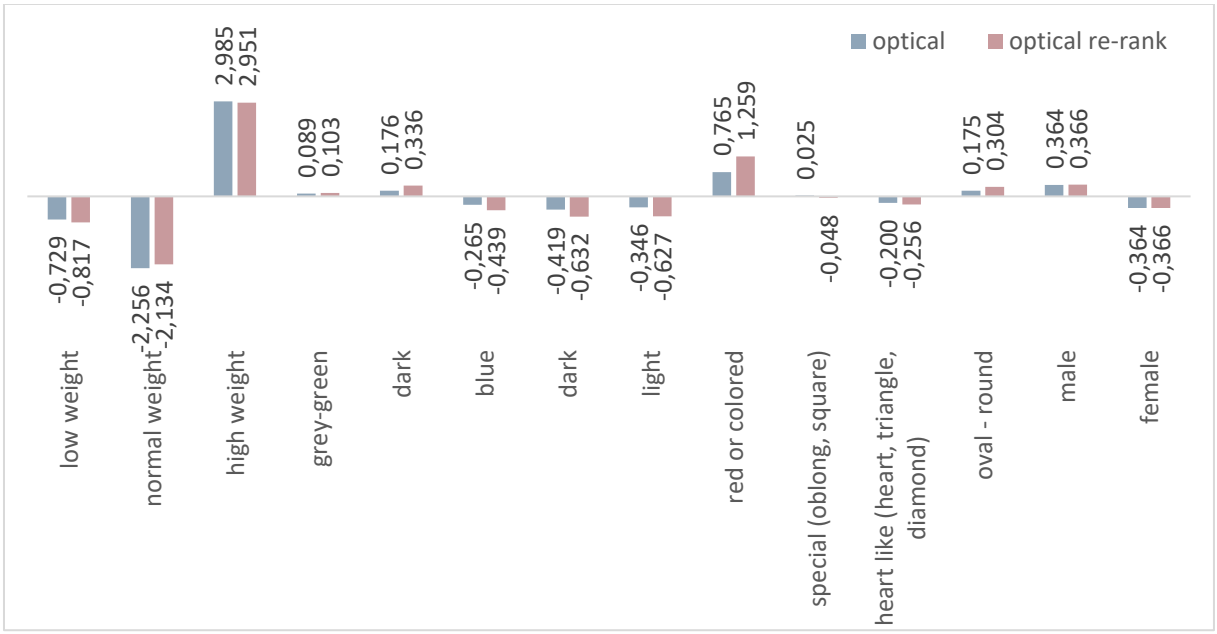


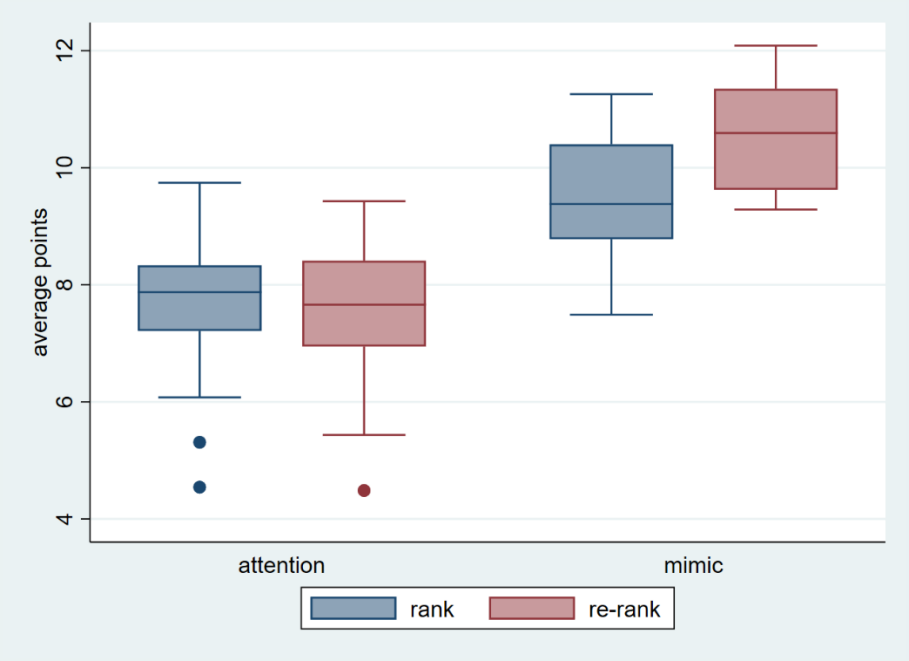
Figure shows that a male person, who has a high BMI, dark eyes, red or colored hair, and an oval-round face, has the highest probability to be excluded from the group.

We also tested for differences among the ratings of female versus male participants. The conjoint results according to the prior results for visual and behavioral design tasks revealed rather stable. Women tend to take the most advantage of ranking a person at number one who is male, a movie lover, a couch potato and shows low trust and high negative reciprocity. Trust is the most important attribute. Also, men benefit the most from ranking a person at number

one who is male, a book lover, a couch potato and shows low trust and high negative reciprocity. Among men, trust is also the most important attribute. The only difference between women or men is one attribute: women tend to exclude movie lovers, men tend to exclude book lovers. For both, however, this is the least important characteristic.

Our second research question was whether having information about the top 5 ranking of another group had an impact on the participants' ranking. For this, we compared the average points a card received in ranking and re-ranking. Although the ranking of another group does not necessarily affect the own group's ranking, we find that in the **mimic** treatment, the average points of the cards revealed to the participants were significantly (p=0.000, Wilcoxon signed rank test) *higher* in the re-ranking. Thus, we see that participants do follow other groups' rankings when being informed about it. This strongly supports Hypothesis 2. Figure 7 depicts the treatment effects.

Figure 7: comparison of ranking and re-ranking (mimic vs. attention)



In contrast, we find no significant (p =0.5639, Wilcoxon signed rank test) difference in the **attention** treatment. We also find that fewer cards were reranked in the attention treatment than in mimic which is in line with the previous finding although being not significant (p=0.1310, Mann-Whitney test).

4. Discussion

The aim of our study was to (i) find patterns for participants to choose the cards that have the highest likelihood to be excluded and (ii) to test whether revealing information about the

excluded resulted in even more participants coordinating on them. We found stable preferences using a conjoint analysis independent from the participants' own characteristics.

We also found that revealing information about excluded cards led to a mimicking process with participants following a previous ranking even though this did not necessarily need to influence their group ranking.

This finding is supported by other studies. Research consistently demonstrates a self-reinforcing cycle of social exclusion, where individuals already isolated or stigmatized face heightened rejection within groups (Wesselmann et al., 2013; Williams, 2007) found that group members tend to intensify bullying behaviors when they perceive a victim as already isolated, reinforcing the individual's exclusion. (Abrams & Hogg, 2001) similarly observed that publicly shaming or belittling low-status individuals worsens their social standing, as others avoid them to maintain their own status. This dynamic shows that exclusion often escalates once an individual becomes the target of negative comments or actions, providing group members with a sense of social relief through collective distancing.

Brauer & Chekroun, (2005) further found that individuals punished for norm violations face a similar cycle of exclusion, with publicized transgressions prompting group members to isolate them to reinforce social norms. Phelan et al. (2008) add that stigmatization fosters a "vicious cycle," especially when targets cannot counter the prevailing social narrative, leading others to view them as socially "dangerous" and avoidant. Salmivalli et al. (1996) captured this effect in a school context, describing a "Bandwagon Effect," where the likelihood of joining bullying behaviors increases as others participate, especially when the exclusion is public. Observing others' exclusion decisions alone can intensify group cohesion, as individuals act preemptively to protect their own social standing.

The visual attribute most strongly linked to group exclusion is a high BMI. Numerous studies support this finding: Neumark-Sztainer et al. (1998) found that overweight adolescents were less frequently chosen in social settings, revealing societal bias. A meta-analysis by Rudolph et al. (2009) further showed that overweight individuals face systemic disadvantages in both personal and professional relationships, often due to prejudices linking overweight with a lack of self-control. Hebl and Heatherton (1998) also found that overweight individuals are perceived as less competent and socially valuable, which hinders their social acceptance. Furthermore, Puhl and Heuer (2009) showed that overweight is often associated with negative stereotypes, such as laziness or lack of self-discipline, which can lead to social rejection.

Research consistently shows that negative behavioral traits, such as unfriendliness, selfishness, poor teamwork, or aggression, significantly increase the likelihood of exclusion from social and

professional groups. Key findings across studies align with this, highlighting that low trust and negative reciprocity heighten exclusion risks. Brauer and Chekroun (2005) found that norm violators are often excluded to preserve social harmony, while Cuddy et al. (2007) demonstrated that "cold" individuals face social rejection. Likewise, a study by Neuberg et al. (2000) confirms that antisocial or uncooperative behaviors lead to exclusion, with cooperative individuals remaining more integrated. Even specific personality traits, like low agreeableness and conscientiousness represent risk factors for being excluded (Rudert et al., 2020). Further, Kurzban and Leary (2001) and Marques et al. (1988) illustrate that individuals displaying deviant behaviors are judged harshly within their own groups—a phenomenon known as the "Black Sheep Effect." Exclusion for antisocial behavior can also create long-term effects, initiating a downward spiral in which excluded individuals show reduced prosocial behaviors and motivation, as Twenge et al. (2007) found, thereby reinforcing the cycle of exclusion.

Both most important attributes, BMI and less social behavior, are also in line with prior research. A study by Puhl and Latner (2007) found that overweight individuals who display unfriendly or aggressive behavior face a "double stigma"—due to both their appearance and their behavior—intensifying social rejection even further.

Furthermore, we found that an oval-round face shape increases the likelihood of group exclusion, consistent with other studies. Zebrowitz and Montepare (2008) found that "baby-like" features, such as round faces, were perceived as friendlier but less competent. In competitive or performance-oriented contexts, oval or round faces are often associated with lower assertiveness and competence, leading to higher exclusion rates (Carré & McCormick, 2008). Similarly, Oosterhof and Todorov (2008) and Rule and Ambady (2008) showed that round faces are viewed as trustworthy but less competent, leading to exclusion in competence-focused groups, though they are more accepted in cooperative settings.

Our findings suggest that while gender is less influential than other attributes, men are generally more likely to be excluded. This aligns with Benenson et al., (2011) who found that men often face isolation in competitive settings when seen as aggressive or uncooperative. The "Precarious Manhood" effect (Bosson & Vandello, 2011) further indicates that men are viewed as burdensome if they fail to meet traditional gender expectations, leading to their exclusion from group benefits. In contrast, women are typically perceived as more cooperative, making them more likely to be included and rewarded in group efforts (Balliet et al., 2011). Bear and Woolley (2011) suggest that men face exclusion when their behavior does not align with stereotypes of harmony, increasing their risk of isolation. The tendency to exclude men poses a further issue, as Courtenay (2000) notes that men are less likely to seek social support, leaving

them more vulnerable to isolation, especially during stressful times, which may lead to a cycle of increasing exclusion for already marginalized men.

Interestingly, oval faces are more often a basis for exclusion due to perceptions of weakness, while men are also more frequently excluded, as studies suggest, for appearing uncooperative. Future research combining these two dimensions—facial shape and gender—could offer valuable insights into how these attributes jointly influence exclusion dynamics.

