Essays on Personnel Economics and Incentives

Zur Erlangung des akademischen Grades eines

Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der KIT-Fakultät für Wirtschaftswissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

André Haas

Tag der mündlichen Prüfung	15. Juli 2019
Referentin	Professor Dr. Petra Nieken
Korreferent	Professor Dr. Alexander Woll

Karlsruhe, Juli 2019

Contents

List of Tables

List of Figures

1 Introduction

1.1 Overview

The essential reason why firms exist is a cost advantage of using resources within the firm over negotiating each transaction on external markets (Coase, 1937). Regarding personnel, efficiency gains can be realized if two or more individuals work together in a team. Team production is typically characterized by an overall output which exceeds the sum of individual outputs if the production function is not separable (Alchian and Demsetz, 1972). Therefore, Lazear and Oyer (2012, p. 506) conclude that "[p]erhaps the greatest value of the firm is that it provides a mechanism for people to work together and take advantage of complementarities in their skills and interests".

This thesis comprises experimental studies on three different aspects which are prevalent in the workplace if individuals are required to work together: cooperation, coordination, and sabotage in tournaments. The former ones focus on individuals' willingness to align common interests which may conflict with individual incentives. Situations which require individuals to cooperate typically provide incentives to free ride at the same time, i.e., due to the fact that individuals who do not contribute in the production stage cannot be excluded from the benefits of the joint surplus, it is favorable from an individual perspective to rely on contributions by others (Andreoni, 1988). Coordination, on the other hand, often depends on a weakest link mechanism, i.e., the lowest individual effort level is decisive for the total output of the group (Hirshleifer, 1983). This situation comprises a set of equilibria in which all individuals choose the same level of effort. These equilibria can be Pareto-ranked; the equilibrium in which all individuals choose the highest possible level of effort is not Pareto-dominated by any of the other equilibria. Yet, strategic uncertainty about other individuals' actions can result in coordination failure which describes a situation in which the Pareto-dominant equilibrium is not attained—or no equilibrium at all (Van Huyck *et al.*, 1990). In contrast to situations focusing on cooperation and coordination, tournaments are used to provide incentives when individual performance in absolute terms is difficult to measure and, hence, is replaced by relative performance evaluation. While rank-order tournaments exhibit both, theoretically and practically, desirable properties (Lazear and Rosen, 1981), they also provide adverse incentives. In particular, not only increasing one's own effort increases an individual's probability of winning the tournament, but also decreasing the output of the opponent by means of sabotage (Lazear, 1989; see also Sheremeta, 2016).

Cooperation Cooperation appeals to individuals' willingness to forgo private benefits to increase the social return. This situation is frequently stylized in the linear public good game. Public goods are characterized by non-rivalry in consumption and non-excludability. The latter property refers to the fact that an individual can benefit from the public good once it is provided without having contributed to its funding before. This kind of free riding behavior arises when private costs exceed private benefits from the provision of the public good. As a consequence,

it is individually rational not to contribute to the funding of public goods assuming standard preferences of money-maximizing individuals and rely on others instead—which eventually results in a situation in which the public good is not provided at all (Olson, 1965). To study this social dilemma in a controlled environment in the laboratory, $n \geq 2$ individuals indexed i form a group.[1] Each individual receives an endowment q_i which can be allocated to the individual's private account or a collective account. The sum of contributions to the collective account by all individuals is multiplied by some factor α and distributed to each individual. Thus, individual i's payoff π_i is given by

$$\pi_i = q_i - c_i + \alpha \sum_{i=1}^{n} c_i,$$

where $c_i \leq q_i$ denotes the individual's contribution to the collective account.[2] α is the marginal per-capita return (MPCR) which represents the relative gain from an individual's contribution of one unit to the collective account as compared to a contribution to her private account. If $\frac{1}{n} < \alpha < 1$, a social dilemma arises as the individual return from a contribution to the collective account is smaller than the associated costs for the individual.[3] Yet, contributing to the collective account is socially beneficial as all individuals—irrespective whether or not and how much they contributed themselves—receive a return $\alpha \sum_{i=1}^{n} c_i$ from the collective account which in total exceeds the individual costs of contributing since $n \cdot \alpha > 1$. Thus, the Nash equilibrium is characterized by total free riding, i.e., all individuals keep their endowment, whereas the Pareto-efficient social optimum is attained if all individuals contribute their entire endowment to the collective account (Andreoni, 1988). In one-shot public good games and the first period of repeated public good games, individuals frequently contribute 40–60% of their endowment to the collective account. Contributions typically decline over the subsequent periods in repeated public good games and approximate the Nash equilibrium with a particularly sharp decline in the final period (Ledyard, 1995; Chaudhuri, 2011). This raises the question why individuals' aggregate behavior initially corresponds neither to the Nash equilibrium nor the social optimum. One explanation proposed by Andreoni (1988) is the learning hypothesis, i.e., individuals do not fully understand the structure of the game and the associated consequences for their payoffs right at the beginning. Instead, they need some experience with the game to learn its incentives. As cooperative behavior is usually observed in public good games with a fixed number of periods until shortly before the final periods, he introduces a second argument which refers to strategic behavior. According to this hypothesis, it can be beneficial to adhere to cooperative behavior to some extent even after learning about the incentive to free ride for some repetitions prior to the final periods. In the related experiment, however, he does not find support for either of these hypotheses. Another explanation for cooperative behavior in public good games despite its incentives to free ride which is frequently discussed in the literature addresses individuals' altruism; in general, there is only weak evidence for this explanation. Therefore, Andreoni (1989, 1990) introduces the concept of "warm-glow" giving which relates impure altruism. In contrast to the notion of pure altruism which increases the individuals' utility as contributing increases the size of the public good for the entire group, impure altruism focuses solely on the act of giving. This extension is by and large consistent with frequently observed behavior in

situations which exhibit similar incentives as public good games, e.g., fundraising for charity. Another explanation for cooperative behavior in public good games is that individuals may not understand the instructions and, as a consequence, split their endowment arbitrarily between the private and the collective account. Indeed, Andreoni (1995) finds evidence that a substantial share of positive contributions to the collective account in public good games occurs due to confusion about the structure of the game. The previous results are largely confirmed by Palfrey and Prisbrey (1997) who find that pure altruism is not decisive for individuals' contributions in the public good game. Instead, impure altruism as reflected by warm-glow giving and ignorance about the incentives of the game explain contributions to a large extent. Whereas these results are based on aggregate behavior, Fischbacher *et al.* (2001) identify different types of contributors in a one-shot public good game using the strategy method (Selten, 1967). Individuals who indicate higher contributions if other group members also contribute higher amounts to the collective account are deemed conditional cooperators which applies to half of the individuals in the study.[4] While contributions of these individuals are highly positively correlated with others' contributions, they do not always match the contributions of the other group members exactly. In fact, negative deviations by conditional cooperators occur more frequently than contributions above the average of the other group members. Approximately one third of the individuals are free riders, i.e., they behave as predicted by standard theory assuming money-maximizing individuals and keep their entire endowment for themselves. The third category of individuals contribute according to a hump-shaped or triangular pattern, i.e., they behave like conditional cooperators up to a certain point and steadily decrease their contributions beyond this threshold. This behavior is observed in approximately one of seven individuals in the study by Fischbacher *et al.* (2001). The remaining individuals cannot be allocated to one of these types.

Coordination In contrast to cooperation which offers individuals incentives to free ride off others' contributions, coordination requires individuals to align their actions. This situation is frequently represented by the minimum effort game (Van Huyck *et al.*, 1990). Each of n individuals indexed i who form a group chooses an effort level e_i at cost b per unit. The return for all individuals is determined by the minimum effort level of the entire group multiplied by some factor a. Hence, individual i's payoff π_i is given by

$$\pi_i = a \min\{e_i, e_{-i}\} - be_i$$

with $a > b > 0$ and $e_{-i} = \min\{e_1, e_2, \ldots, e_{i-1}, e_{i+1}, \ldots, e_n\}$ as minimum effort level of all individuals except i. There are several equilibria in this game in which all individuals choose the same effort level. Although these equilibria can be Pareto-ranked with the equilibrium in which all individuals choose the highest possible effort level as Pareto-dominant option, it is not straightforward to predict which equilibrium will actually be chosen. All outcomes in which the Pareto-dominant equilibrium is not attained—including other equilibrium points—are considered coordination failure. The concepts of payoff dominance and risk dominance (Harsanyi and Selten, 1988) allow to identify choices which may be salient for all individuals of a group in absence of com-

munication. The former alternative points at the highest possible effort level as coordinating on this level yields the highest payoff for all individuals. In contrast, risk dominance favors the other extreme of possible effort levels as this choice is consistent with a maximin-strategy which avoids incurring excess cost if another individual chooses a lower effort level than the individual herself. The results of Van Huyck *et al.* (1990) show that groups converge to the lowest possible effort level after a few rounds. They explain this observation with strategic uncertainty, i.e., an individual does not know whether or not the other group members choose the same effort level and, therefore, the individual harks back to the risk dominant option. There are some approaches discussed in the literature to overcome coordination failure in minimum effort games. Brandts and Cooper (2006a) introduce financial incentives which increase the monetary return from successful coordination. While this approach is effective in overcoming coordination failure, its success does not depend on the intensity of financial incentives. Furthermore, even temporary increases of financial incentives are sufficient to achieve a lasting effect. In a related study, Brandts and Cooper (2006b) directly compare changes in financial incentives and different information policies to overcome coordination failure. In the first dimension, they switch from high incentives to low incentives at one point during the study and *vice versa*. Additionally, they vary whether individuals are informed about all effort levels chosen in their group or only about the minimum effort level. Extensive feedback about the effort levels of all group members turns out almost as effective as financial incentives in overcoming coordination failure. When trying to prevent coordination failure after a cut of financial incentives, however, the resolution of information provided to the individuals is only of minor relevance. To study the impact of heterogeneity on groups' ability to coordinate successfully, Brandts *et al.* (2007) introduce different costs which individuals have to bear for each unit of their effort choice. The idea is that individuals with low costs choose a higher effort level and serve as role model for the remaining individuals in the group which eventually overcomes coordination failure. Quite contrary, it turns out that those individuals whose type is most prominent within a group drive the outcome. Put differently, heterogeneity is detrimental for successful coordination and groups of homogeneous individuals are likely to have an advantage in this respect.

Besides minimum effort games discussed above, coordination can alternatively be studied in a stylized way in threshold public good games (Weber *et al.*, 2001). In contrast to linear public good games, a public good is only provided if at least some pre-defined amount of contributions to the collective account has been raised. Individual i's payoff π_i is, thus, given by

$$\pi_i = q_i - c_i + \alpha \sum_{i=1}^{n} c_i \cdot \mathbb{1}_{[\sum_{i=1}^{n} c_i \geq T]},$$

where $\mathbb{1}_{[\sum_{i=1}^{n} c_i \geq T]}$ is an indicator function which assumes a value of one if the group's total contributions $\sum_{i=1}^{n} c_i$ meet or exceed the threshold T and zero otherwise (Isaac *et al.*, 1989). The specific properties of the game depend on the refund rule and the rebate rule. The former one defines what happens to contributions if the threshold is not reached, i.e., whether contributions are refunded or not.[5] The rebate rule determines the utilization of contributions above the

threshold. Common alternatives are a no-rebate rule which means the size of the public good is fixed *ex ante* and contributions beyond the threshold are wasted, or a linear rebate rule which increases the size of the public good above the threshold along with the contributions—as in the linear public good game. As in the minimum effort game, the threshold public good game exhibits multiple equilibria. The full free riding equilibrium corresponds to the unique Nash equilibrium in the linear public good game, i.e., all individuals keep their endowment. Additionally, there is a set of threshold equilibria which include all combinations of contributions by the group which equal the value of the threshold. Unlike in the minimum effort game, these threshold equilibiria cannot be Pareto-ranked. Finally, a third set of equilibria exists if the threshold is not met in case of a full-refund policy (Cadsby and Maynes, 1999).[6]

Sabotage in Tournaments Tournaments are particularly beneficial if absolute performance is difficult to measure as they rely on a relative ranking using ordinal information. Following Lazear and Rosen (1981), two individuals compete for a prize, e.g., a bonus or promotion. The principal does not receive a precise measure about the individuals' effort levels. Instead, she receives a noisy signal and awards the prize to the individual with the higher signal. It can be shown that both individuals choose the same effort levels in the equilibrium and, thus, the winner of the tournament is determined by noise in the individuals' signals. On the other hand, Lazear (1989) shows that an individual can improve her relative position in a tournament by inflicting sabotage on her opponent. Sabotage is indeed frequently observed in experimental studies.[7] Harbring and Irlenbusch (2005) examine the impact of the principal's discretion to choose the winner prize on effort and sabotage in a tournament. The choice of the principal in this experiment influences both, the prize spread and the total sum of wages. Whereas higher incentives are expected to increase effort and sabotage from a theoretical point of view, a higher winner prize can also trigger reciprocity in agents resulting in lower levels of sabotage. Although both, effort and sabotage, increase with a higher winner prize, the increase in sabotage activities is more pronounced. Moreover, agents punish their principal by using higher levels of sabotage if she decides to award a winner prize from the lower end of the possible range. While this experimental design does not allow to disentangle the effect of a higher prize spread from an increase in the total sum of wages, Harbring and Irlenbusch (2011) provide additional insights. As before, they find that both, effort and sabotage, increase with a higher prize spread with a steeper increase for sabotage activities. Due to reciprocity, however, a higher sum of total wages yields more productive effort. Other potential determinants for behavior in tournaments are the number of competitors and the number of available winner prizes. These issues are addressed by Harbring and Irlenbusch (2008). Whereas theory predicts no influence of these factors on effort and sabotage, the results of the related experiment indicate that this is only true for the number of competitors. An unbalanced distribution in the number of winner and loser prizes leads to more intensive sabotage activities. To reduce the extent of sabotage, a balanced share of winner and loser prizes seems favorable as agents perceive a proper chance to receive one of the winner prizes. In a Tullock contest with heterogeneous agents, Harbring *et al.* (2007) introduce treatments with and without disclosing the identity of individuals who inflict sabotage on other

agents. Whether or not this information is provided is not expected to influence the decision of payoff maximizing individuals as no monetary consequences emerge. Yet, sabotage activities decrease if the identity of the saboteur is revealed to the sufferer compared to the corresponding treatment maintaining anonymity. In contrast to the studies discussed so far, Vandegrift and Yavas (2010) use a real-effort task in a tournament setting. More specifically, individuals in their setting are supposed to predict the value of a fictitious stock in a multiple-cue-probability-learning task. The individual with the more accurate estimation is awarded the winner prize. Besides changes in the prize spread, Vandegrift and Yavas (2010) distinguish between settings with a partner and stranger matching, respectively. While they find no impact of a higher prize spread on performance, sabotage activities increase. In the partner setting, sabotage is generally used more reticently; in subsequent rounds, however, sabotage is retaliated emphatically among partners. Finally, individuals who exhibit a higher ability suffer more frequently from sabotage, whereas there is no difference in the levels of sabotage individuals inflict on their opponent depending on their own ability. In a more elaborate setting, Carpenter *et al.* (2010) let individuals in their study prepare envelopes for mail shipping. Individuals in all treatments are paid a piece-rate per prepared envelope. In the tournament setting, the individual who prepared the highest number of envelopes is additionally paid a bonus. Performance is measured along two dimensions, quantity which can be assessed objectively and quality which is predominantly a subjective domain. Both ratings are either based on an assessment by an employee of the US Postal Service to ensure objectivity or other individuals of the same session. In some settings individuals have, thus, the opportunity to sabotage their opponents by rating the quality of their work low or reporting a lower number of envelopes than actually prepared by the respective individual. While the bonus increases performance compared to the piece-rate, the opportunity to inflict sabotage in the tournament discourages individuals from exerting effort and results in an average performance even below the plain piece-rate. Irrespective of the opportunity to sabotage the opponents, quality decreases in both tournament settings relative to the baseline with the piece-rate. Furthermore, financial incentives in the tournament induce individuals to sabotage their opponents severely. This relates predominantly to the quality rating due to its subjective nature and only to a lower extent to the reported number of envelopes prepared by the opponents. Furthermore, some degree of sabotage is observed in the tournament setting even in absence of any financial incentives when the assessment of the employee of the US Postal Service is relevant for determining the recipient of the bonus.

1.2 Outline of the Thesis

To address different aspects which are decisive for success in the labor market and the workplace, this thesis comprises three experimental studies on coordination, cooperation, and—as an example of adverse incentives—sabotage behavior in tournaments.

Section 2 presents the results of a laboratory experiment on coordination in a threshold public good game. The experimental design employs several collective accounts of which at most one is feasible with the given endowment. Hence, coordination is required in two domains as

individuals have to agree independently on one alternative and contribute sufficient funds to reach the corresponding threshold. In absence of communication coordination is particularly difficult in this situation. To alleviate this issue, a first-moving leader allocates her endowment to her private account and the collective accounts. The remaining individuals of the group are informed about this allocation before they decide on the allocation of their endowment simultaneously and independently. While individuals have no reason to object to the first mover's choice of a particular collective account in case of homogeneous incentives, a selfish first mover who exploits her strategic position in case of heterogeneous incentives by choosing the alternative which yields the highest payoff for herself may decrease followers' inclination to act on her example—even though followers are better off when they support a selfish leader compared to a situation in which no threshold is reached at all. The results, however, show that there are no material differences in groups' ability to coordinate under homogeneous and heterogeneous incentives, respectively. Moreover, first movers do not systematically exploit their strategic position within their groups. Quite contrary, first movers are in a substantial number of instances willing to bear a fair share in their groups' effort to reach a threshold and, thereby, act as a good example which induces followers to contribute as well.

The results of an intervention study focusing on cooperation are reported in Section 3. Children aged 10–13 years at upper secondary schools in Germany are provided with enhanced physical education lessons for six weeks. These lessons address a broad range of social skills, including cooperation. To study the effect of the intervention which conveys *inter alia* examples on challenges and advantages of cooperative behavior in a playful way on actual behavior, children participate in an incentivized linear public good game at different times during the course of the study. The first measure of cooperativeness in terms of contributions to the collective account in the public good game is elicited before the beginning of the intervention. To examine the effect of the intervention, this procedure is repeated right after the end of the intervention and nine weeks later. Besides the individuals who participate in the intervention, other children from the same schools form a control group to allow for the identification of the causal effect of the intervention on cooperative behavior. The results regarding cooperative behavior indicate only weak effects of the intervention which may be attributed to its short duration. However, a subset of individuals in the treatment group who live in households with low socio-economic status respond positively to the intervention, i.e., their cooperative behavior in the incentivized public good game is improved as a result of enhanced physical education lessons. This particular effect even occurs in the third measurement nine weeks after the end of the intervention in the Treatment group.

Finally, Section 4 provides evidence on adverse incentives in a tournament setting. Whereas tournaments exhibit advantageous properties such as inducing individuals to choose the first-best level of effort even under asymmetric information which make their application common in the workplace (Prendergast, 1999; Lazear and Shaw, 2007), sabotage is also prevalent in settings which rely on ordinal rankings. The present study considers in particular the impact of private and social image concerns on the decision to inflict sabotage on the opponent. To

disentangle these effects, the experimental design employs several treatments which differ in their information policies, i.e., either the individual herself or both, the individual and the opponent, are or may be informed about the implication of the individual's sabotage activities. In a second stage, individuals have the opportunity to prevent information disclosure by paying a price. Whereas the decision to inflict sabotage on one's opponent does not vary between different information policies, individuals' willingness to pay does depend on whether or not information may be disclosed to the opponent. More specifically, individuals exhibit a significantly higher willingness to pay to prevent disclosure about the implication of their sabotage activities to their opponent which may be attributed to maintaining a positive social image.

Notes

1. The prisoner's dilemma (Luce and Raiffa, 1957) can be considered as special case of the linear public good game with $n = 2$ individuals (Roth, 1995).
2. Thus, $q_i - c_i$ is the amount kept by the individual for her private account.
3. Since $\frac{\partial \pi_i}{\partial c_i} = -1 + \alpha < 0$. Note that in case of $\alpha < \frac{1}{n}$ even the social return from the public good is lower than the individual costs of contributing. On the other hand, $\alpha > 1$ yields a private return from the public good which exceeds the associated costs of contributing.
4. For a different use of the term conditional cooperation, see Keser and van Winden (2000).
5. There is also the possibility that contributions are only partly refunded.
6. Note that both, the full free riding equilibrium and the equilibria which do not reach the threshold in case of the full-refund rule, are Pareto-dominated by the threshold equilibria.
7. Providing reliable figures about the extent of unethical behavior is challenging. Alem *et al.* (2018) provide evidence that individuals' revealed behavior differs from their stated behavior which suggests that non-experimental approaches such as surveys are not fully appropriate to study determinants of unethical behavior.

2 First Movers in a Threshold Public Good Game*

We study coordination among individuals in a threshold public good game with several collective accounts. As resources are scarce, only one public good is feasible. Coordination, therefore, is required in two domains: *(i)* Individuals have to choose one out of several public goods and *(ii)* they have to coordinate their contributions to reach the corresponding threshold. To facilitate coordination, we employ a semi-sequential structure with a first-moving leader and second-moving followers who decide simultaneously and independently after being informed about the first mover's allocation. The main treatment variable distinguishes groups with homogeneous and heterogeneous incentives for particular public goods. The institution of a semi-sequential move order fosters coordination even in the case of heterogeneous incentives. Overall, we do not observe substantial differences between both treatments.