In our study, exclusion decisions were largely independent of participants' own characteristics, a finding supported by other studies. These studies suggest that group members often base exclusion on stereotypical traits or behaviors of others rather than on their own appearance or behavior (e.g. (Fiske et al., 2002; Neuberg et al., 2000; Palermo & Rhodes, 2007), which might be more an evolutionary mechanism reading based on whether a person deviate in appearance or behavior (Asch, 1955; Kurzban & Leary, 2001).

5. Limitations

Despite OLS assumptions of interval-scale data and residual homoscedasticity, numerous studies confirm its validity, even with rank data, yielding results comparable to non-metric techniques (Green & Krieger, 1993). Since we focus on aggregated data, we see no methodological limitations in this approach.

Several limitations should be noted. First, we examined visual and behavioral attributes separately, so we cannot conclusively determine which is more relevant or if BMI or low trust is the primary exclusion factor. A combined approach in future research could clarify this.

Second, additional attributes may be relevant. We included attributes reported by participants, most of whom are University of Magdeburg students, which may limit generalizability. A broader dataset might capture other attribute distributions. For instance, we did not include glasses as a variable; although glasses affect how we perceive the faces of the people wearing them (Edwards, 1987; Harris et al., 1982; Hasart & Hutchinson, 1993; Leder et al., 2011) including differences for type of gender (Terry & Hall, 1989).

Third, we excluded attributes that might raise ethical concerns, such as disability, visible scars, or ethnic background, though these factors likely influence exclusion. (Fine & Asch, 1988) found that visible disabilities led to higher exclusion, and Houston and Bull (1994) observed similar effects for scars and skin conditions. Ethnicity also affects group inclusion, as shown by Bigler et al. (2001).

Fourth, cultural factors should be examined to better understand exclusion origins. Fifth, the study’s small sample size warrants caution, as larger samples are needed to ensure replicability

and reliability. Lastly, we used verbal descriptions of attributes rather than images, which could affect participant perceptions, as seen in previous studies (e.g. (Fine & Asch, 1988; Houston & Bull, 1994).

6. Conclusion

Our findings reveal a significant tendency to exclude individuals with higher BMI, underscoring deep-rooted societal biases. In contrast to the fairy tale, it wasn't the qualities associated with a "slim, friendly character" like Cinderella that predicted exclusion from group benefits, but rather characteristics linked to her stepsisters — unsociable behavior, low trust, and higher weight.

The implications of this research underscore a significant exclusion bias against individuals with higher BMI, aligning with previous studies that reveal deep societal prejudice. The global increase in overweight populations suggests this issue is not merely about appearance but reflects a broader social dynamic. Our findings show that stable physical features—such as face shape and eye color—alongside BMI also contribute to group exclusion. Since most of these traits are unchangeable, it's vital to raise awareness and promote sensitivity toward such biases. Social cognition theories, like Weiner's Attribution Theory (1985), indicate that people often attribute behaviors to internal causes, which may lead individuals to view higher BMI as a matter of personal responsibility. When combined with unsociable behavior, this attribution can intensify negative perceptions and reinforce exclusion. Furthermore, behaviors like low trust and negative reciprocity only heighten exclusion risk, suggesting that adopting prosocial behaviors could help maintain group connections. Excluded individuals often experience reduced motivation for social interaction, perpetuating a cycle of exclusion that can have severe long-term consequences, such as social isolation, depression, and low self-esteem.

Overall, this study shows that both behavioral and physical factors significantly impact group exclusion, and these criteria remain consistent across genders and contexts. This research advances our understanding of group exclusion dynamics, highlighting a consistent preference for excluding certain individuals and emphasizing the need to address and counteract such biases within social settings.

Acknowledgements

We thank the faculty of Economics and Management of the Otto-von-Guericke University for funding this project and Christian Stahr for his support in preparing the data.

References

Abbink, K., & Doğan, G. (2019). How to choose your victim. *Games and Economic Behavior*, *113*, 482–496. https://doi.org/10.1016/j.geb.2018.10.006

Abrams, D., & Hogg, M. A. (2001). Collective Identity: Group Membership and Self-Conception. In M. A. Hogg & R. S. Tindale (Eds.), *Blackwell Handbook of Social Psychology: Group Processes* (pp. 425–460). Blackwell Publishers Ltd. https://doi.org/10.1002/9780470998458.ch18

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 31–35. https://www.jstor.org/stable/24943779

Balliet, D., Li, N. P., Macfarlan, S. J., & van Vugt, M. (2011). Sex differences in cooperation: A meta-analytic review of social dilemmas. *Psychological Bulletin*, *137*(6), 881–909. https://doi.org/10.1037/a0025354

Baumeister, R. F., & Leary, M. R [Mark R.] (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Bear, J. B., & Woolley, A. W. (2011). The role of gender in team collaboration and performance. *Interdisciplinary Science Reviews*, *36*(2), 146–153. https://doi.org/10.1179/030801811X13013181961473

Benenson, J. F., Markovits, H., Thompson, M. E., & Wrangham, R. W. (2011). Under threat of social exclusion, females exclude more than males. *Psychological Science*, *22*(4), 538–544. https://doi.org/10.1177/0956797611402511

Bershadskyy, D., & Seidel, A. (2024). Choosing a victim you know: Introducing communiation to the mobbing game. *Journal of Behavioral and Experimental Economics*, *112*, 102265. https://doi.org/10.1016/j.socec.2024.102265

Bigler, R. S., Brown, C. S., & Markell, M. (2001). When groups are not created equal: Effects of group status on the formation of intergroup attitudes in children. *Child Development*, *72*(4), 1151–1162. https://doi.org/10.1111/1467-8624.00339

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review*, *71*, 117–120. https://doi.org/10.1016/j.euroecorev.2014.07.003

Boehm, C. (1986). Capital punishment in tribal Montenegro: Implications for law, biology, and theory of social control. *Ethology and Sociobiology*, *7*(3-4), 305–320. https://doi.org/10.1016/0162-3095(86)90056-7

Bosson, J. K., & Vandello, J. A. (2011). Precarious Manhood and Its Links to Action and Aggression. *Current Directions in Psychological Science*, *20*(2), 82–86. https://doi.org/10.1177/0963721411402669

Brauer, M., & Chekroun, P. (2005). The Relationship Between Perceived Violation of Social Norms and Social Control: Situational Factors Influencing the Reaction to Deviance. *Journal of Applied Social Psychology*, *35*(7), 1519–1539. https://doi.org/10.1111/j.1559-1816.2005.tb02182.x

Cahnman, W. J. (1968). The Stigma of Obesity. *The Sociological Quarterly*, *9*(3), 283–299. https://doi.org/10.1111/j.1533-8525.1968.tb01121.x

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings. Biological Sciences*, *275*(1651), 2651–2656. https://doi.org/10.1098/rspb.2008.0873

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Chen, Z., Williams, K. D., Fitness, J., & Newton, N. C. (2008). When hurt will not heal: Exploring the capacity to relive social and physical pain. *Psychological Science*, *19*, 789–795. https://doi.org/10.1111/j.1467-9280.2008.02158.x

Courtenay, W. H. (2000). Constructions of masculinity and their influence on men's well-being: A theory of gender and health. *Social Science & Medicine (1982)*, *50*(10), 1385–1401. https://doi.org/10.1016/S0277-9536(99)00390-1

Crocker, J., Cornwell, B., & Major, B. (1993). The stigma of overweight: Affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, *64*(1), 60–70. https://doi.org/10.1037//0022-3514.64.1.60

Crocker, J., Major, B., & Steele, C. (1998). Social Stigma. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 504–553). McGraw-Hill.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648. https://doi.org/10.1037/0022-3514.92.4.631

DeJong, W. (1980). The Stigma of Obesity: The Consequences of Naive Assumptions Concerning the Causes of Physical Deviance. *Journal of Health and Social Behavior*, *21*(1), 75. https://doi.org/10.2307/2136696

DeWall, C. N., & Baumeister, R. F. (2006). Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *Journal of Personality and Social Psychology*, *91*(1), 1–15. https://doi.org/10.1037/0022-3514.91.1.1

Edwards, K. (1987). Effects of Sex and Glasses on Attitudes toward Intelligence and Attractiveness. *Psychological Reports*, *60*(2), 590. https://doi.org/10.1177/003329418706000201

Eisenberger, N. I., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, *8*(7), 294–300. https://doi.org/10.1016/j.tics.2004.05.010

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global Evidence on Economic Preferences *The Quarterly Journal of Economics*, *133*(4), 1645–1692. https://doi.org/10.1093/qje/qjy013

Falk, A., Becker, A., Dohmen, T., Huffman, D., & Sunde, U. (2016). The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences. *IZA Discussion Paper No. 9674*.

Feinberg, M., Willer, R., & Keltner, D. (2012). Flustered and faithful: Embarrassment as a signal of prosociality. *Journal of Personality and Social Psychology*, *102*(1), 81–97. https://doi.org/10.1037/a0025403

Felps, W., Mitchell, T. R., & Byington, E. (2006). How, When, and Why Bad Apples Spoil the Barrel: Negative Group Members and Dysfunctional Groups. *Research in Organizational Behavior*, *27*, 175–222. https://doi.org/10.1016/S0191-3085(06)27005-9

Fine, M., & Asch, A. (1988). Disability Beyond Stigma: Social Interaction, Discrimination, and Activism. *Journal of Social Issues*, *44*(1), 3–21. https://doi.org/10.1111/j.1540-4560.1988.tb02045.x

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Goette, L., Huffman, D., & Meier, S. (2006). The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups. *American Economic Review*, *96*(2), 212–216. https://doi.org/10.1257/000282806777211658

Goffman, E. (1963). *Stigma: Notes on the Management of spoiled Identity*. Prentice Hall Inc.

Green, P. E., & Krieger, A. M. (1993). Conjoint analysis with product-positioning applications. In J. Eliashberg & G. L. Lilien (Eds.), *Handbooks in OR & MS: Marketing* (Vol. 5, pp. 467–515). https://doi.org/10.1016/S0927-0507(05)80033-7

Gruter, M. W., & Masters, R. (1986). Ostracism as a social and biological phenomenon: An introduction. *Ethology and Sociobiology*, *7*, 149–158.

Harris, M. B., Harris, R. J., & Bochner, S. (1982). Fat, Four-Eyed, and Female: Stereotypes of Obesity, Glasses, and Gender 1. *Journal of Applied Social Psychology*, *12*(6), 503–516. https://doi.org/10.1111/j.1559-1816.1982.tb00882.x

Hasart, J. K., & Hutchinson, K. L. (1993). The Effects of Eyeglasses on Perceptions of Interpersonal Attraction. *Journal of Social Behavior and Personality*, *8*(3), 521–528.

Heatherton, T. F [T. F.], Kleck, R. E., Hebl, M. R [M. R.], & Hull, J. G. (Eds.). (2000). *The social psychology of stigma*. The Guilford Press.

Hebl, M. R [Michelle R.], & Heatherton, T. F [Todd F.] (1998). The Stigma of Obesity in Women: The Difference is Black and White. *Personality and Social Psychology Bulletin*, *24*(4), 417–426. https://doi.org/10.1177/0146167298244008

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*, 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Houston, V., & Bull, R. (1994). Do people avoid sitting next to someone who is facially disfigured? *European Journal of Social Psychology*, *24*(2), 279–284. https://doi.org/10.1002/ejsp.2420240205

Jones, D. N., & Paulhus, D. L. (2014). Introducing the short Dark Triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*(1), 28–41. https://doi.org/10.1177/1073191113514105

Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development*, *84*(3), 772–790. https://doi.org/10.1111/cdev.12012

Krupka, E. L., & Weber, R. A. (2013). Identifying Social Norms using Coordination Games: Why does Dictator Game Sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524. https://doi.org/10.1111/jeea.12006

Kurzban, R., & Leary, M. R [M. R.] (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin*, *127*(2), 187–208. https://doi.org/10.1037/0033-2909.127.2.187

Leder, H., Forster, M., & Gerger, G. (2011). The Glasses Stereotype Revisited. *Swiss Journal of Psychology*, *70*(4), 211–222. https://doi.org/10.1024/1421-0185/a000059

Lelieveld, G.-J., Harris, L. T., & van Dillen, L. F. (2020). Jumping on the 'bad'wagon? How group membership influences responses to the social exclusion of others. *Social Cognitive and Affective Neuroscience*, *15*(5), 571–586. https://doi.org/10.1093/scan/nsaa070

Levin, J., & Nalebuff, B. (1995). An Introduction to Vote-Counting Schemes. *Journal of Economic Perspectives*, *9*(1), 3–26. https://doi.org/10.1257/jep.9.1.3

Levine, M., Cassidy, C., Brazier, G., & Reicher, S. (2002). Self-Categorization and Bystander Non-intervention: Two Experimental Studies 1. *Journal of Applied Social Psychology*, *32*(7), 1452–1463. https://doi.org/10.1111/j.1559-1816.2002.tb01446.x

Marques, J. M., Yzerbyt, V. Y., & Leyens, J.-P. (1988). The “Black Sheep Effect”: Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*(1), 1–16. https://doi.org/10.1002/ejsp.2420180102

Neuberg, S. L., Smith, D. M., & Asher, T. (2000). Why people stigmatize: Toward a biocultural framework. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma* (pp. 31–61). The Guilford Press.

Neumark-Sztainer, D., Story, M., & Faibisch, L. (1998). Perceived stigmatization among overweight African-American and Caucasian adolescent girls. *The Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, *23*(5), 264–270. https://doi.org/10.1016/s1054-139x(98)00044-5

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75–92. https://doi.org/10.1016/j.neuropsychologia.2006.04.025

Phelan, J. C., Link, B. G., & Dovidio, J. F. (2008). Stigma and prejudice: One animal or two? *Social Science & Medicine (1982)*, *67*(3), 358–367. https://doi.org/10.1016/j.socscimed.2008.03.022

Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: A review and update. *Obesity (Silver Spring, Md.)*, *17*(5), 941–964. https://doi.org/10.1038/oby.2008.636

Puhl, R. M., & Latner, J. D. (2007). Stigma, obesity, and the health of the nation's children. *Psychological Bulletin*, *133*(4), 557–580. https://doi.org/10.1037/0033-2909.133.4.557

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2013). A Short Scale for Assessing the Big Five Dimensions of Personality: 10 Item Big Five Inventory (BFI-10). *Methoden, Daten, Analysen*, *7*(2), 233–249. https://doi.org/10.12758/mda.2013.013 (17 Pages / methods, data, analyses, Vol 7, No 2 (2013)).

Rudert, S. C., Keller, M. D., Hales, A. H., Walker, M., & Greifeneder, R. (2020). Who gets ostracized? A personality perspective on risk and protective factors of ostracism. *Journal of Personality and Social Psychology*, *118*(6), 1247–1268. https://doi.org/10.1037/pspp0000271

Rudolph, C. W., Wells, C. L., Weller, M. D., & Baltes, B. B. (2009). A meta-analysis of empirical studies of weight-based bias in the workplace. *Journal of Vocational Behavior*, *74*(1), 1–10. https://doi.org/10.1016/j.jvb.2008.09.008

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science*, *19*(2), 109–111. https://doi.org/10.1111/j.1467-9280.2008.02054.x

Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior*, *22*(1), 1–15. https://doi.org/10.1002/(SICI)1098-2337(1996)22:1<1::AID-AB1>3.0.CO;2-T

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Terry, R. L., & Hall, C. A. (1989). Affective responses to eyeglasses: Evidence of a sex difference. *Journal of the American Optometric Association*, *60*(8), 609–611.

Twenge, J. M., Baumeister, R. F., DeWall, C. N., Ciarocco, N. J., & Bartels, J. M. (2007). Social exclusion decreases prosocial behavior. *Journal of Personality and Social Psychology*, *92*(1), 56–66. https://doi.org/10.1037/0022-3514.92.1.56

Voci, A. (2006). The link between identification and in-group favouritism: Effects of threat to social identity and trust-related emotions. *The British Journal of Social Psychology*, *45*(Pt 2), 265–284. https://doi.org/10.1348/014466605X52245

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573. https://doi.org/10.1037/0033-295X.92.4.548

Wesselmann, E. D., Wirth, J. H., Pryor, J. B., Reeder, G. D., & Williams, K. D. (2013). When Do We Ostracize? *Social Psychological and Personality Science*, *4*(1), 108–115. https://doi.org/10.1177/1948550612443386

Williams, K. D. (1997). Social Ostracism. In R. M. Kowalski (Ed.), *Aversive Interpersonal Behaviors* (pp. 133–170). Springer US. https://doi.org/10.1007/978-1-4757-9354-3_7

Williams, K. D. (2007). Ostracism: The Power of Silence. In A. W. Kruglanski & E. T. Higgins (Eds.), *Handbook of Social Psychology (2nd ed.).* Guilford Press.

Williams, K. D., Cheung, C. K. T., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology*, *79*(5), 748–762. https://doi.org/10.1037/0022-3514.79.5.748

Zebrowitz, L. A., & Montepare, J. M. (2008). Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, *2*(3), 1497. https://doi.org/10.1111/j.1751-9004.2008.00109.x

Zippelius, R. (1986). Exclusion and shunning as legal and social sanctions. *Ethology and Sociobiology*, *7*(3-4), 159–166. https://doi.org/10.1016/0162-3095(86)90044-0

Appendix

Instructions part one – the questionnaire

Instructions

Dear participant.

Welcome to part one of our experiment. In this part, you will have to answer a questionnaire.

We ask you to complete the questionnaire in one attempt. It is not possible to pause and resume the questionnaire.
Please make sure to concentrate on the questions asked as it will not be possible to go back to previous pages once you have finished them.

You will earn 16 Euro for the successful completion of the questionnaire. The amount will be paid at the end of the second part of this experiment. The second part of the experiment will be conducted online and will take place in March/April. We will invite you separately.

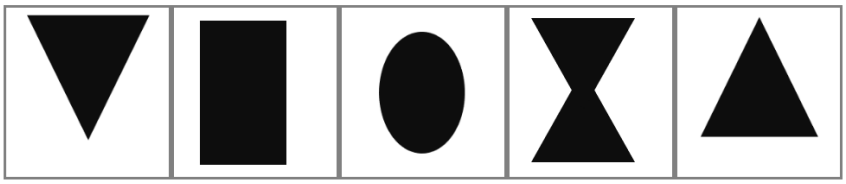
Thank you for participating!

Weiter

Visual attributes

Body

Please select the tile that represents your body type most.



Weiter

Height, Weight

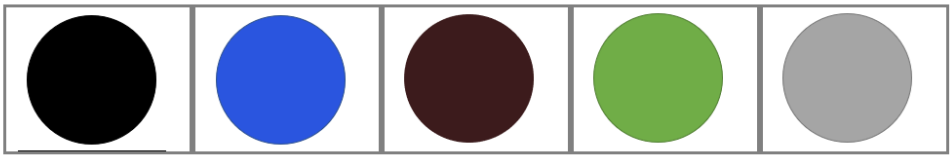
What is your body height in cm?

What is your body weight in kg?

Weiter

Eyes

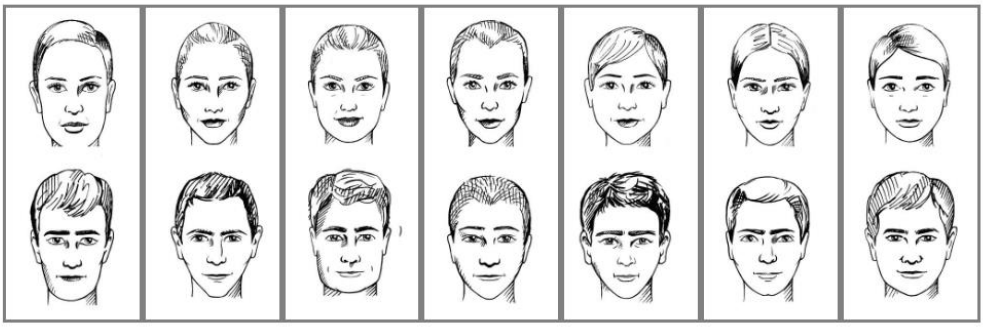
Please select the tile that fits your eye colour most. In case you have more than one eye colour, plese choose the dominant colour. In case you have two different coloured eyes, please select the colour of your right eye.



Weiter

Face

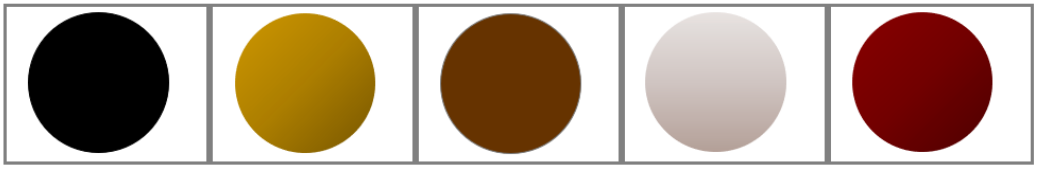
Please select the tile that represents your face type most.



Weiter

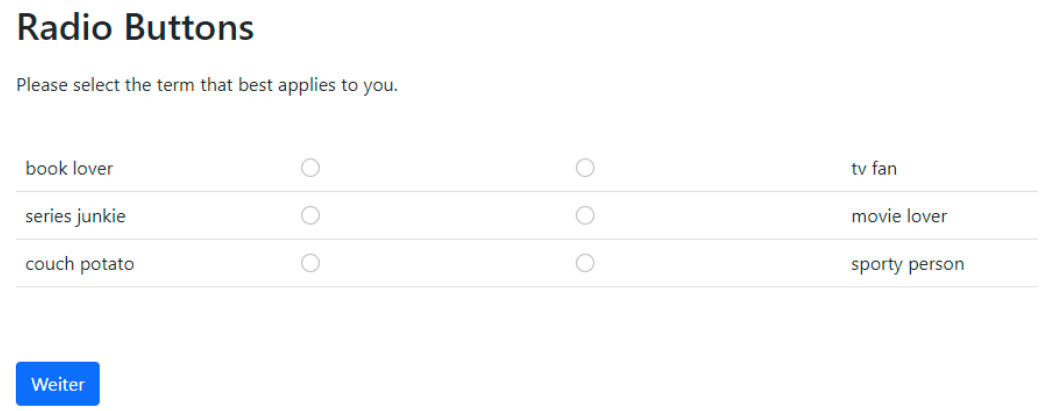
Hair

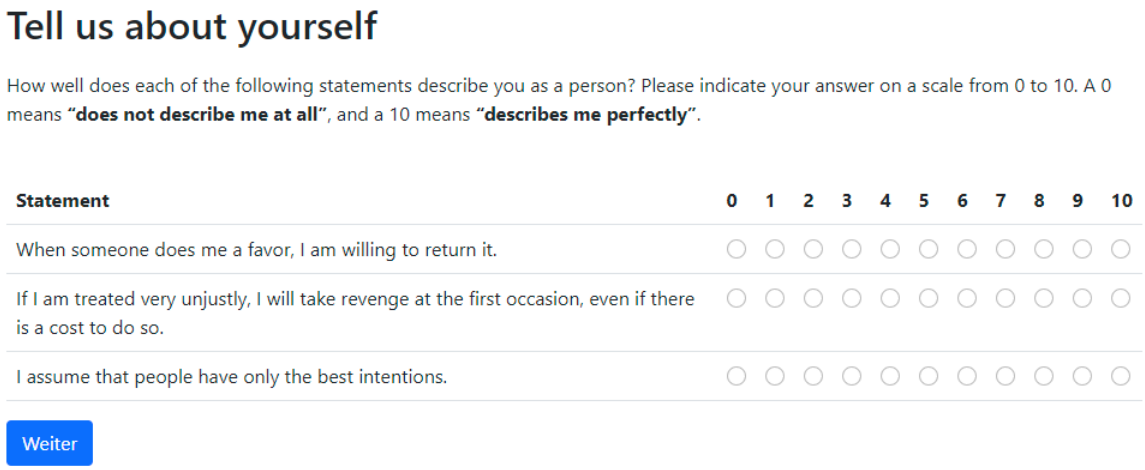
Please select the tile that fits your hair colour most.