2.1 Introduction

The notion that individuals are sometimes not able to accomplish extensive projects on their own is prevalent. Moreover, projects often exhibit step-level character, i.e., success is only achieved if a certain level of input has been contributed. If there is more than one alternative it is essential that individuals not only contribute to a project, but also coordinate their contributions. In cases with simultaneous contributions, coordination is difficult in absence of communication. When alternatives are virtually indistinguishable and individuals have identical incentives, the allocation of a first mover who contributes prior to other individuals is expected to serve as a blueprint for subsequent contributions by second-moving followers and, therefore, can foster coordination. If individuals have, however, divergent incentives and the first mover behaves in a selfish manner by choosing the alternative which yields a higher return just for herself, it is less clear whether second-moving followers still act on the first mover's suggestion.

We use a threshold public good game with several collective accounts. The parameters are set in a way that only one threshold is feasible. Moreover, coordination is aggravated by the fact that no individual can reach a threshold on her own. The institution of a first mover who allocates her endowment before all other individuals do is expected to foster coordination by making one collective account salient, i.e., once the first mover has contributed a decent amount to one collective account, followers are concerned with only a subset of the initial coordination problem, namely how to raise the remaining input which is necessary to reach the corresponding threshold. In case of homogeneous incentives with symmetric monetary returns from all collective accounts for all individuals, followers have no reason to object to the first mover's choice. Heterogeneous incentives, on the other hand, may induce divergent attitudes towards the choice of a specific collective account by the first mover. In particular, followers' reactions to a selfish first mover who chooses the collective account which yields a higher return for herself can be ambiguous: A first mover who contributes a decent amount of input herself can set a good a example for

*This section is based on Haas (2019). Submitted to *Experimental Economics* on November 26, 2018 (manuscript no. EXEX-D-18-00300); rejected on January 21, 2019.

followers who, in turn, are inclined to act similarly. The mere fact that followers are better off if they act on the suggestion of a selfish first mover and receive a monetary return from the collective account compared to a situation in which the threshold is not reached, however, does not guarantee that they behave accordingly. The reason for this conjecture is the way followers may perceive actions of the first mover. Unlike in the case of homogeneous incentives, followers cannot rely on the notion that contributions by the first mover are an act of pure generosity. If they suspect the first mover only to be focused on maximizing her own monetary payoff and trying to lure followers into contributing substantial inputs, the institution of a first mover can have detrimental effects on overall contributions (Glöckner *et al.*, 2011).

Hence, the aim of this paper is to study whether groups of several individuals with a first-moving leader and several second-moving followers are able to coordinate contributions in a threshold public good game with several collective accounts if individuals have divergent incentives for particular collective accounts. To address this research question, we conduct a laboratory experiment with two different treatments in which individuals interact repeatedly in fixed groups. One individual in each group is exogenously assigned the role of the first-moving leader. Individuals can allocate their endowment between their private account and several collective accounts which only generate a monetary return if a pre-defined amount of input is raised. As mentioned above, only one threshold is feasible and all individuals are required to contribute in order to reach the respective threshold. The latter constraint refrains individuals from "cheap riding" which relates to individuals who try to reduce their own contribution in the group's effort to reach a threshold at the expense of other individuals (Isaac *et al.*, 1989). This phenomenon must not be confused with free riding which is frequently observed in *linear* public good games (Ledyard, 1995; Chaudhuri, 2011). *Threshold* public good games, in contrast, are concerned with coordination, i.e., each individual's contribution is essential for a successful outcome of the entire group which imposes a strong incentive to contribute on each individual. In particular, coordination games require individuals to behave similarly to achieve a successful outcome, whereas it is individually optimal to free ride irrespective of the other individuals' actions in linear public good games (Weber *et al.*, 2001).

The results of this study show no substantial differences between both treatment groups. Individuals in groups with heterogeneous incentives are even slightly more successful in coordinating on a threshold than individuals in the treatment with homogeneous incentives although the difference is not statistically significant. On average, first movers in both treatments contribute almost the same amount as followers. In more than 45% of all instances, first movers in both treatments contribute their entire endowment, i.e., first movers are indeed willing to set a good example for second-moving followers rather than exploiting their strategic position within the group. The so-called first mover's curse which can occur in experiments with sequential contributions due to followers whose contributions fall short of the first mover's contribution (e.g., Cappelen *et al.*, 2016; Gächter and Renner, 2018; see also Murnighan *et al.*, 1993) is not observed in this setting. Implementing a semi-sequential move order with a first-moving leader and second-moving followers is, hence, beneficial for coordination in a threshold public good game

with several collective accounts—even if individuals have divergent incentives for particular collective accounts.

2.2 Related Literature

Absence of leadership can be blamed for coordination failure among individuals (Brandts and Cooper, 2006a; Brandts *et al.*, 2007). When an individual takes the responsibility and emerges as leader, there are different approaches how to convince followers. Two notable examples frequently discussed in the literature are leading by words and leading by example. Whereas the first alternative relies on a non-binding announcement of intentions, leading by example necessitates a binding commitment of the leader which is usually costly and irrevocable. Leading by words, on the other hand, can be classified primarily as cheap talk and has only limited influence on followers' behavior (e.g., Pogrebna *et al.*, 2011; Dannenberg, 2015).

In a team production setting with asymmetric information a leader can induce followers to exert high levels of effort by using costly signals. Hermalin (1998) describes a model with hidden information in which a leader with private information exerts effort before all others individuals do. Followers interpret the effort level of the leader as signal about her private information and choose their effort levels subsequently. The payoff generated by the group depends on a state of the world which is only known by the leader and the total amount of effort exerted by the group. Followers, therefore, do best by adapting to their leader's effort level. Empirical evidence by Meidinger and Villeval (2002) indeed shows that followers imitate the leader's effort choice. Differences in situations with symmetric and asymmetric information occur due to different behavior of the leader. Whereas symmetric information induce leaders to strive for high levels of coordination which is reciprocated by followers, leaders who face an environment with asymmetric information behave contrarily and prefer to free ride which is detrimental for coordination within groups.

While a large number of contributions emphasize the positive effects of sequential contributions on cooperation a linear public good games (e.g., Moxnes and van der Heijden, 2003; Vesterlund, 2003; Potters *et al.*, 2005, 2007; Güth *et al.*, 2007; Levati *et al.*, 2007; Gächter *et al.*, 2010, 2012), the effect of leadership in coordination settings has received little attention so far. Based on these results, Sahin *et al.* (2015) study the effectiveness of leading by example compared to leading by words on both, cooperation in a linear public good game and coordination in a minimum effort game. In the latter case, both leadership styles result in substantially higher effort choices compared to the benchmark setting with simultaneous decisions after the first period. Leaders' effort choices in the case of leading by example are slightly higher than in the case of leading by words, albeit the difference is not statistically significant. The relation between both treatments is reversed for effort choices of followers. Overall, effort choices of first-moving leaders and followers are not different. Under either leadership style, followers react reciprocally to the example or suggestion by the first-moving leaders and almost match their increase in effort choices.

Cartwright *et al.* (2013) study the effect of first movers on preventing coordination failure in minimum effort games. In particular, they compare the minimum effort choice in groups of three and four individuals who decide simultaneously to groups of four individuals in which the role of the first-moving leader is either exogenously or endogenously assigned. The presence of a first-moving leader who sets an example for second-moving followers reduces strategic uncertainty and is, hence, expected to facilitate coordination on higher effort levels (Cartwright *et al.*, 2013). Over the course of the experiment, minimum effort choices in both, the treatments with a first-moving leader and the three-player treatment with simultaneous effort choices, are stable, while the trend in the four-player treatment with simultaneous effort choices is decreasing. This limited effect of leading by example on the prevention of coordination failure is to due to the fact that leaders do not persistently choose higher effort levels than individuals in the treatments with simultaneous decisions. Only for later periods differences in leaders' effort choices in the treatments with a first-moving leader and the three-player treatment with simultaneous effort choices on the one hand and the four-player treatment with simultaneous effort choices on the other hand are observed. To summarize, even though Cartwright *et al.* (2013) observe a positive correlation between first-moving leaders' and followers' effort choices, coordination failure can be attributed to leaders' low effort choices to a large extent.

More recently, Dong *et al.* (2018) compared the effect of different leadership styles—leading by example and leading by words—on restoring coordination after coordination failure and on preventing coordination failure in a minimum effort game, respectively. The experiment consists of two blocks with ten periods each. Between the treatments, the timing of the introduction of the mechanisms varies: In the restoration treatment, individuals play a minimum effort game with simultaneous decisions in the first ten periods and one of the leadership mechanisms is introduced for the subsequent block—and *vice versa* for the prevention treatments. The benchmark scenario consists of two blocks with simultaneous decisions. Regarding the timing of introduction, the effect of leadership on coordination in a minimum effort game is more pronounced when introduced at the beginning to prevent coordination failure as compared to restoring coordination after a series of unsuccessful attempts. Effort choices in the restoration setting with a mechanism of leadership being introduced at the beginning of the second block quickly converge to effort choices made in the benchmark case with simultaneous decisions. In contrast to this observation, effort levels in the prevention setting with leadership mechanisms in the first block are higher than in the benchmark case with simultaneous decisions and in the restoration case after the first two rounds. A comparison of the leadership mechanisms reveals that followers are more inclined to adapt their decision to a higher effort choice if the first mover's decision is binding compared to a non-binding suggestion.

Kaplan *et al.* (2018) study a situation in which coordination is crucial for efficient cooperation. In a two-player game, each individual is assigned an integer between 1 and 5 and has to choose between the two alternatives to enter or leave the game, respectively. An individual who leaves the game receives a payoff of zero. If one individuals enters the game, whereas the other individual does not, the former one receives the value of her number. Finally, if both individuals enter

the game, each individual receives on third of the value of her number. In the related experiment Kaplan *et al.* (2018) use three different treatments in which individuals can either decide simultaneously, sequentially, or wait until the other individual has made her decision. One cooperative solution of this setting is that one individual enters the game and the other individual does not. Although the treatments employing two stages lead to an increase in uncooperative behavior, payoffs in the sequential treatment with an exogenous move order are substantially higher than in case of simultaneous decisions. It should be noted, however, that this effect is driven by a subset of cooperative groups who use the information due to the sequential structure and coordinate on outcomes which are efficient for the group.

Another setting in which coordination matters are threshold public good games. Unlike in linear public good games, the public good is only provided if contributions reach at least a pre-defined level (e.g., Isaac *et al.*, 1989; Croson and Marks, 2000, see also Section 2.4.1). Early evidence on the effectiveness of a sequential move order in the provision of threshold public goods is found by Erev and Rapoport (1990) who let individuals in their experiment make a binary decision whether or not to contribute. In the sequential case, individuals decide one after another with perfect information about the decisions of preceding individuals. Compared to the case of simultaneous decisions, Erev and Rapoport (1990) observe a higher frequency of public good provision if individuals decide sequentially. Similarly, Coats *et al.* (2009) study the effects of sequential and simultaneous contributions under a full refund and a no refund scheme, respectively. As before, a sequential structure with information about preceding individuals' contributions turns out more successful regarding the provision of threshold public goods compared to simultaneous contributions.