Weiter

Behavioral attributes





Tell us about yourself

Imagine the following situation: Today you unexpectedly received 1,600 U.S. dollars.

How much of this amount would you donate to a good cause?

Weiter

Tell us about yourself

In the next block, we will show you statements. There will be no wrong or right answers - so feel free to answer spontaneously.

Please rate the following statements from **1 (strongly disagree)** to **5 (strongly agree)**.

Statement	1	2	3	4	5
I'm rather cautious, reserved.	○	○	○	○	○
I trust others easily and believe in the good in people.	○	○	○	○	○
I am comfortable and tend to be lazy.	○	○	○	○	○
I am relaxed and don't let stress upset me.	○	○	○	○	○
I have very little interest in art.	○	○	○	○	○
I come out of my shell, I'm sociable.	○	○	○	○	○
I tend to criticise others.	○	○	○	○	○
I complete tasks thoroughly.	○	○	○	○	○
I quickly become nervous and insecure.	○	○	○	○	○
I have an active imagination, I am very fantasy orientated.	○	○	○	○	○

Weiter

Tell us about yourself

In the next block, we will show you statements. There will be no wrong or right answers - so feel free to answer spontaneously.

Please rate the following statements from **1 (strongly disagree)** to **5 (strongly agree)**.

Statement	1	2	3	4	5
It is not wise to tell your secrets.	○	○	○	○	○
People see me as a natural leader.	○	○	○	○	○
I like to get revenge on authorities.	○	○	○	○	○
I like to use clever manipulation to get my way.	○	○	○	○	○
I hate being the center of attention.	○	○	○	○	○
I avoid dangerous situations.	○	○	○	○	○
Whatever it takes, you must get the important people on your side.	○	○	○	○	○
Many group activities tend to be dull without me.	○	○	○	○	○
Payback needs to be quick and nasty.	○	○	○	○	○
Avoid direct conflict with others because they may be useful in the future.	○	○	○	○	○
I know that I am special because everyone keeps telling me so.	○	○	○	○	○
People often say I am out of control.	○	○	○	○	○
It is wise to keep track of information that you can use against people later.	○	○	○	○	○
I like to get acquainted with important people.	○	○	○	○	○
It is true that I can be mean to others.	○	○	○	○	○
You should wait for the right time to get back at people.	○	○	○	○	○
I feel embarrassed if someone compliments me.	○	○	○	○	○
People who mess with me always regret it.	○	○	○	○	○
There are things you should hide from other people to preserve your reputation.	○	○	○	○	○
I have been compared to famous people.	○	○	○	○	○
I have never gotten into trouble with the law.	○	○	○	○	○
Make sure your plans benefit yourself, not others.	○	○	○	○	○
I am an average person.	○	○	○	○	○
I enjoy having sex with people I hardly know.	○	○	○	○	○
Most people can be manipulated.	○	○	○	○	○
I insist on getting the respect I deserve.	○	○	○	○	○
I will say anything to get what I want.	○	○	○	○	○

Weiter

Instructions part two

Welcome (all)

Welcome to the second part of the experiment that started in January with a questionnaire. With participating in this experiment, you can earn money based on your performance. Please read the following instructions carefully.

Important: Once you have started the experiment, you will not be able to pause and resume it.

Button: next

Instructions: ethical and neutral

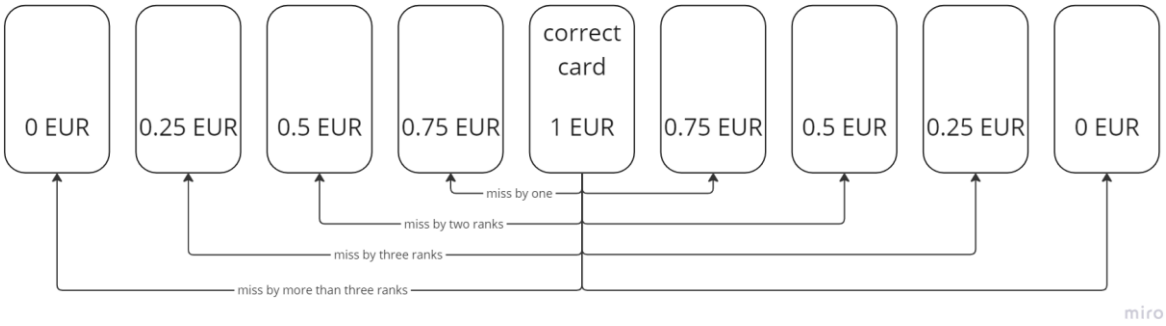
Imagine that there is a group of people, in which every group member – with the exception of one person – will receive an additional payoff. The group selects one person to be excluded from receiving that additional payoff in an anonymous and covered voting. Your task is to make guesses on how likely it is for individuals, each with a specific set of characteristics, to be excluded by the group.

You will see two sets of 16 person cards each depicting a person's characteristics. Please, rank them according to the **likelihood that your group will exclude them from receiving an additional payoff**. Position number 1 is the person you think is most likely to be excluded by the group. Your group members will also rank the same person cards. After all group members have completed their ranking tasks, the group's overall ranking will be compared to your ranking.

Your payment depends on the accuracy of your guess concerning the group's overall ranking. For each person card that you have ranked, you receive a payment as follows:

If you place a person card on the exactly the same rank as the group's overall ranking does, you receive 1 Euro. If you miss the group's overall rank by one position, you receive 0.75 Euro. If you miss by two positions, you receive 0.50 Euro. If you miss by three positions, you receive 0.25 Euro. If you miss the group's overall rank for that person card by four or more positions, you receive nothing for that card. (Please, find an illustration of the payment scheme for each of the 16 ranked person cards below.)

Additionally, you will receive a fixed amount of 16 Euro for completing the experiment.

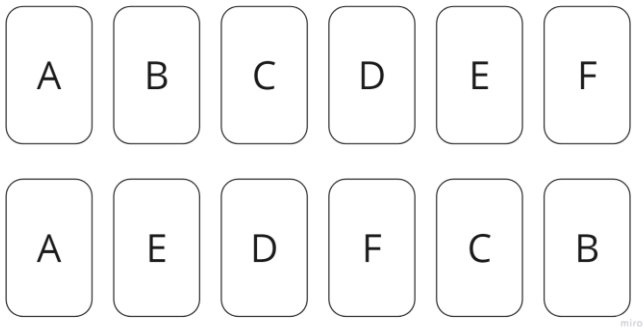


Please, answer the following comprehension questions correctly before the experiment begins. **Attention:** You can fail once. However, if you answer a question incorrectly twice, you will not be able to take part in the experiment.

Button: comprehension questions

---

1. Imagine, you rank the three person cards A, B, and C as follows: 1. B, 2. C, 3. A According to the instructions, which person card has the highest likelihood to be excluded from receiving further payoffs by the group?
2. Please, take a look at the illustration below. In the first line, you see the group's overall ranking. In the second line, you see an individual's ranking. How much will the individual earn with this ranking? (Please, do not include the fixed payment of 16 Euro for completing the experiment.)



Button: ready

Scenario: correct: You have answered all questions correctly and can now start the experiment.

Button: Start

Scenario: at least one mistake (mistakes will be highlighted red): Please check the highlighted question(s) again.

Two groups: visual first, behavior first

(after first ranking)

Group neutral: This is your ranking. You can now change your ranking once, if you want to.

Buttons: change; continue

Group ethical: This is your ranking. Please, think about stereotypes that people have about other people. You can now change your ranking once, if you want to.

Buttons: change; continue

(same procedure for the second set of 16 cards)

Thank you for submitting your ranking. The experimenter will contact you as soon as all group members have completed their ranking tasks. You will then be informed on the group's overall raking and receive your payoff.

You can now close the browser window. Thank you for participating.

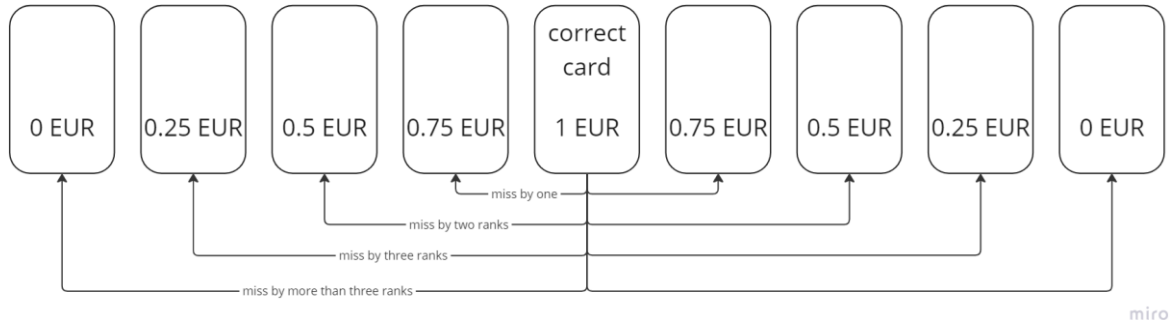
Instructions: mimic and attention

Imagine that there is a group of people, in which every group member – with the exception of one person – will receive an additional payoff. The group selects one person to be excluded from receiving that additional payoff in an anonymous and covered voting. Your task is to make guesses on how likely it is for individuals, each with a specific set of characteristics, to be excluded by the group.

You will see two sets of 16 so called person cards each that depict a person's characteristics and will be asked to rank them according to the **likelihood of your group to exclude them**. Position number 1 is the person card you think to be most likely to be excluded by the group. Your group members will also do their ranking and after the group finished the ranking, the complete ranking will be compared to your ranking.

Your payment will be according to the accuracy of your perception of the group's ranking. This will hold as follows:

If you place a card on the correct rank, you will receive 1 Euro. If you miss the correct rank by one, you will receive 0.75 Euro etc. Please find an illustration below. Additionally, you will receive 16 Euro for completing part one and two of the experiment.



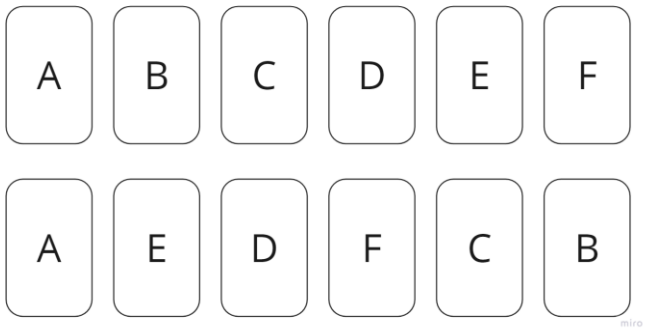
Please answer the following comprehension questions before the real experiment begins. **Attention:** You can fail once. However, if you answer a question incorrectly twice, you will not be able to take part in the experiment.