Whereas in the studies discussed above situations with only one threshold level are considered, Normann and Rau (2015) add a second threshold level in their setting. More precisely, they study simultaneous and sequential contributions in groups of two individuals in environments with one or two threshold levels, respectively. One potential weakness of sequential moves in coordination problems is the fact that individuals who decide early can exploit their position and try to make successive individuals bear the major part of the burden. Individuals who decide later, however, have the opportunity to punish a first mover by contributing zero. The introduction of a second threshold level by Normann and Rau (2015) is likely to induce first movers to make higher contributions such that the second threshold level is feasible for the successive individual. The follower, in turn, has no incentive to punish the first mover. In line with other studies on sequential contributions in threshold public good games, Normann and Rau (2015) find this approach to be more successful regarding the provision of threshold public goods than the simultaneous alternative. As expected, having a second threshold level increases contributions. Yet, the frequency of public good provision is not affected by the additional threshold level.

The approach which is most closely related to the present study is introduced by Corazzini *et al.* (2015) who do not use several threshold levels but several threshold public goods. Groups of

four individuals can simultaneously allocate their endowment to either their private account or one or more of four collective accounts. The collective accounts only yield a return to each individual if contributions meet or exceed the corresponding threshold. As individuals decide simultaneously, coordination is difficult. Therefore, different treatments try to make one collective account salient either by assigning a marginal per-capita return[1] which is different from the other collective accounts or by using a computer-generated message which appears on individuals' screens. Without saliency as coordination device, the frequency of public good provision is lower than in the benchmark case with a single threshold public good. If the salient alternative is identical to the threshold public good which yields the highest potential payoff, the frequency of provision is the same as in the benchmark case. If saliency is at odds with efficiency, however, individuals tend to ignore the coordination device and try to reach the threshold of any of the collective accounts which yield a higher potential payoff—which eventually results in a lower frequency of public good provision. Providing a randomly generated signal as coordination device results in a higher frequency of successful coordination than in absence of a coordination device or with a coordination device which is at odds with efficiency. At the same time, the randomly generated signal is less successful regarding the provision of threshold public goods compared to situations with a single threshold public good or when the salient alternative is also the most efficient one.

2.3 Conceptual Framework

2.3.1 Set-Up

Consider a group of $n > 2$ individuals indexed $i = 1, \ldots, n$ with $i = 1$ as first-moving leader and $i \neq 1$ as second-moving followers. Individual i's endowment is denoted by q_i. Each individual can either keep her endowment or allocate some fraction to one or more of n collective accounts indexed k. The contribution of individual i to collective account k is denoted by $c_{i,k}$. The individual's total contribution to the collective accounts is denoted by $C_i = \sum_k c_{i,k} \leq q_i$. For each collective account k, there is a threshold T_k. For simplicity, the thresholds are identical for all collective accounts, i.e., $T_k = T \; \forall k$. Each threshold T satisfies the following properties: *(i)* the threshold is feasible: $T \leq \sum_i q_i$; *(ii)* the group can reach at most one threshold: $\sum_i q_i < 2T$; *(iii)* all individuals have to contribute to reach the threshold: $T > \frac{n-1}{n} \sum_i q_i$. Hence, the thresholds are defined by $\frac{n-1}{n} \sum_i q_i < T \leq \sum_i q_i$. The third property imposes strong incentives to contribute in order to reach a threshold as it confines individuals' incentives to cheap ride (Isaac *et al.*, 1989).

Whenever the group's contribution to a collective account meets or exceeds the corresponding threshold, all individuals of the group receive a return $\alpha_{i,k} C_k$, where $\alpha_{i,k}$ represents the marginal per-capita return of individual i from collective account k with $\frac{1}{n} < \alpha_{i,k} < 1$ and $C_k = \sum_i c_{i,k}$ denotes the group's total contribution to the corresponding collective account k. Thus, individual i's payoff is given by

$$\pi_i = q_i - C_i + \sum_k \alpha_{i,k} \cdot C_k \cdot \mathbb{1}_{[C_k \geq T]},$$

where $q_i - C_i$ corresponds to the amount she transfers to her private account. $\mathbb{1}_{[C_k \geq T]}$ is an indicator function which equals one if the group's total contribution to collective account k meets or exceeds the corresponding threshold T and zero otherwise.

The game employs a semi-sequential move order. In the first stage, the first mover allocates her endowment to her private account and the collective accounts. After they have been informed about the first mover's allocation, followers decide simultaneously and independently about the allocation of their endowment. If the group does not reach a threshold, contributions to the corresponding collective account are not refunded. Contributions above the threshold are treated like in a linear public good game, i.e., the size of the public good increases in contributions. This reduces the setting to a linear public good game if the number of collective accounts is set to one and the threshold is zero (Corazzini *et al.*, 2015).

2.3.2 Analysis

The following analysis is limited to a one-shot interaction and pure strategies (Isaac *et al.*, 1989). The properties of the thresholds require coordination in two domains: *(i)* As resources are scarce, at most one threshold is feasible. *(ii)* Furthermore, contributions by all individuals are required for successful coordination as a threshold is not feasible if one or more individuals decide to contribute only a negligibly small share of their endowment or nothing at all. In a simultaneous setting predictions about individual contributions are not very constructive (Isaac *et al.*, 1989). The semi-sequential move order discussed here, however, exhibits advantages due to the observability of the first mover's allocation.

Regarding the choice of a collective account as one domain of the coordination problem, the contribution of the first mover is decisive for the group. Since at most one threshold is feasible, the first mover contributes—if at all—to one collective account $k' \in k$. Contributions to all other collective accounts $k \neq k'$ are, therefore, zero for all individuals (Corazzini *et al.*, 2015). If the first mover transfers the entire endowment to her private account or contributes only a negligibly small share of her endowment below a critical value c_1^{crit} to collective account k' such that the followers' endowments are not sufficient to cover the remaining amount to reach the threshold, i.e., $c_{1,k'} < c_1^{\text{crit}} \equiv T - \sum_{i \neq 1} q_i$, followers do best by not contributing at all. The first mover anticipates followers' behavior in this case and does not contribute to the collective account either. This results in individual contributions $c_{i,k} = 0$ for all collective accounts k (including k') and for all individuals i which constitutes an inefficient outcome.

If the first mover contributes $c_{1,k'} \geq c_1^{\text{crit}}$ to collective account k', the corresponding threshold is feasible for the group and the choice of a collective account is determined by the first mover's contribution. Hence, followers also contribute—if at all—only to collective account k'. The second part of the coordination problem concerning individual contributions is mainly directed toward the followers who play a simultaneous threshold public good game in the second stage. In this subgame the threshold is defined by $\widetilde{T} \equiv T - c_{1,k'}$. Possible equilibria of the subgame are the inefficient equilibrium with $\sum_{i \neq 1} c_{i,k'} = 0$ and numerous threshold equilibria with $\sum_{i \neq 1} c_{i,k'} = \widetilde{T}$.

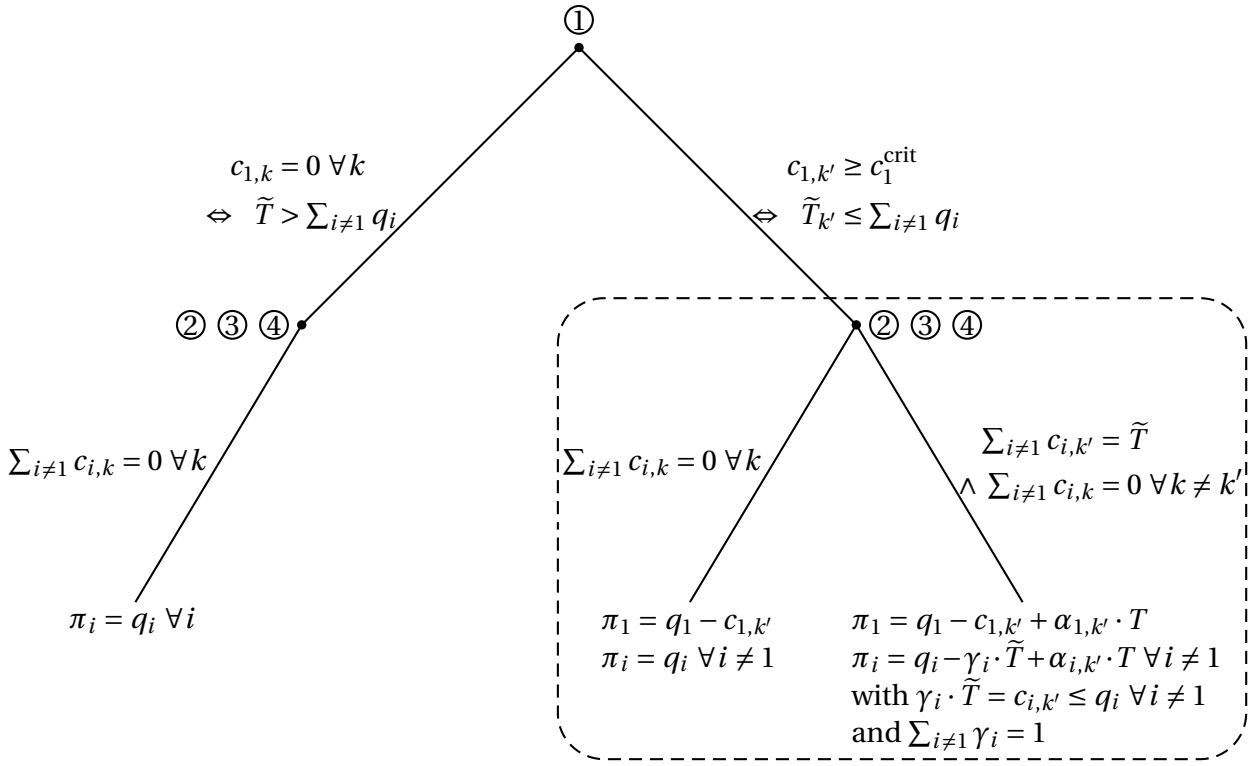


Figure 2.1: Illustration of the game.

In particular, each follower's contribution in the latter case reflects a fraction of the threshold $\widetilde{T}$ in the subgame, i.e., $c_{i,k'} = \gamma_i \widetilde{T} \leq q_i\ \forall i \neq 1$ with $\sum_{i\neq 1} \gamma_i = 1$.[2] In the full game, therefore, both contribution vectors $\mathbf{c} = \left(c_{1,k'}, \sum_{i\neq 1} c_{i,k'} = 0\right)$ and $\mathbf{c} = \left(c_{1,k'}, \sum_{i\neq 1} c_{i,k'} = \widetilde{T}\right)$ for collective account k' are equally plausible with individual contributions $c_{i,k} = 0\ \forall k \neq k'$ (Andreoni, 1998). Note that outcomes in which no threshold is reached are Pareto-dominated by outcomes which reach a threshold. In the latter case, however, outcomes cannot be Pareto-ranked among each other (Isaac *et al.*, 1989).

The game features multiple equally plausible outcomes which exacerbates coordination. Yet, some outcomes are salient. Notable instances are cases with symmetric contributions such as the inefficient outcome with individual contributions $c_{i,k} = 0\ \forall k$ and the outcome which reaches a threshold with individual contributions $c_{i,k'} = 40$ and $c_{i,k} = 0\ \forall k \neq k'$. Note that in the homogeneous treatment both types of symmetric contributions also result in symmetric payoffs, whereas payoffs in the heterogeneous case are asymmetric for the presented threshold outcome. In general, equal payoffs cannot be achieved in the heterogeneous treatment if the group reaches a threshold.

The concept of equilibrium selection introduced by Harsanyi and Selten (1988) provides additional classifications. The numerous outcomes in which a threshold is reached are Pareto-superior to other outcomes in which the threshold is not reached and are, thus, classified payoff-dominant. The inverse argument of risk dominance applies to inefficient outcomes in which all individuals transfer their entire endowment to their private account and are not at risk of losing their contributions to a collective account if the corresponding threshold is not reached

(Corazzini *et al.*, 2015). Moreover, the approach by Harsanyi and Selten (1988) provides useful insights to the simultaneous threshold public goods game played by the followers in the second stage after the first mover has contributed c_1^{crit}. In this situation, followers face the decision whether or not to contribute. Due to a high degree of strategic uncertainty about the actions of other individuals who decide simultaneously, the risk-dominant strategy in this subgame is to contribute zero tokens to the collective accounts as opposed to the payoff-dominant alternative which requires positive contributions to collective account k' by followers to reach the threshold $\widetilde{T}$ in the subgame. In fact, a contribution by the first mover at or above the critical value c_1^{crit} is a necessary, but not a sufficient requirement to reach the corresponding threshold in the full game.

Having two equally plausible sets of outcomes, the actual outcome remains predominantly an empirical question. Depending on the first mover's contribution, however, the relation between risk imposed on an individual follower and potential payoffs changes. If the first mover contributes a small amount of tokens equal to or marginally larger than the minimum required amount to reach a threshold, c_1^{crit}, followers have to cover a large amount to reach the threshold which imposes not only risk on an individual follower but also reduces potential payoffs. On the other hand, if the first mover contributes a large amount of tokens close or equal to q_1, the remaining gap to reach the corresponding threshold is smaller than in the previous case which results in higher potential payoffs for the followers at lower risk and, hence, alleviates the assurance problem (see Isaac *et al.*, 1989, for a related argument). To summarize, followers' inclination to contribute to the collective account in the subgame played in the second stage can be influenced by the first mover's contribution. In particular, higher contributions by the first mover can increase followers' willingness to act on her example (see Cartwright *et al.*, 2013; Dong *et al.*, 2018, for evidence on coordination on higher effort levels in minimum effort games).

2.4 Experimental Design

2.4.1 Procedures

A total of 76 individuals participated in six sessions between April and July 2016 at Karlsruhe Decision & Design Lab. Individuals were recruited via ORSEE (Greiner, 2015) for the first and second session and via hroot (Bock *et al.*, 2014) for subsequent sessions. Experimental procedures were computerized using z-Tree (Fischbacher, 2007).

Upon arrival, individuals were randomly allocated to cubicles equipped with computer terminals. After they had read the instructions and answered a set of control questions, individuals played three trial periods with no interaction in which they had the opportunity to familiarize with the payoff mechanism of the first part of the experiment. At the end of the trial periods, groups of four individuals were randomly matched and one individual was randomly assigned the role of the first mover. Individuals were not informed about the other group members' identities. Eventually, individuals played a threshold public good game for ten periods using a partner matching protocol. In the second part of the experiment, individuals were matched in new

groups of two with another individual they had not interacted with before and played a Battle of the Sexes game (e.g., Camerer, 2003) using the strategy method (Selten, 1967). Each individual indicated an independent decision as row player and two dependent decisions as column player. The purpose was gathering additional information about individuals' basic understanding of the structure of a situation with multiple, yet conflicting equilibria.[3] In the third part of the experiment, we elicited risk preferences using the staircase approach by Falk *et al.* (2016). The final questionnaire included open questions about individuals' decisions in the previous parts of the experiment, questions from the German socio-economic panel about impulsiveness, risk incentives, and attitudes about trust in other people (DIW/SOEP, 2014; Dohmen *et al.*, 2011), a subset of the competitiveness index by Smither and Houston (1992), and questions about demographic characteristics.

During the experiment the fictitious currency tokens was used. At the end of the experiment, one period of the first part was randomly selected for payoff. For the second part of the experiment, it was randomly determined which player of a group was the row player and the column player, respectively. Payoffs were determined according to the row player's independent decision and the corresponding dependent decision of the column player. Additionally, individuals received a payoff from the incentivized elicitation of risk incentives in the third part of the experiment. The total amount of tokens earned in the experiment was converted into Euro at an exchange rate of 10 tokens = 1.00 Euro and paid out privately to each individual in cash. The experiment lasted about 75 minutes and individuals earned on average 15.82 Euro including a show-up fee of 4.00 Euro.

2.4.2 The Game

The design of the first part of the experiment is based on a threshold public good game (e.g., Isaac *et al.*, 1989; Croson and Marks, 2000) with several collective accounts and a semi-sequential move order. Groups consist of four individuals endowed with 50 tokens each. One individual is randomly assigned the role of the first mover for the entire experiment. Individuals can allocate their endowment in increments of 5 tokens to their private account and to one or more of four collective accounts. The collective accounts are labeled by colors. The order of the collective accounts is randomly determined in each period.

In the first stage of each period, the first mover allocates her endowment. Meanwhile, followers are asked to state their beliefs as to which account the first mover allocates the largest share of her endowment. A correct belief is rewarded with an additional payoff of 10 tokens. In the second stage, followers are informed about the allocation of the first mover and then allocate their endowment simultaneously and independently. As before, the first mover is asked to state her belief as to which account the followers allocate on average the largest share of their endowment which is also incentivized.

Whereas the whole group is entitled to payoffs generated by the collective accounts, only the individuals themselves benefit from the payoff of the respective private accounts. After all

		Collective account			
		Red	Blue	Yellow	Green
Individual	1	0.75	0.5	0.5	0.5
	2	0.5	0.75	0.5	0.5
	3	0.5	0.5	0.75	0.5
	4	0.5	0.5	0.5	0.75

Table 2.1: Marginal per-capita returns in the heterogeneous treatment. In the homogeneous treatment, marginal per-capita returns are equal to 0.5 for all individuals and all collective accounts.

individuals have allocated their endowment to the collective accounts and their private accounts, they receive a return from the collective accounts—provided the corresponding threshold of 160 tokens for a particular collective account has been reached. The return is determined by the marginal per-capita return as fraction of the group's total contribution to this particular collective account. In the homogeneous treatment, the marginal per-capita return assumes a value of 0.5 for all individuals and all collective accounts. In the heterogeneous case, however, there is always one collective account which is more beneficial for one particular individual with a marginal per-capita return of 0.75. At the same time, an individual's preferred option is different from the others' most favorable collective accounts (see Table 2.1).

At the end of each period, individuals receive detailed feedback for all collective accounts. They are informed about the other individuals' contributions to each collective account and they are told whether a threshold has been reached. Finally, they learn the return from their private account and the collective account as well as their potential earnings from the current period.

2.4.3 Conjectures

If a threshold is reached, the payoffs of all individuals of a group are higher than in a situation in which individuals fail to reach a threshold. A group of payoff-maximizing individuals, therefore, strives to coordinate on a threshold. The construction of the thresholds allows the group to reach at most one threshold. Moreover, each individual is required to contribute. In particular, the semi-sequential move order requires the first mover to contribute at least c_1^{crit} tokens such that the followers can cover the remaining amount with their endowment.

(C1) The first mover contributes a positive amount equal to or larger than c_1^{crit} to one collective account k'.

The game also requires coordination in another domain which concerns the choice of one out of four collective accounts. As discussed in Section 2.3.2, the properties of the thresholds allow the first mover to determine which collective account k' is feasible for the group. In the homogeneous treatment, the collective accounts are indistinguishable regarding their potential payoffs. Thus, the first mover can choose one collective account randomly. In the heterogeneous case, however, there is one particular collective account for each individual which yields a higher

return than the other collective accounts. As the first mover is expected to maximize her own monetary payoff, she chooses the collective account which yields the highest payoff for herself.

(C2) When marginal per-capita returns are heterogeneous, the first mover contributes to her preferred collective account.

Although the second movers are informed about the first mover's allocation, an individual second mover does not know how the others will behave as all second movers decide simultaneously and independently. Nevertheless, the game provides incentives for the followers to contribute. If the contribution by the first mover increases, the remaining amount to be covered by the followers decreases which is supposed to enhance followers' inclination to contribute.

(C3a) Followers are more inclined to contribute to the collective account if the first mover's contribution increases.

Even though followers have divergent incentives toward the choice of the collective account than their first mover in the heterogeneous case, they do not oppose the first mover's choice because they are still better off if the corresponding threshold is reached compared to a situation in which coordination is not successful.

(C3b) Followers act on the first mover's suggestion even in the heterogeneous treatment.

2.5 Results

2.5.1 Overview

The data set contains observations of 76 individuals in groups of four which remained unchanged over the entire course of the experiment. Thus, there are 19 independent observations on group level, nine in the homogeneous treatment and ten in the heterogeneous treatment. Non-parametric tests are based on averages by groups over all periods. To control for the impact of additional factors on individual contributions, results of regressions on random effects models are also reported. Robust standard errors in these regressions are clustered on group level.