Button: comprehension questions

---

1. Imagine, you rank three person cards A, B, and C. You rank as follows: 1. B, 2. C, 3. A According to the instructions, which person card has the highest likelihood to be excluded by the group?

2. Please look at the illustration below. In the first line, you see the ranking of the group. In the second line, you see an individual's ranking. How much will the individual earn with this ranking? (Please *exclude* the 16 Euro for completing both parts of the experiment)



Button: ready

Scenario: correct: You have answered all questions correctly and can now start the experiment.
Button: Start

Scenario: at least one mistake (mistakes will be highlighted red): Please check the highlighted question(s) again.

Two groups: visual first, behavior first
(after first ranking)

Group mimic: This is your ranking.

Behavior first:

Behavior: In an earlier experiment, the cards **12** and **13** have been ranked among the top 5. You can change your ranking once, if you want to.

Visual: In an earlier experiment, the cards **1** and **8** have been ranked among the top 5. You can change your ranking once, if you want to.

Visual first:

Behavior: In an earlier experiment, the cards 5 and 14 have been ranked among the top 5. You can change your ranking once, if you want to.

Visual: In an earlier experiment, the cards **7** and **8** have been ranked among the top 5. You can change your ranking once, if you want to.

Buttons: change; continue

Group attention: This is your ranking.

Behavior first:

Behavior: In an earlier experiment, the cards **3** and **7** have not been ranked among the top 5. You can change your ranking once, if you want to.

Visual: In an earlier experiment, the cards **4** and **15** have not been ranked among the top 5. You can change your ranking once, if you want to.

Visual first:

Behavior: In an earlier experiment, the cards **2** and **16** have not been ranked among the top 5. You can change your ranking once, if you want to.

Visual: In an earlier experiment, the cards **6** and **11** have not been ranked among the top 5. You can change your ranking once, if you want to.

Buttons: change; continue

(same procedure for the second set of 16 cards)

Thank you for submitting your ranking. The experimenter will contact you as soon as the whole group completed their ranking. You will then also receive your payoff.

You can now close the browser window. Thank you for participating.

Choosing a Victim you know

-Introducing communication to the Mobbing Game-

Dmitri Bershadskyy[1], Alexandra Seidel*[1]

[1] Otto-von-Guericke-University Magdeburg, Germany

* corresponding author: Alexandra Seidel (alexandra.seidel@ovgu.de)

Abstract

Mobbing causes severe damages to the victims and is a prime example of antisocial coordination. Using the 'mobbing game' by Abbink and Doğan (2019), we investigate the role of communication and incremental incentives on mobbing in a laboratory experiment. Doing so, we vary the degree of strategic vs. social communication on the one hand and the pecuniary incentives of repeatedly bullying a certain victim on the other hand. Results indicate that incremental incentives increase nomination rates (i.e., attempts to mob another player) and mobbing rates (i.e., successful group coordination to reduce payoffs of one player). In contrast, communication decreases nomination rates without having significant effect on mobbing rates. Further, communication analysis indicates our approach to eliminate strategic communication was successful and can be applied in other setups.

JEL codes

C91, D91

Keywords

Antisocial

Communication

Coordination

Mobbing

Bullying

1. Introduction

Mobbing or bullying undoubtedly causes severe damage to victims. Psychological literature discusses different contributing factors to bullying, such as individual character traits (e.g., psychopathic tendencies, neuroticism), socio-ecological factors (e.g., peer group, family), and demographics (e.g., gender, age) or focuses on specific oversight policies and the effect of bullying on the victims (Swearer & Hymel, 2015, Farrington, 1993; Olweus, 1997, Wolke & Lereya, 2015)[1]. In contrast to the highly detailed psychological perspective, from a standard economic point of view, mobbing can be depicted as a setup where a group of people coordinates on one group member, reducing her utility and simultaneously increasing their own. This formal abstraction is central to the game-theoretic and experimental analysis of Abbink and Doğan (2019) (hereinafter referred to as AD). In the experiment of AD, the players act anonymously. A group coordinates on a victim if three of the four players *nominate* the same other player. Then, the victim's money is split evenly among the perpetrators. In this abstract environment, bullying can be investigated as an antisocial coordination problem without any social context. Thus, the game theoretical equilibria will be found in either full coordination (*mobbing*) or no coordination at all (AD, 2019). Whenever one player is nominated by two others, the remaining player (not the victim) can benefit by joining the bullies in the next period in order to increase his payoff.

However, since social context is an important feature of bullying, we extend the experiment in two directions: communication (since communication is important in bullying) and dynamic benefits (since bullying one victim is more tempting than continuously switching victims in the group). By doing so, we do neither change the structure of the game itself nor the equilibrium. First, we focus on communication and then illustrate the case of incremental incentives for mobbing.

From the perspective of experimental economics, communication can positively affect coordination (Bershadskyy, 2023; Brosig et al., 2003; Cooper et al., 1992; Demichelis & Weibull, 2008; Ellingsen & Östling, 2010) and social preferences (Andreoni & Rao, 2011; Bershadskyy et al., 2023; Brosig et al., 2004; Mohlin & Johannesson, 2008). However, it is unclear how communication affects subjects in an antisocial coordination problem. Introducing communication to AD in a laboratory experiment extends both strands of literature. This leads to our first hypothesis.

[1] In this article, we will use mobbing and bullying synonymously.

1. *Social Communication-Hypothesis:*
    a. Social communication reduces nomination rates (i.e., attempts to coordinate on one victim) as compared to the baseline.
    b. Social communication reduces mobbing rates (i.e., successful coordination on one victim) as compared to the baseline.

The question of how communication changes subjects' behaviour in the mobbing game is even more intriguing as experimental literature indicates that communication can lead to type detection of subjects (He et al., 2017) while individual characteristics could be important for finding a victim in the AD experiment (Abbink & Doğan, 2019). To obtain a more detailed understanding of the communication effect on antisocial coordination, we distinguish between social and strategic communication. Since communication is a complex process which can transmit either social or strategic information (Greiner et al., 2012; Zultan, 2012), we make strategic communication impossible in two of our treatments. The approach is based on findings from Andersson and Wengström (2012) which indicate that people use communication for different purposes depending on the information stages it appears in and is described in more detail in the Experimental Design section. Such a distinction enables the analysis of our second hypothesis.

2. *Strategic Communication-Hypotheses:*
    a. Strategic communication will lead to higher nomination rates than non-strategic communication.
    b. Strategic communication will lead to higher mobbing rates than non-strategic communication.

The second dimension of our experiment deals with incremental incentives of mobbing, i.e. the utility a mob gains when sticking to a certain victim in contrast to continuously changing who is the victim within a group. Following psychological literature that considers reputation (or social dominance) being one of the most central benefits of bullying (Volk et al., 2014), we also introduce a different payoff scheme for the subjects than in AD. This refers to the idea that bullying is a long-term strategy (Volk et al., 2012; Wiertsema et al., 2023). A group can use bullying to gradually deepen bonds and rise in the hierarchy of the general network. Therefore, we introduce an increasing benefit of mobbing. Yet, in line with classical utility functions we assume that the size of such benefits has diminishing returns and induce this structure to the payoff function. Focusing on the role of incentives of mobbing, the original experiment indicates a significant increase of mobbing after increasing the incentives for the players to mob (Abbink & Doğan, 2019). This is in line with other literature finding that non-linear incentives

can lead to higher cooperation rates (Cason & Gangadharan, 2015, 2016). In total, this leads to our third hypothesis.

3. *Incentive Hypothesis:*
    a. Incremental incentives of repeatedly coordinating on one victim lead to higher nomination rates.
    b. Incremental incentives of repeatedly coordinating on one victim lead to higher mobbing rates.

Summing up, we introduce six treatments: one practically replicating one treatment of AD as baseline, one adding incremental incentives to the baseline, and a 2x2 factorial design that varies in payoff structure and timing of communication (see Table 1).

In the next section, we explain the experimental design, the treatments, and the experimental procedure. The third section, we present the results of our work which we then discuss in section four. We close this paper with the concluding remarks in section five.

2. Experimental Design

Our baseline treatment is AD's Medium Treatment. The participants are randomly matched in groups of four and remain in these groups together for the whole session that consists of 20 playing periods. In each period, the participants have to make the decision on whether or not they nominate a victim from the players in their group. If three players nominate the same other player, those three will receive 32 laboratory dollars (1 LD = 0.01 Euro) and the chosen victim will receive no payment. If no player receives three votes, all players are paid 24 LD for this period. If participants do not want to nominate another player, they can abstain. As it is our goal to analyse the coordination ability of communication, we have to reduce chances of accidental coordination. To exclude accidental coordination on one position in the displayed list of players (e.g. always nominating the first player in the list), we intensified the randomization compared to AD by randomizing the order in which the player names are displayed not only for each player but for each player and in every round.

After each period, the voting result will be displayed to all showing *how many* votes each player got but not *who* voted for them. At the end of the experiment, players receive their payoff without having information about the identity of their group members. Additionally, they receive a show up fee of 5.00 Euro.

To simulate a continuously growing reputation due to mobbing as a group, we implement an increasing payoff structure. This means that players who mob one victim in consecutive rounds will receive increasing payoffs for up to seven rounds. In case the mob keeps coordinating on

the same victim for more than seven rounds, they earn the highest mobbing payoff without any further increase. The **incremental treatments** depict the positive yet decreasing marginal utility of mobbing one victim repeatedly.

In addition to the incremental expansion of our experiment, we introduce communication by providing participants the opportunity of a group chat for one minute. To exclude strategic elements of communication, we vary the sequence of communication and game information. If subjects communicate before they learn the rules of the game – communication before instructions (**CBI**) – they by definition cannot discuss any game related strategy. This contrasts communication after instructions (**CAI**).

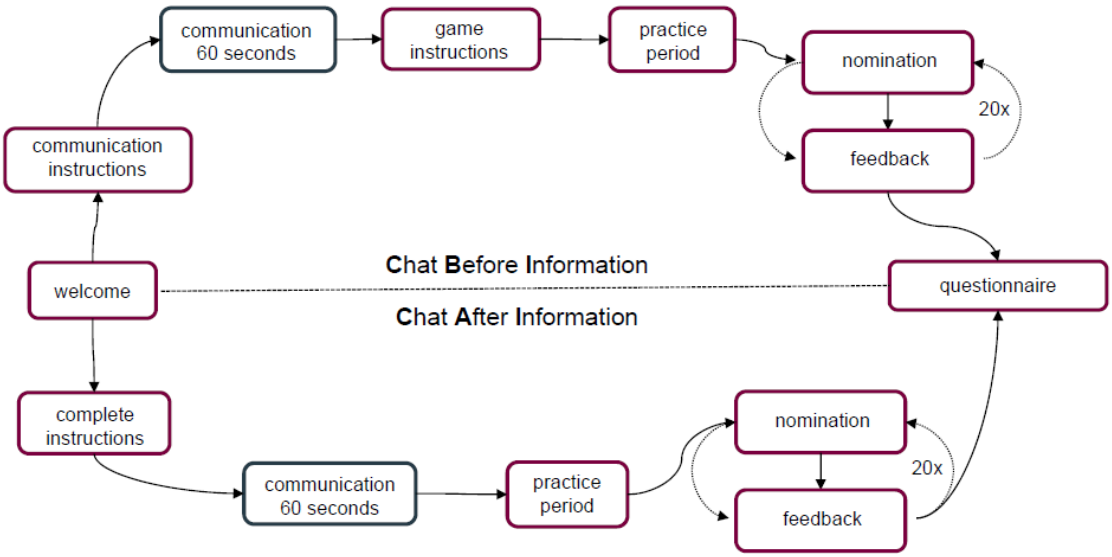
Table 1: Overview over treatments

	no communication	communication before information	communication after information
constant mobbing payoff	**Basic**	**CBI**	**CAI**
incremental mobbing payoff	**IncrBasic**	**IncrCBI**	**IncrCAI**

Experimental procedure

The sessions of this experiment took place in 2021 and 2022. Due to Covid regulations, all sessions were conducted online using *oTree* (Chen et al., 2016). This is in contrast to AD who conducted their sessions offline in the laboratories of Amsterdam and Cologne. Subjects were recruited using *hroot* (Bock et al., 2014). Independent of the treatment, every session started via zoom where participants entered and stayed in the waiting room (where one finds herself alone and without any information about other people in the call) to ensure that they would not receive any information about other participants. After being provided some general instructions about the procedure, they then received the link to start the experiment. The zoom session remained open throughout the experiment so that participants could separately contact the experimenter in case of technical problems or queries.

Figure 1: Procedure of treatments with strategic/non-strategic communication



As depicted in Figure 1, the procedure of the treatments that included communication varied. Please find the instructions for all treatments in the supplemental material.

In the beginning, all participants received general information about the experiment itself and the show-up fee that they will receive. Participants in the **(Incr)CBI** treatments then started communication time without being informed about the decision needed nor payoff structure. This information was displayed after the communication period to ensure that communication cannot be used to discuss the strategy for the game. In contrast, participants of the **(Incr)CAI** treatments got full instructions in the beginning and started then the communication period. After this stage, all treatments had the same structure: the groups played one practice period and proceeded with the twenty payoff relevant periods afterwards.

In the end, participants were asked to answer the dark triad questionnaire and provide some demographic information. They then came back to the zoom session for some final information concerning their payoffs and were paid in private. The sessions lasted between 15 and 25 minutes and the average payoff was 10.11€.

3. Results

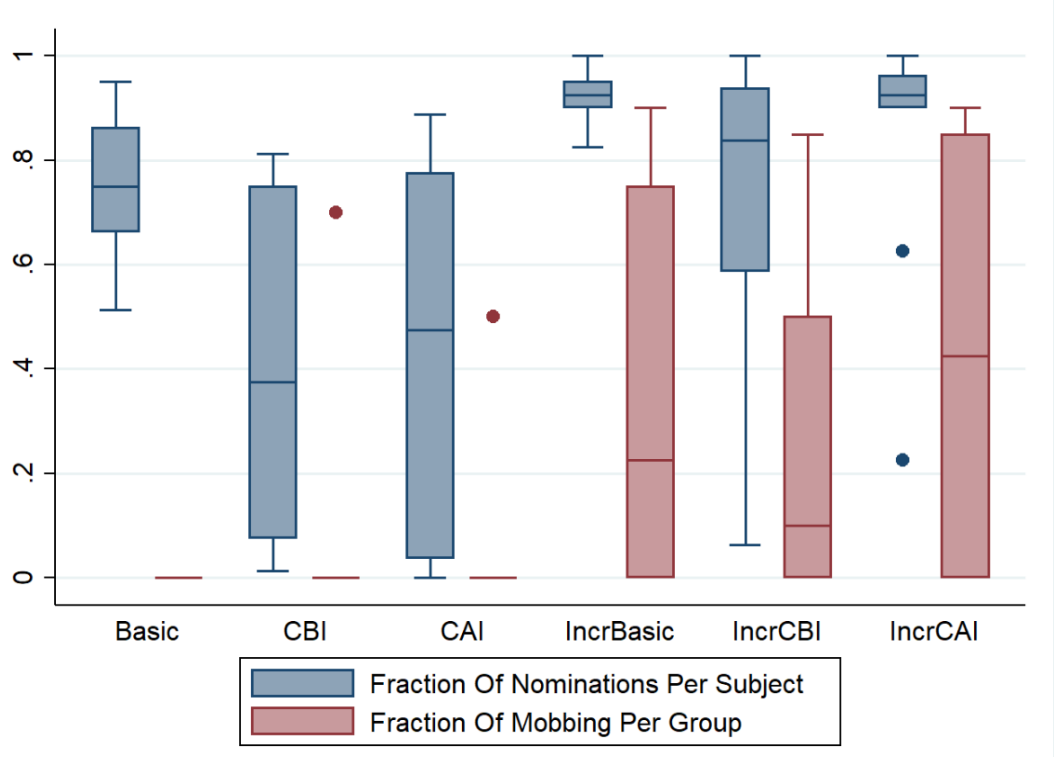
For a first overview, we depicted information about the sample we worked with in Table 2. All participants were students of the Otto-von-Guericke University Magdeburg.

Table 2. Descriptive statistics of the sample

	Basic	CBI	CAI	IncrBasic	IncrCBI	IncrCAI
Number of subjects	40	40	40	40	44	40
Number of groups	10	10	10	10	11	10
Male/Female/Diverse	19/20/0[2]	17/22/1	20/20/0	19/21/0	15/29/0	16/23/1
Average age	24.44	24.63	24.63	23.73	23.55	23.53

Similar to the original article, we focus on two central variables (nomination and mobbing rates)[3]. To test the hypotheses, we calculate group averages over 20 rounds and apply the two-sided Mann-Whitney-Test. This means that every group can have up to 80 nominations (4 subjects * 20 nominations/subject) and up to 18 (20-2) rounds with mobbing. The results are displayed in Figure 2.

Figure 2. Boxplot for nomination and mobbing behaviour with outliers (as dots)



[2] One person did not answer this question.
[3] In line with the definition in AD, mobbing is defined as consecutive nomination of the same player for *at least three periods*.

Note: The figure displays the median fraction of individual nominations and mobbing aggregated on group level. For example, the median fraction of nominations in Basic is 0.75 and the median fraction of rounds with mobbing in IncrCAI is 0.43.

First, concerning the effect of communication on nomination rates, we see a significant decrease for CBI (p=0.0155) and a non-significant one for CAI (p=0.0958) compared to basic. Yet, we see no significant effect on mobbing rates. However, this is not surprising. As shown in Figure 2, we did not observe such mobbing rates as in the AD paper. This means, that the level of mobbing was already very low and thus, communication could not decrease it due to the floor effect. Conducting the same type of analysis for the treatments with incremental payoffs, displays the same picture for nominations. Communication decreases nomination significantly from IncrBasic fro IncrCBI (p=0.0475) and non-significantly for IncrCAI (p=0.9085). Concerning mobbing rates, the displayed behavior is in line with the hypotheses yet on a non-significant level. Compared to IncrBasic mobbing rates decrease to IncrCBI (p=0.6942) but increase to IncrCAI (p=0.7881).

Second, concerning the distinction between strategic and non-strategic communication, we do not see statistically different results. Independent of the incentive structure (CBI vs. CAI or IncrCBI vs. IncrCAI), nomination and mobbing rates are similar. Since this could be either due to subjects not having strategic communication in (Incr)CAI or due to strategic communication being insufficient to provide coordination, this needs additional analysis.

To further investigate this issue, we analysed the chat protocols of the groups. Using two independent coders and a predefined code book we assessed several variables for all 41 chats: *number of people communicating, greeting, attempt for strategic communication, coordination on one player, agreement to nominate nobody, small talk*[4]. For all variables, the coders achieved a high interrater agreement rate (>95%). We observe that by changing the timing of communication, we can exclude the strategic element of communication to a certain degree. The coders found significantly more attempts of strategic communication in CAI and IncrCAI as compared to CBI and IncrCBI (MW-Test, p=0.0008)[5]. However, we did not observe a lot of successful coordination on one specific player. The summary of chat classification is displayed in Table 3.

[4] Please find the code book in the supplementary data.

[5] Note that the percentage of groups classified as having strategic communication is not zero because of our broad definition of strategic communication in the code book (see supplementary information). Therefore, coders labeled groups discussing the broad goal to "maximize the profit" as strategic communication. Coders were asked to mark a conversation as strategic, if one or more members have made statements on the joint approach.

Table 3. Summary of chat classification

	CBI	**CAI**	**IncrCBI**	**IncrCAI**
Number of chat protocols	10	10	11	10
Greeting	1.00	0.90	0.95	0.80
Attempted strategic communication	0.35	0.70	0.00	0.65
Coordination on one player	0.00	0.05	0.00	0.05
Agreement on not nominating anyone	0.05	0.20	0.00	0.00
Small Talk	0.25	0.06	0.27	0.10

Note: The numbers depict how often coders classified a chat as 1 given the description in the codebook. For example, in CBI only in one of ten chats did one of the two coders stated there was an agreement on not nominating anyone.

The highest rate of agreeing not to vote for anyone has been found in CAI. However, we see that this vanished in the incremental CAI treatment although the rate of attempted strategic conversation was similar. Also, coders found that there was less small talk in the CAI treatments.

In total, the results indicate that subjects did attempt strategic communication, yet were not successful in doing so. This is a possible reason why we did not observe significant increases in nomination or mobbing rates with strategic communication.

Third, concerning the incremental incentives, we are able to confirm and to extend prior findings by AD in the context of communication. Introducing incremental incentives increased nomination rates significantly (MW-Tests) in all three cases: comparing Basic to IncrBasic ($p=0.0050$), CBI to IncrCBI ($p=0.0201$) and CAI to IncrCAI ($p=0.0040$). Similar findings hold for the analysis of mobbing rates comparing Basic to IncrBasic ($p=0.0019$), CBI to IncrCBI ($p=0.0139$) and CAI to IncrCAI ($p=0.0072$).

To investigate how robust our findings are, we implement several types of regression analyses (see Table 4). Please note that as there was absolutely no mobbing in the Basis treatment, it cannot be used a base in a regression model and test statistics cannot be calculated. This is the reason we chose CBI as reference to display all regressions.

First, in models (1-3) we use a tobit regression (from 0 to 20) to analyse individual number of nominations clustered at the group level. The results indicate that the treatment effect remains highly significant after controlling for demographics and psychological parameters from the

dirty dozen (Jonason & Webster, 2010). We observe higher nomination rates in the **Basic** treatment than in **CBI**. This supports our *Social Communication-Hypothesis* (1a). Additionally, we find support for our *Incentives-Hypothesis* (3a) since we find significantly higher nomination rates in the **incremental treatments**[6]. Participants with a machiavellian profile tend to nominate more often.

Second, we investigate the nomination behaviour (yes or no) in the first round as this constitutes a higher level of independence of the observation. The results from the logit models (4-6) show strong treatment effects. There are significantly more participants in **Basic** and the **incremental treatments** that nominate someone in the very first round. People with higher age nominate less often in the first round – an effect that remains significant independent from the dirty dozen. Participants with psychopathic profile are less likely to nominate in this first round.

Finally, we shed some light on the mobbing behaviour using tobit model (from 0 to 18). Due to the floor effect, we see a negative yet non-significant treatment effect to the **Basic** treatment. However, there is significantly more mobbing in the **incremental treatments**, which again supports our *Incentive Hypothesis* (3b), although the effect decreases after including our control variables. We find no significant differences between CBI and CAI in all of our regressions and thus find no support for our *Strategic Communication-Hypothesis* (2a, b).