Figure 2.2 displays average contributions by first movers and followers in both treatments by period. Overall, individuals contribute 30.38 tokens in the homogeneous treatment and 33.35 tokens in the heterogeneous treatment to the collective accounts. The difference between both treatments is statistically not significant ($p = 0.6831$, two-sided Mann-Whitney U-test). Relative contributions are in the range which is frequently observed in threshold public good experiments with one collective account (see Croson and Marks, 2000, for a meta analysis). Moreover, Corazzini *et al.* (2015) report a slightly lower level of average contributions (about 55% of the initial endowment) in their treatment 4G_RS which is similar to the present setting as it uses four collective accounts and a computer-generated signal to make one collective account salient. Within treatments, there is only a minor difference between first movers (30.72 tokens) and followers (30.26 tokens) in the homogeneous treatment as well as between first movers (33.15 tokens) and followers (33.42 tokens) in the heterogeneous treatment. None of

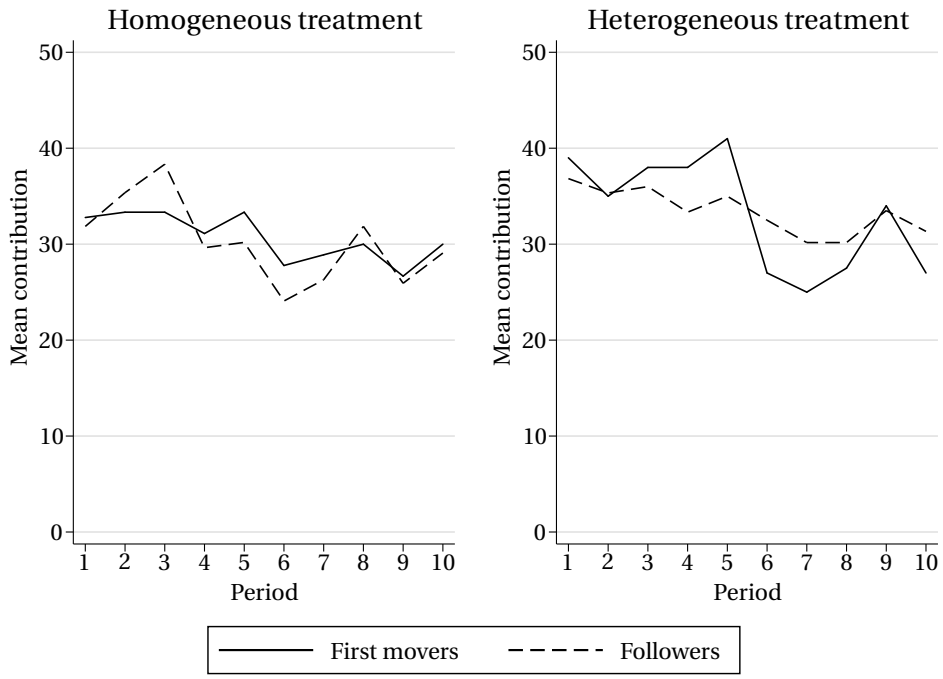


Figure 2.2: Average contributions in the homogeneous and the heterogeneous treatment by first movers and followers across periods.

these differences by type and treatment is statistically significant, neither within treatments (homogeneous treatment: $p = 0.6350$, heterogeneous treatment: $p = 0.3583$, two-sided Wilcoxon signed-rank test) nor between treatments (first movers: $p = 0.9345$, followers: $p = 0.4622$, two-sided Mann-Whitney U-test).

The fraction of successful groups which are able to reach a threshold by period and treatment is depicted in Figure 2.3. In the first period, three out of nine groups in the homogeneous treatment and four out of ten groups in the heterogeneous treatment are able to coordinate their contributions to reach a threshold. By the end of the experiment, the number of successful groups increases to five groups in the homogeneous treatment and seven groups in the heterogeneous treatment.[4] Over all periods, coordination on a threshold is successful in 50% of all cases in the homogeneous treatment and in 58% of all cases in the heterogeneous treatment. Compared to 63.2% in the treatment 4G_RS of Corazzini *et al.* (2015) success rates are slightly lower in both treatments of the present study which may be attributed to different threshold levels in both experiments.[5] The difference in the average fraction of successful periods between treatments is statistically not significant ($p = 0.7267$, two-sided Z-test (two-sample)).

Figure 2.4 provides an overview of average payoffs by type and period in both treatments. On average, individuals earn 65.38 tokens per period in the homogeneous treatment and 75.32 tokens in the heterogeneous treatment. The difference is statistically not significant ($p = 0.6242$, two-sided Mann-Whitney U-test). However, it can be explained by the fact that there is always one collective account for each individual in the heterogeneous treatment which yields a marginal per-capita return which is 50% higher than for the remaining collective accounts and

Figure 2.3: Frequency of thresholds reached in the homogeneous treatment and the heterogeneous treatment by period.

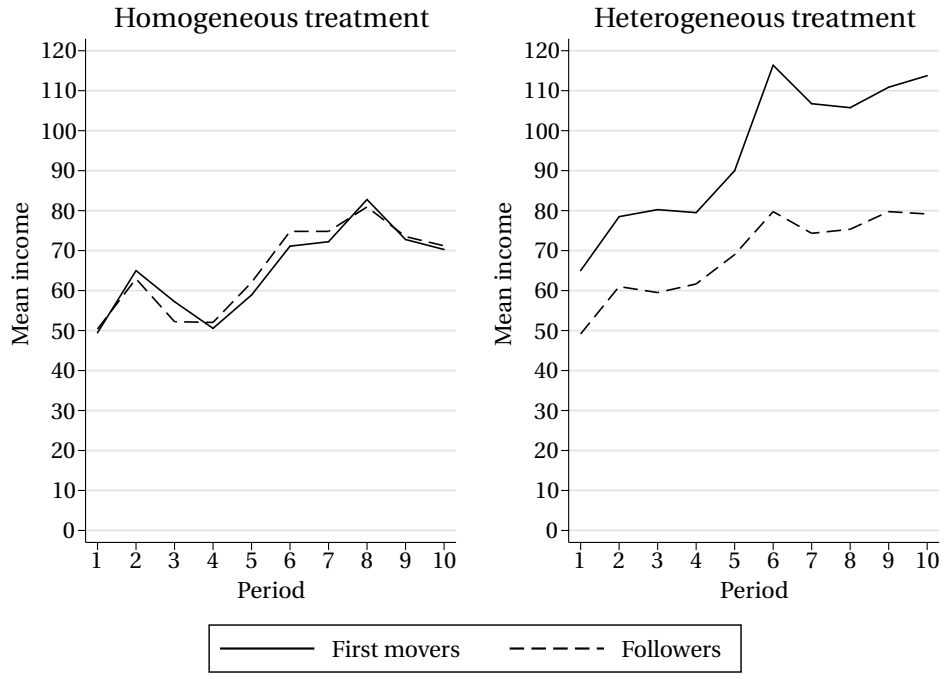


Figure 2.4: Average income in the homogeneous treatment and the heterogeneous treatment by type and period.

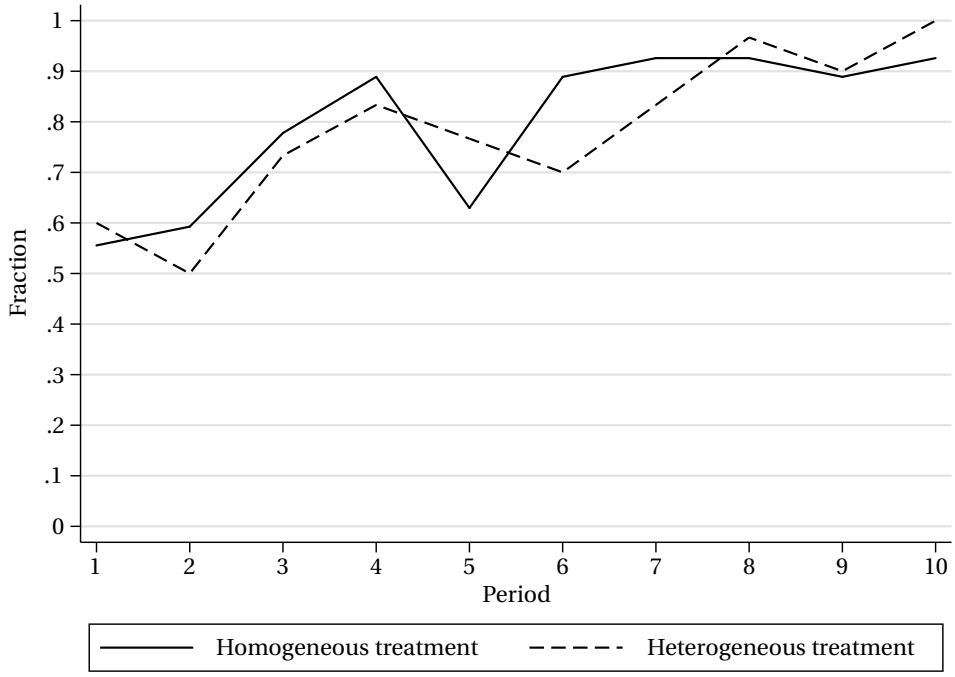


Figure 2.5: Average fraction of followers' correct beliefs by treatment and period.

for the other individuals. In the homogeneous treatment, the difference in average payoffs of first movers (65.03 tokens) and followers (65.49 tokens) is statistically not significant ($p = 0.6350$, two-sided Wilcoxon signed-rank test). The difference between first movers' (94.68 tokens) and followers' average payoffs (68.87 tokens) in the heterogeneous treatment, in contrast, is substantial and statistically significant ($p = 0.0284$, two-sided Wilcoxon signed-rank test). This is an indication of first movers choosing the collective account which yields the highest potential payoff for themselves in the heterogeneous treatment. Indeed, first movers in the heterogeneous treatment allocate the highest contribution to a particular collective account to the alternative which is most beneficial for themselves regarding the potential payoff in 74% of all cases—which makes this collective account also salient for followers.

The fraction of followers' correct beliefs in the homogeneous and the heterogeneous treatment in each period is depicted in Figure 2.5. The differences between both treatments are statistically not significant ($p = 0.9017$, two-sided Mann-Whitney U-test). The graph indicates that already in the first period more than half of the individuals in both treatments correctly anticipate the account which the first mover allocates the largest share of her endowment to. Whereas followers in the homogeneous treatment frequently expect the first mover to choose the first collective account on the screen (from left to right; see Figure A.3), followers in the heterogeneous treatment in many cases anticipate that the first mover chooses the collective account which yields the highest payoff for herself. Over the course of the experiment, the fraction of correct beliefs increases such that almost all followers state a correct belief by the final periods which indicates that the repeated interaction in fixed groups fosters learning.

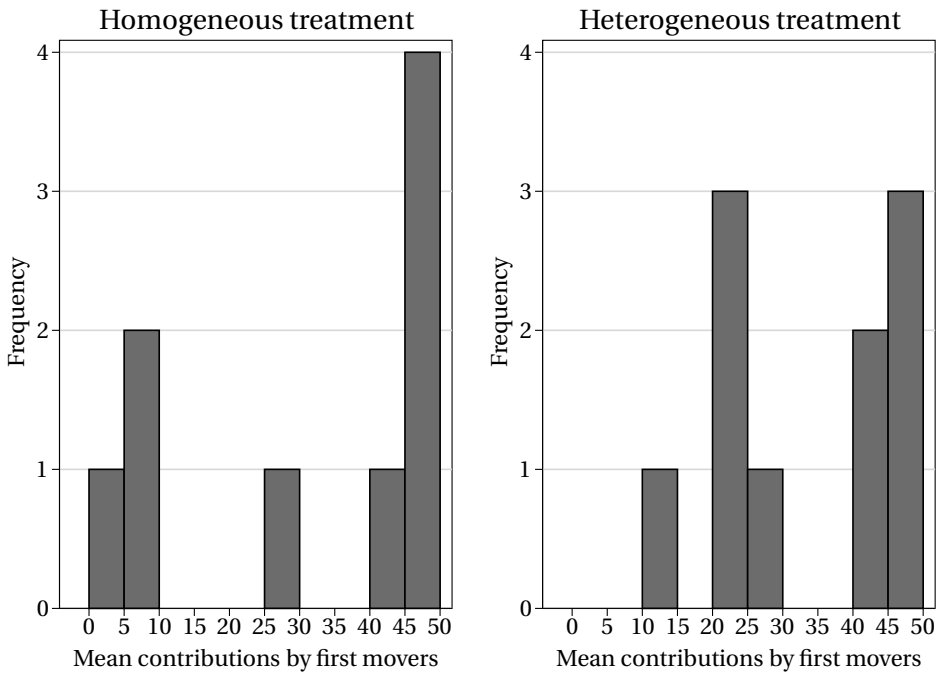


Figure 2.6: First movers' mean contributions to the collective accounts over all periods.