[6] We use CBI as comparison but the treatments effects are also significant for a comparison between Basic and IncrBasic or CAI and IncrCAI.

Table 4. Summary of the regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Number of Nominations	Number of Nominations	Number of Nominations	Nomination Round 1	Nomination Round 1	Nomination Round 1	Mobbing	Mobbing	Mobbing
Basic	9.265***	9.202***	8.803***	1.170**	1.120**	1.054**	-2.510	-2.471	-2.708
	(3.131) [0.003]	*(3.135) [0.004]*	*(3.095) [0.005]*	*(0.532) [0.028]*	*(0.537) [0.037]*	*(0.526) [0.045]*	*(0)*	*(0)*	*(0)*
CAI	0.193	0.161	0.383	-0.205	-0.246	-0.243	-0.049	-0.019	-0.069
	(4.400) [0.965]	*(4.389) [0.971]*	*(4.294) [0.929]*	*(0.623) [0.742]*	*(0.609) [0.686]*	*(0.585) [0.678]*	*(0.345) [0.888]*	*(0.332) [0.954]*	*(0.322) [0.830]*
IncrBasic	16.32***	15.98***	15.05***	3.145***	3.090***	3.013***	0.768**	0.703**	0.527*
	(3.259) [0.000]	*(3.252) [0.000]*	*(3.125) [0.000]*	*(0.769) [0.000]*	*(0.761) [0.000]*	*(0.738) [0.000]*	*(0.303) [0.014]*	*(0.291) [0.019]*	*(0.299) [0.084]*
IncrCBI	9.014**	8.773**	8.366**	0.860*	0.824	0.754	0.674**	0.572**	0.429
	(3.916) [0.022]	*(3.876) [0.025]*	*(3.813) [0.030]*	*(0.521) [0.099]*	*(0.528) [0.099]*	*(0.528) [0.119]*	*(0.299) [0.028]*	*(0.284) [0.049]*	*(0.270) [0.118]*
IncrCAI	13.60***	13.27***	12.72***	2.147***	2.140***	2.230***	0.860***	0.761**	0.614**
	(4.089) [0.001]	*(4.089) [0.001]*	*(3.925) [0.001]*	*(0.697) [0.002]*	*(0.768) [0.005]*	*(0.775) [0.004]*	*(0.303) [0.006]*	*(0.288) [0.011]*	*(0.275) [0.030]*
Gender		-0.739	-0.358		-0.394	-0.577*		0.0892	0.0824
		(1.443) [0.609]	*(1.526) [0.815]*		*(0.294) [0.180]*	*(0.337) [0.087]*		*(0.336) [0.791]*	*(0.369) [0.824]*
Age		-0.319*	-0.351*		-0.0811**	-0.0836**		-0.0804*	-0.105**
		(0.193) [0.099]	*(0.196) [0.075]*		*(0.038) [0.034]*	*(0.037) [0.023]*		*(0.042) [0.063]*	*(0.042) [0.016]*
Machiavellianism			1. 353***			0.212			0.190
			(0.519) [0.009]			*(0.147) [0.149]*			*(0.136) [0.168]*
Psychopathy			-0.131			-0.257**			0.042
			(0.615) [0.831]			*(0.127) [0.043]*			*(0.138) [0.760]*
Narcissism			0.151			0.0102			-0.170*
			(0.419) [0.719]			*(0.102) [0.920]*			*(0.095) [0.080]*
Constant	7.432**	15.71***	12.99**	-0.201	2.025*	2.226**	-0.513*	1.424	2.068*
	(2.934) [0.012]	*(5.032) [0.002]*	*(5.710) [0.024]*	*(0.376) [0.593]*	*(1.054) [0.055]*	*(1.110) [0.041]*	*(0.263) [0.056]*	*(1.047) [0.179]*	*(1.135) [0.074]*
Observations	244	243	243	244	243	243	61	61	61

Note: *The first three models use number of nominations of individuals as dependent variable in a Tobit regression. Models 4-6 only investigate whether subjects nominated someone in the first round using the logit model. All models 1-6 are clustered at group level. Models 7-9 investigate mobbing behaviour which can only be analysed on group level. Therefore, all control variables are aggregated at the group level (i.e. average age in the group). Further, for all models the table displays the treatment effects compared to CBI, which is due to the Basic treatment having no mobbing. This further implies that the coefficients for e.g. IncrCAI cannot be used for our hypothesis tests as they include two changes (incentives and type of communication). Nonetheless, they are informative on the total difference between the treatments.*

Robust standard errors in parentheses, p-values are in brackets.

**** p<0.01, ** p<0.05, * p<0.1*

4. Discussion

In line with prior research, we show that the effect of higher incentives on mobbing behaviour is stronger than the effect of communication or demographic attributes. As we have seen, the mobbing rates in our baseline treatment that held as replication of AD, are zero. One possible reason can be the hard coordination on one victim because in contrast to AD, we randomized the order of appearance of players' names in each period and for each player. A lack of communication will not hold for explanation since in the AD experiment, participants were not able to communicate either. We stress that conducting the experiment online due to Covid regulations is another deviation from the original setup of AD. However, it seems implausible to assume that the higher level of anonymity of our online experiment leads to less mobbing given prior literature on specifically cyber mobbing (Barlett & Gentile, 2012; Moore et al., 2012) or more general economic literature (Bohnet & Frey, 1999; Hoffman et al., 1994). An alternative solution would be to consider the initial payoffs in our basic treatment as insufficient to induce mobbing. Compared to AD we have similar financial incentives. While the total payoff is slightly lower than in AD, the hourly wage is slightly higher. Yet, we can not rule out completely that this caused the difference in mobbing.

Because of observing no mobbing in our baseline treatment, a reduction of mobbing rates via introducing communication is impossible for constant incentives. For incremental incentives we see the anticipated difference between strategic and non-strategic communication. Yet, the results are not significant. However, we do see a decrease in nomination rates that might indicate that communication supports prosocial behaviour more than antisocial coordination. Further, we indicate that our results might differ if it was possible to exclude some members from communication as in (Abbink et al., 2022; Bershadskyy et al., 2023). However, this requires additional research.

Finally, the effect of incremental incentives is in line with the original findings from AD, previous research on non-linear incentives (Cason & Gangadharan, 2015, 2016), and the findings that the higher the incentives for acting in an anti-social way, the more often this behaviour will be observed (Bolle et al., 2014; Charness et al., 2014; Harbring et al., 2007).

Still, we stress two major limitations of our study. First, the low total number of observations. Concerning the effect of incremental incentives, this threat is limited, as the results appear to be very robust independent of the statistical method and the experimental condition of communication. Concerning communication, our results are more ambiguous. A larger number of observations could have helped investigate whether the decreases in mobbing rates induced

by communication in incremental treatments differ depending on the type of communication. Since the effects went in the anticipated direction, future analysis may be valuable. Second, our goal was to introduce more structural and social aspects to the original experiment on mobbing. In doing so, we implemented a payoff structure in which mobbing the same person has a benefit compared to alternating through all group members and where mobbing starts small and increases in intensity. Both changes depict certain characteristics of mobbing in the literature. Yet, in line with the original study by AD, we model mobbing purely monetarily, meaning that there may be other psychological and or social elements missing in our setup. Nonetheless, within this limitation, our results indicate the importance of structural payoffs from mobbing.

5. Concluding remarks

Despite not being able to replicate the mobbing rates found by AD in one of our treatments, our findings shed light on the role of communication and incentives in the mobbing game. Our results indicate that communication decreases nomination rates, but is not as effective as personal benefits are in increasing nomination and mobbing rates. Our approach to exclude strategic communication worked. Yet, distinguishing between social and strategic communication, we observe that subjects did not use the strategic communication channel to effectively coordinate in an antisocial way. One plausible explanation may be that the social aspects of communication counteract the efforts to use communication for anti-social coordination. In conclusion, our results show that communication in an antisocial coordination problem can have ambiguous effects that need to be disentangled in future investigations.

Acknowledgements

The authors thank Gönül Doğan and Klaus Abbink for their support in replicating the study, Joachim Weimann and Abdolkarim Sadrieh for their insights and comments, Jannik Greif for assistance in conducting the experiment, and two anonymous referees for helpful reports. We also thank the Faculty of Management and Economics of the Otto-von-Guericke University Magdeburg for funding this research project.

List of References

Abbink, K., & Doğan, G. (2019). How to choose your victim. *Games and Economic Behavior, 113*, 482–496. https://doi.org/10.1016/j.geb.2018.10.006

Abbink, K., Dong, L., & Huang, L. (2022). Talking Behind Your Back: Communication and Team Cooperation. *Management Science*, *68*(7), 5187–5200. https://doi.org/10.1287/mnsc.2021.4143

Andersson, O., & Wengström, E. (2012). Credible communication and cooperation: Experimental evidence from multi-stage Games. *Journal of Economic Behavior and Organization, 81*(1), 207–219. https://doi.org/10.1016/j.jebo.2011.10.002

Andreoni, J., & Rao, J. M. (2011). The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of Public Economics*, *95*(7–8), 513–520. https://doi.org/10.1016/j.jpubeco.2010.12.008

Barlett, C. P., & Gentile, D. A. (2012). Attacking others online: The formation of cyberbullying in late adolescence. *Psychology of Popular Media Culture*, *1*(2), 123–135. https://doi.org/10.1037/a0028113

Bershadskyy, D. (2023). Reverberation effect of communication in a public goods game. *PLOS ONE*, *18*(2), e0281633. https://doi.org/10.1371/JOURNAL.PONE.0281633

Bershadskyy, D., Sachs, F. E., & Weimann, J. (2023). Collective bargaining in a shrinking group game: The role of information and communication. *Journal of Economic Behavior & Organization*, *209*, 391–410. https://dvoi.org/10.1016/J.JEBO.2023.03.009

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review*, *71*, 117–120. https://doi.org/10.1016/J.EUROECOREV.2014.07.003

Bohnet, I., & Frey, B. S. (1999). Social Distance and Other-Regarding Behavior in Dictator Games: Comment. *American Economic Review*, *89*(1), 335–339. https://doi.org/10.1257/aer.89.1.335

Bolle, F., Tan, J. H. W., & Zizzo, D. J. (2014). Vendettas. *American Economic Journal: Microeconomics*, *6*(2), 93–130. https://doi.org/10.1257/mic.6.2.93

Brosig, J., Weimann, J., & Ockenfels, A. (2003). The Effect of Communication Media on Cooperation. *German Economic Review*, *4*(2), 217–241. https://doi.org/10.1111/1468-0475.00080

Brosig, J., Weimann, J., & Yang, C.-L. (2004). Communication, Reputation, and Punishment in Sequential Bargaining Experiments. *Journal of Institutional and Theoretical*

Economics (JITE) / Zeitschrift Für Die Gesamte Staatswissenschaft, *160*(4), 576–606. https://doi.org/10.2307/40752481

Cason, T. N., & Gangadharan, L. (2015). Promoting cooperation in nonlinear social dilemmas through peer punishment. *Experimental Economics*, *18*(1), 66–88. https://doi.org/10.1007/s10683-014-9393-0

Cason, T. N., & Gangadharan, L. (2016). Swords without covenants do not lead to self-governance. *Journal of Theoretical Politics*, *28*(1), 44–73. https://doi.org/10.1177/0951629814539199

Charness, G., Masclet, D., & Villeval, M. C. (2014). The Dark Side of Competition for Status. *Management Science*, *60*(1), 38–55. https://doi.org/10.1287/mnsc.2013.1747

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1992). Communication in Coordination Games. *The Quarterly Journal of Economics*, *107*(2), 739–771. https://doi.org/10.2307/2118488

Demichelis, S., & Weibull, J. W. (2008). Language, meaning, and games: A model of communication, coordination, and evolution. *American Economic Review*, *98*(4), 1292–1311. https://doi.org/10.1257/aer.98.4.1292

Ellingsen, T., & Östling, R. (2010). When Does Communication Improve Coordination? *American Economic Review*, *100*(4), 1695–1724. https://doi.org/10.1257/aer.100.4.1695

Greiner, B., Güth, W., & Zultan, R. (2012). Social communication and discrimination: A video experiment. *Experimental Economics*, *15*(3), 398–417. https://doi.org/10.1007/s10683-011-9305-5

Harbring, C., Irlenbusch, B., Kräkel, M., & Selten, R. (2007). Sabotage in Corporate Contests – An Experimental Analysis. *International Journal of the Economics of Business*, *14*(3), 367–392. https://doi.org/10.1080/13571510701597445

He, S., Offerman, T., & van de Ven, J. (2017). The Sources of the Communication Gap. *Management Science*, *63*(9), 2832–2846. https://doi.org/10.1287/mnsc.2016.2518

Hoffman, E., McCabe, K., Shachat, K., & Smith, V. (1994). Preferences, Property Rights, and Anonymity in Bargaining Games. *Games and Economic Behavior*, *7*(3), 346–380. https://doi.org/10.1006/game.1994.1056

Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, *22*(2), 420–432. https://doi.org/10.1037/A0019265

Mohlin, E., & Johannesson, M. (2008). Communication: Content or relationship? *Journal of Economic Behavior and Organization*, *65*(3–4), 409–419. https://doi.org/10.1016/j.jebo.2005.12.003

Moore, M. J., Nakano, T., Enomoto, A., & Suda, T. (2012). Anonymity and roles associated with aggressive posts in an online forum. *Computers in Human Behavior*, *28*(3), 861–867. https://doi.org/10.1016/j.chb.2011.12.005

Volk, A. A., Camilleri, J. A., Dane, A. V., & Marini, Z. A. (2012). Is Adolescent Bullying an Evolutionary Adaptation? *Aggressive Behavior*, *38*(3), 222–238. https://doi.org/10.1002/AB.21418

Volk, A. A., Dane, A. V., & Marini, Z. A. (2014). What is bullying? A theoretical redefinition. *Developmental Review*, *34*(4), 327–343. https://doi.org/10.1016/J.DR.2014.09.001

Wiertsema, M., Vrijen, C., van der Ploeg, R., Sentse, M., & Kretschmer, T. (2023). Bullying perpetration and social status in the peer group: A meta-analysis. *Journal of Adolescence*, *95*(1), 34–55. https://doi.org/10.1002/JAD.12109

Zultan, R. (2012). Strategic and social pre-play communication in the ultimatum game. *Journal of Economic Psychology*, *33*(3), 425–434. https://doi.org/10.1016/J.JOEP.2011.12.009

Appendix

Instruktionen "Choosing a Victim you know" Deutsch

Basis

Willkommen zu diesem Experiment zur Entscheidungsbildung.
Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.
Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.
Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.
Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.
Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.
Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.
Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.
Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!

Button: <<Weiter>>

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Es gibt zwei verschiedene Szenarien, die sich aus den Entscheidungen der Spieler ergeben können.
Entweder wählen drei Gruppenmitglieder dieselbe Person, dann erhalten diese drei 32 Punkte während die nominierte Person 0 Punkte erhält (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In diesem Fall erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Runde.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.

Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.

Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.

Button: <<Ich habe die Instruktionen verstanden>>

IncrBasis

Willkommen zu diesem Experiment zur Entscheidungsbildung.
Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.
Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.
Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.
Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.
Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.
Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.
Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.
Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!

Button: <<Weiter>>

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Entweder wählen drei Gruppenmitglieder dieselbe Person (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In Szenario 1 erhält die nominierte Person 0 Punkte. Die anderen drei Personen erhalten 32 Punkte.

Wenn sich drei Personen in aufeinanderfolgenden Runden stets für die gleiche nominierte Person entscheiden, steigt ihre Auszahlung wie folgt:
Runde 1: 32 Punkte, Runde 2: 38 Punkte, Runde 3: 44 Punkte, Runde 4: 49 Punkte, Runde 5: 54 Punkte, Runde 6: 58 Punkte, Runde 7: 61 Punkte und ab Runde 8 63 Punkte.

Beachten Sie: sobald sich drei Personen auf eine andere Person koordinieren, fällt der erhaltene Betrag auf 32 Punkte zurück und die Steigerung beginnt Runde für Runde mit der neuen nominierten Person.
In Szenario 2 erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Periode.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.

Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.

Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.
Button: <<Ich habe die Instruktionen verstanden>>

CBI

Teil 1:

Willkommen zu diesem Experiment zur Entscheidungsbildung.
Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.
Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.
Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.
Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.
Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.
Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.
Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.
Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!

Button: <<Weiter>>

Teil 2 (nach der Kommunikation):

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Es gibt zwei verschiedene Szenarien, die sich aus den Entscheidungen der Spieler ergeben können.
Entweder wählen drei Gruppenmitglieder dieselbe Person, dann erhalten diese drei 32 Punkte während die nominierte Person 0 Punkte erhält (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In diesem Fall erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Runde.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.

Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.

Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.

Button: <<Ich habe die Instruktionen verstanden>>

CAI

Willkommen zu diesem Experiment zur Entscheidungsbildung.
Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.
Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.
Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.
Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.
Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.
Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.
Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.
Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!

Button: <<Weiter>>

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Es gibt zwei verschiedene Szenarien, die sich aus den Entscheidungen der Spieler ergeben können.
Entweder wählen drei Gruppenmitglieder dieselbe Person, dann erhalten diese drei 32 Punkte während die nominierte Person 0 Punkte erhält (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In diesem Fall erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Runde.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.

Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.

Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.

Button: <<Ich habe die Instruktionen verstanden>>

IncrCBI

Teil 1

Willkommen zu diesem Experiment zur Entscheidungsbildung.

Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.

Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.

Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.

Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.

Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.

Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.

Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.

Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!

Button: <<Weiter>>

Teil 2 (after communication)

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Es gibt zwei verschiedene Szenarien, die sich aus den Entscheidungen der Spieler ergeben können. Entweder wählen drei Gruppenmitglieder dieselbe Person (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In Szenario 1 erhält die nominierte Person 0 Punkte. Die anderen drei Personen erhalten 32 Punkte.

Wenn sich drei Personen in aufeinanderfolgenden Runden stets für die gleiche nominierte Person entscheiden, steigt ihre Auszahlung wie folgt:
Runde 1: 32 Punkte, Runde 2: 38 Punkte, Runde 3: 44 Punkte, Runde 4: 49 Punkte, Runde 5: 54 Punkte, Runde 6: 58 Punkte, Runde 7: 61 Punkte und ab Runde 8 63 Punkte.

Beachten Sie: sobald sich drei Personen auf eine andere Person koordinieren, fällt der erhaltene Betrag auf 32 Punkte zurück und die Steigerung beginnt Runde für Runde mit der neuen nominierten Person.
In Szenario 2 erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Periode.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.

Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.
Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.

Button: <<Ich habe die Instruktionen verstanden>>

IncrCAI

Willkommen zu diesem Experiment zur Entscheidungsbildung.
Bitte lesen Sie zunächst die Instruktionen aufmerksam durch. Wenn Sie eine Frage haben, wenden Sie sich bitte via Chat an uns und wir werden Ihnen helfen.
Wenn alle die Instruktionen verstanden haben, beginnt das Experiment.
Zu Übungszwecken beginnt das Experiment mit einer Testrunde, die nicht auszahlungsrelevant ist.
Erst danach beginnt das Experiment.

Wenn Sie die Anweisungen sorgfältig befolgen, können Sie Geld verdienen. Während des Experiments wird Ihr Verdienst in Punkten angegeben.
Ihr Gesamtverdienst ist die Summe aller Punkte, die Sie in allen Runden verdienen.
Nach dem Experiment wird Ihr Verdienst in Geld umgerechnet, wobei 1 Punkt 1 Eurocent entspricht.
Zusätzlich erhalten Sie eine Aufwandsentschädigung von 5 Euro.
Sie können Ihre Auszahlung in bar oder via Paypal erhalten – bitte kontaktieren Sie den Experimentator dazu im Anschluss an das Experiment. Viel Erfolg!
Button: <<Weiter>>

In diesem Experiment erhalten Sie eine Kennzeichnung als Spieler M, Spieler T, Spieler P oder Spieler G. Ihre Spielerkennzeichnung wird zu Beginn der Übungsphase zufällig bestimmt und bleibt während des gesamten Experiments gleich. Ihnen wird stets angezeigt, welcher Spieler Sie sind. Diese Bezeichnungen haben keine Bedeutung und daher keinen Bezug zu den Aktionen, die Sie im Experiment ausführen werden. Auch die Zusammensetzung Ihrer Gruppe bleibt während des gesamten Experiments gleich.

Sie werden nicht wissen, wer in Ihrer Gruppe ist. Das Experiment besteht aus 20 Runden. In allen 20 Runden werden Sie vor der gleichen Aufgabe stehen. Die Aufgabe besteht darin, eine Entscheidung darüber zu treffen, ob Sie einen Spieler in Ihrer Gruppe wählen. In jeder Periode werden Sie und jedes andere Mitglied Ihrer Gruppe entscheiden, ob Sie ein Gruppenmitglied auswählen. Sie können entweder ein Gruppenmitglied oder niemanden auswählen, sich also enthalten.
Sich selbst wählen ist nicht möglich.

Es gibt zwei verschiedene Szenarien, die sich aus den Entscheidungen der Spieler ergeben können.Entweder wählen drei Gruppenmitglieder dieselbe Person (Szenario 1) oder niemand erhält drei Stimmen (Szenario 2).
In Szenario 1 erhält die nominierte Person 0 Punkte. Die anderen drei Personen erhalten 32 Punkte.

Wenn sich drei Personen in aufeinanderfolgenden Runden stets für die gleiche nominierte Person entscheiden, steigt ihre Auszahlung wie folgt:
Runde 1: 32 Punkte, Runde 2: 38 Punkte, Runde 3: 44 Punkte, Runde 4: 49 Punkte, Runde 5: 54 Punkte, Runde 6: 58 Punkte, Runde 7: 61 Punkte und ab Runde 8 63 Punkte.

Beachten Sie: sobald sich drei Personen auf eine andere Person koordinieren, fällt der erhaltene Betrag auf 32 Punkte zurück und die Steigerung beginnt Runde für Runde mit der neuen nominierten Person.
In Szenario 2 erhält jeder Spieler eine Auszahlung von 24 Punkten für diese Periode.

Ein Spieler kann sich nicht selbst nominieren, aber kann sich enthalten. Nach jeder Runde wird jeder in der Gruppe darüber informiert, wie oft jeder Spieler nominiert wurde, aber nicht darüber, wer die Nominierung vorgenommen hat.
Nach den 20 Runden wird Ihr Gesamtergebnis die Summe aller Punkte sein, die Sie über alle Runden hinweg verdienen.

Für jeden Punkt, den Sie bei dem Experiment verdienen, erhalten Sie 1 Eurocent.

Button: <<Ich habe die Instruktionen verstanden>>