2.5.2 Contributions of First Movers

Mean contributions by first movers over all periods are depicted in Figure 2.6. Two thirds of first movers in the homogeneous treatment and even all first movers in the heterogeneous treatment contribute on average over all periods an amount of at least 10 tokens to the collective accounts which is the minimum required contribution of first movers such that followers can cover the remaining amount to reach a threshold with their endowment.

In line with conjecture (C1), we find that first movers in the heterogeneous treatment contribute on average over all periods an amount which equals at least the minimum required contribution of 10 tokens ($p = 0.0010$, one-sided sign test (one-sample)). For the homogeneous treatment, however, the null hypothesis of first movers' contributions below the minimum required threshold cannot be rejected ($p = 0.2539$, one-sided sign test (one-sample)).

The institution of a first mover is expected to facilitate coordination in an environment with several alternatives and scarce resources which permits the group to coordinate at most on one alternative successfully. By contributing a considerable amount to one particular collective account, one alternative becomes salient for followers who can align with the first mover's choice of the collective account which resolves one domain of the coordination problem. In both treatments, first movers make use of their strategic position to facilitate coordination in their groups: In 74.4% of all cases in the homogeneous treatment and in 83% of all cases in the heterogeneous treatment, first movers make one particular collective account salient by their contribution. The difference in proportions between both treatments is statistically not significant ($p = 0.6463$, two-sided Z-test (two-sample)). As collective accounts in the homogeneous treatment are indistinguishable regarding potential payoffs, first movers can choose, e.g., by label, by position on

the screen, or randomly. In 50% of all cases, first movers in the homogeneous treatment choose the collective account which appears on the first position on the screen (from left to right) after the private account.[6] This proportion is statistically significantly different from 25% ($p = 0.0416$, one-sided Z-test (one-sample)) which is the probability of choosing the collective account on the first position if one out four collective account is chosen randomly. In the heterogeneous treatment, there is one collective account for each individual which yields a higher potential payoff due to its marginal per-capita return of 0.75 compared to the marginal per-capita returns of 0.5 for all other collective accounts. Hence, first movers in the heterogeneous treatment choose their preferred collective account in 74% of all cases which is statistically significantly different from 25% ($p = 0.0002$, one-sided Z-test (one-sample)) and, therefore, supports conjecture (C2).

2.5.3 Contributions of Followers

As discussed in the previous section, the majority of first movers in both treatments make an effort to facilitate coordination in their groups by making one out of four collective accounts salient. Moreover, they contribute on average considerable amounts above the minimum required contribution such that followers can cover the remaining amount to reach a threshold with their endowment. While making one collective account salient resolves the underlying coordination problem in one domain, higher contributions by first movers reduce the remaining amount which has to be covered to reach the corresponding threshold and, thereby, also reduce strategic uncertainty of followers who decide simultaneously to some extent. Yet, contributing still inhibits risk for followers as they do not know whether and how much the other followers contribute and contributions are not refunded if a threshold is not reached. It is, therefore, under question whether followers respond to their first mover's contribution and support the attempt to reach a particular threshold.

As discussed in Section 2.5.1, there are only minor differences between first movers' and followers' contribution in both treatments. Figure 2.7 depicts followers' average contributions in a period contingent on the respective first mover's contribution. Followers in the homogeneous treatment respond positively to first movers' contributions which is also confirmed by a Spearman rank correlation coefficient on average contributions over all periods ($\rho = 0.9412$, $p = 0.0002$), whereas the rank correlation coefficient for the heterogeneous treatment is lower in magnitude and statistically insignificant ($\rho = 0.4787$, $p = 0.1617$). Nevertheless, there is a tendency in both treatments—albeit to different degrees—that followers are indeed willing to respond positively to their first movers' contributions which provides (weak) support for conjecture (C3a). While followers in the homogeneous treatment have no reason to object to their first mover's choice of a particular collective account, followers in the heterogeneous treatment face the decision whether or not to support the first mover if she chooses the collective account which yields the highest payoff for herself. Focusing only on a subset of cases in the heterogeneous treatment in which the first mover acts selfishly by making her preferred collective account salient, the correlation between first movers' and followers' contributions is slightly negative but statistically insignificant ($\rho = -0.0988$, $p = 0.7868$). The right panel of Figure 2.7, however, indicates that

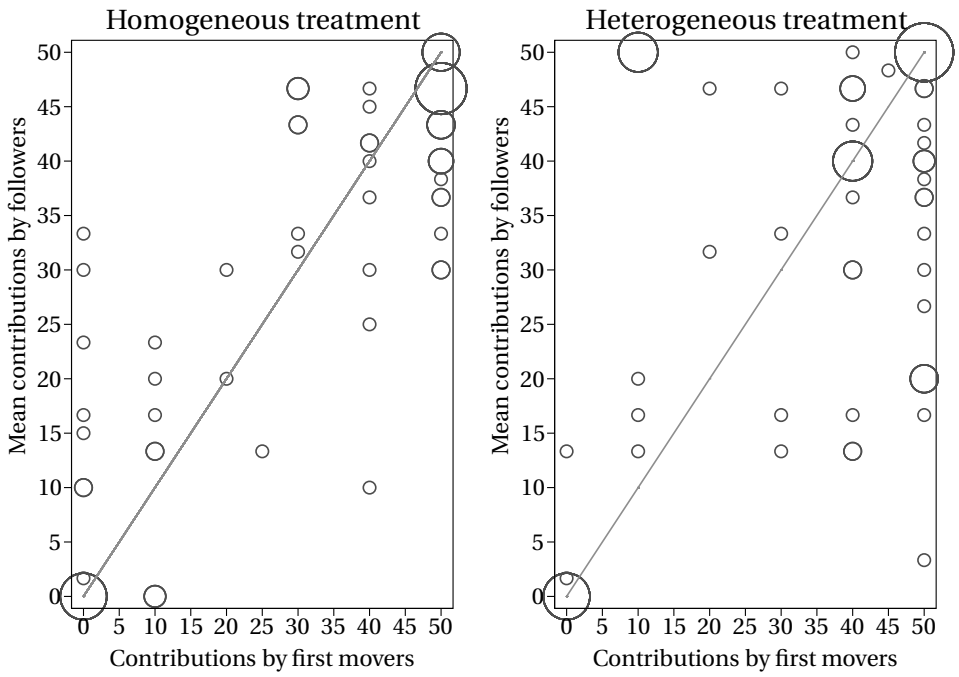


Figure 2.7: First movers' contributions and followers' mean contributions by period. The size of the bubbles indicates the frequency of distinct observations.

in 10 of 74 instances in which the first movers chooses the alternative which yields the highest potential payoff for herself *and* only contributes the minimum required amount of 10 tokens such that followers can still reach the threshold, followers do indeed contribute their entire endowment. This suggests—in line with conjecture (C3b)—that followers in general do not have a strong desire to punish first movers who exploit their position strategically and focus on reaching the corresponding threshold instead.

2.5.4 Regression Analysis

To get additional insights in individual contribution behavior, the following subsection reports results of panel regressions with random effects. To account for interdependencies between periods, robust standard errors are clustered on group level. Additional regressions on averages by groups over all periods as robustness checks are reported in Appendix A.2.

The dependent variables in Table 2.2 are individual contributions by all players. Columns (1) and (2) contain all periods, whereas column (3) only contains periods 2–10. Het and FM denote indicator variables for the heterogeneous treatment and the role of the first mover, respectively. Period is a continuous variable for the time trend and LagThr is a dummy variable which indicates a group's successful coordination on a threshold in the previous period. The dummy variables for the treatment and the first mover are statistically not significant; neither is the interaction term between these explanatory variables. The time trend is negative as commonly observed in repeated public good experiments but statistically not significant in specifications (1) and (2) either. When controlling for groups' successful coordination in the previous period, the

	(1)	(2)	(3)
Het	2.975	3.157	1.464
	(7.284)	(7.535)	(4.299)
FM	0.079	0.463	0.412
	(2.602)	(1.294)	(1.793)
Period	−0.796	−0.796	−1.461***
	(0.485)	(0.485)	(0.432)
Het × FM		−0.730	−0.949
		(4.977)	(5.104)
LagThr			20.460***
			(3.960)
Constant	34.735***	34.639***	28.745***
	(4.939)	(4.941)	(3.800)
Obs.	760	760	684
No. Groups	19	19	19
Wald-χ^2	3.045	6.462	31.33
Prob. > χ^2	0.385	0.167	$8.05 \cdot 10^{-6}$

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.2: Individual contributions by all players. Generalized least squares estimates with random effects.

time trend is statistically highly significant which indicates a decreasing effect of failure in the previous period on individual contributions over time. The dummy variable for successful coordination in the previous period itself is positive and statistically highly significant which reflects individuals' incentives to contribute a considerable share of their endowment in the present period and, thereby, increase the prospect of sustaining successful coordination. This observation is confirmed by column (2) of Table A.1 when using average contributions by group and period as dependent variable.

To analyze the inclination of first movers to contribute to the collective accounts to set a good example and guide followers towards successful coordination of the group, Table 2.3 includes several control variables on individual characteristics and attitudes elicited in the post-experimental questionnaire using primarily questions from the German Socio-economic Panel (DIW/SOEP, 2014). Econ is a dummy variable for individuals who study Industrial Engineering and Management, Information Engineering and Management, Economics Engineering, Economathematics, or a related subject. Risk is the value of the answer to the question "Would you describe yourself as someone who tries to avoid risks (risk-averse) or as someone who is willing to take risks (risk-prone)?" on an end-labeled Likert-scale ranging from 0 (risk-averse) to 10 (risk-prone) (Dohmen *et al.*, 2011). Q06 and Q07 refer to the answers on a four-point Likert-scale[7] to the questions "People can generally be trusted." and "Nowadays you cannot rely on anyone.", respectively. Finally, Q08 is a dummy variable for individuals who choose the second alternative as answer

	(1)	(2)	(3)	(4)	(5)
Het	0.046	–0.007	0.187	3.218	0.920
	(7.764)	(7.392)	(7.895)	(9.495)	(7.958)
Period	–0.992	–0.992	–0.992	–0.992	–0.992
	(0.637)	(0.639)	(0.641)	(0.641)	(0.641)
Econ	–5.972	–6.109	–5.973	0.499	–5.878
	(11.398)	(11.404)	(11.504)	(14.165)	(11.234)
Female	–2.650	–2.818	–3.121	–7.107	–4.740
	(12.545)	(12.703)	(15.117)	(10.778)	(13.333)
Risk		–0.067	–0.149	–1.112	–0.144
		(1.718)	(2.680)	(1.849)	(1.814)
Q06			–0.353		
			(7.244)		
Q07				7.922	
				(7.670)	
Q08					3.289
					(9.396)
Constant	42.076***	42.673***	43.828	21.715	41.210**
	(13.280)	(16.428)	(32.173)	(28.942)	(16.853)
Obs.	190	190	190	190	190
No. Groups	19	19	19	19	19
Wald-χ^2	2.665	2.736	2.775	6.110	2.786
Prob. > χ^2	0.615	0.741	0.837	0.411	0.835

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.3: First movers' contributions. Generalized least squares estimates with random effects.

to the question "Do you believe that most people *(i)* would use you if they had the chance or *(ii)* that they would try to be fair to you?". As can be seen in Table 2.3, various specifications including control variables on individual characteristics fail to explain first movers' contribution behavior. In particular, the statistics of the Wald-χ^2-test are very low, i.e., the null hypothesis of joint insignificance of all explanatory variables included in the respective specification cannot be rejected for any of these models. We can, therefore, not derive a reliable and meaningful conclusion on what drives first movers' contributions. The robustness check presented in table Table A.2 using first movers' average contributions over all periods as dependent variable does not improve this result.

Focusing only on contributions of followers, Table 2.4 includes Salient as a dummy variable which assumes a value of one if the first mover contributes a decent amount of her endowment to a single collective account and makes this option salient instead of spreading an equal amount on two or more collective accounts which is not informative for followers when trying to solve the first domain of the coordination problem as to which collective account they are supposed to contribute to. Additionally, ContFM corresponds to the contribution of the first mover to the collective account in the present period. Correct Belief is a dummy variable which indicates whether a follower correctly anticipated the account which the first mover allocated the largest share of her endowment to. As before, neither the coefficients of the dummy variable for the heterogeneous treatment nor the coefficients for the time trend are statistically significant. First movers who make use of the semi-sequential order and alleviate the coordination problem by making one collective account salient induce followers to contribute statistically significantly higher amounts to the collective accounts. Moreover, contributions of followers to the collective accounts rise statistically significantly in first movers' contributions which emphasizes the positive effect of good examples set by first movers. The interaction term between these two explanatory variables is negative, but statistically only marginally significant in column (6) when also controlling for correct beliefs and successful coordination the previous period. Followers who correctly anticipate the target of the first mover's allocation contribute statistically marginally more to the collective account in specification (4), although the coefficient loses its significance when controlling for successful coordination in the previous period. Again, the dummy variable for successful coordination in the previous period itself is positive and statistically highly significant. Changing the dependent variable to followers' average contributions by group as in Table A.3 does not change these results except the fact that the coefficient for the dummy variable indicating a correct belief in column (5) is statistically not significant any more.

Unlike the regressions of Table 2.4 which take followers' total contributions to the collective accounts into account, the results reported in Table 2.5 are based on contributions to the salient account—given that the first mover provides his group with a salient alternative. The coefficients of the dummy variable for the heterogeneous treatment and the variable for the time trend in columns (1) and (2) are statistically insignificant as before. The signs of the coefficients for the time trend, however, change to positive which can be interpreted as followers' improved

	(1)	(2)	(3)	(4)	(5)	(6)
Het	3.157	1.013	1.170	5.583	5.388	5.937
	(7.528)	(4.947)	(4.428)	(5.505)	(5.453)	(4.588)
Period	–0.731	–0.011	–0.012	–0.026	–0.176	–0.576
	(0.457)	(0.290)	(0.258)	(0.268)	(0.291)	(0.382)
Salient		25.066***	14.900***	14.743**	16.175***	14.019**
		(2.590)	(5.570)	(5.735)	(6.146)	(6.739)
ContFM			0.294**	0.382**	0.353**	0.344**
			(0.134)	(0.151)	(0.158)	(0.174)
Het × ContFM				–0.139	–0.133	–0.168*
				(0.111)	(0.110)	(0.099)
Correct Belief					3.895*	2.366
					(2.205)	(2.381)
LagThr						11.756***
						(4.087)
Constant	34.280***	11.662***	10.216***	7.686*	5.231	4.904
	(4.781)	(3.418)	(3.239)	(4.271)	(4.218)	(4.574)
Obs.	570	570	570	570	570	513
No. Groups	19	19	19	19	19	19
Wald-χ^2	2.732	119.5	161.6	223.3	265.6	721.3
Prob. > χ^2	0.255	0.000	0.000	0.000	0.000	0.000

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.4: Followers' contributions. Generalized least squares estimates with random effects.

	(1)	(2)	(3)	(4)	(5)
Het	1.220	0.620	19.263*	20.598**	20.585**
	(5.620)	(5.054)	(11.396)	(10.362)	(9.675)
Period	0.538	0.462	0.492	0.239	–0.151
	(0.409)	(0.357)	(0.380)	(0.346)	(0.289)
ContSalFM		0.368**	0.654***	0.628***	0.525***
		(0.156)	(0.138)	(0.146)	(0.131)
Het × ContSalFM			–0.479*	–0.508**	–0.500**
			(0.250)	(0.230)	(0.222)
Correct Belief				6.379**	4.130*
				(2.497)	(2.332)
LagThr					10.005***
					(3.613)
Constant	32.787***	19.061***	7.961	5.402	7.181
	(4.635)	(6.867)	(6.437)	(6.123)	(4.505)
Obs.	450	450	450	450	399
No. Groups	18	18	18	18	18
Wald-χ^2	1.781	9.745	27.94	39.69	74.07
Prob. > χ^2	0.410	0.021	$1.28 \cdot 10^{-5}$	$1.73 \cdot 10^{-7}$	0.000

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.5: Followers' contributions to the salient account. Generalized least squares estimates with random effects.

understanding to act on the first mover's suggestion over time. Contributions to the salient account by the first mover increase followers' contributions statistically (highly) significantly. The coefficients of the interaction term between the dummy variable for the heterogeneous treatment and contributions to the salient account by first movers is negative and statistically (marginally) significant indicating a weakly negative effect of first movers' contributions in the heterogeneous treatment. Yet, when controlling for this interaction, the coefficients of the dummy variable for the heterogeneous treatment themselves become statistically (marginally) significant and increase substantially in magnitude, i.e., the overall effect when controlling for this interaction term is still positive. Individuals who stated a correct belief as to which account the first mover is going to allocate the largest share of her endowment contribute statistically (marginally) more to the salient account. The coefficient of successful coordination on a threshold in the previous period is still positive and statistically highly significant. Considering the same specification with followers' average contributions to the salient account by group in Table A.4 only leads to changes in column (3) in which the coefficients of the dummy variable for the heterogeneous treatment and the corresponding interaction term with first movers' contributions lose their significance.

To investigate the relation between first movers' and followers' contribution in the heterogeneous

treatment more detailed, the results reported in Table 2.6 are based on followers' contributions to the collective accounts in the heterogeneous treatment. As in all previous regressions, the coefficients for the time trend are statistically not significant. A notable exception is the specification reported in column (5) containing a control for success in the previous period, i.e., this declining time trend can be attributed to groups which fail to coordinate successfully on a threshold. The dummy variable Selfish indicates whether the first mover chooses the collective account which yields the highest potential payoff for herself due to a higher marginal per-capita return. The corresponding coefficients are positive and statistically highly significant. As the choice of the preferred account by the first mover frequently coincides with making one out of four collective accounts salient, the positive reaction of followers is not surprising. In particular, followers understand the benefit of supporting even a selfish first mover compared to a situation in which no threshold is reached. As seen above, contributions by the first mover also increase followers' inclination to contribute to the collective accounts statistically significantly. The interaction term between a selfish choice of the first mover and her contributions yields negative and statistically significant coefficients when controlling for correct beliefs and successful coordination in the previous period although the overall effect is still positive. The coefficient of the dummy variable for correct beliefs is positive and statistically marginally significant in column (4) but loses its significance when controlling for successful coordination in the previous period which itself yields a positive and statistically highly significant coefficient. These results are by and large confirmed by additional regressions based on average contributions of followers to the collective accounts by group in Table A.5 with differences in the specification reported in column (5).

2.6 Conclusion

The present study investigates coordination in an environment with several alternatives and scarce resources such that only one of these alternatives is feasible. The coordination problem in this case comprises two domains: Firstly, individuals have to choose one particular alternative independently and secondly, they have to decide how much to contribute to reach the corresponding threshold. To facilitate coordination, we use a semi-sequential move order with a first-moving leader and second-moving followers who make their decisions simultaneously and independently. This structure can foster coordination as the allocation of the first mover serves as signal for followers. In particular, the first mover's choice of the collective account is decisive for the group which resolves one domain of the underlying coordination problem.

On the other hand, the institution of a first mover provides a large degree of discretion for one individual which can be exploited. In the present case, the first mover can make a contribution so low that followers have to contribute their entire endowment to reach the corresponding threshold. If individuals face divergent incentives regarding the collective accounts, the first mover can also strive for the alternative which yields the highest potential payoff for herself. Yet, followers have to recognize that supporting a selfish first mover makes them better off than a situation in which no threshold is reached. The first mover, in turn, bears the risk of losing her contribution to the collective account if followers do not act on her example.

	(1)	(2)	(3)	(4)	(5)
Period	–0.298	0.005	0.149	–0.035	–0.616***
	(0.473)	(0.304)	(0.366)	(0.392)	(0.220)
Selfish	24.703***	14.751***	29.174***	30.571***	23.898***
	(3.729)	(5.208)	(10.545)	(9.408)	(8.071)
ContFM		0.326***	0.455***	0.498***	0.438***
		(0.112)	(0.053)	(0.059)	(0.074)
Selfish × ContFM			–0.384	–0.466**	–0.451**
			(0.252)	(0.226)	(0.203)
Correct Belief				4.706*	1.189
				(2.524)	(3.567)
LagThr					15.810***
					(5.259)
Constant	16.775***	11.675***	7.201**	4.490	8.346***
	(4.508)	(4.137)	(2.865)	(2.871)	(2.974)
Obs.	300	300	300	300	270
No. Groups	10	10	10	10	10
Wald-χ^2	63.080	50.47	184.3	188.4	1422
Prob. > χ^2	0.000	$6.34 \cdot 10^{-11}$	0.000	0.000	0.000

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2.6: Followers' contributions in the heterogeneous treatment. Generalized least squares estimates with random effects.

Followers themselves also face strategic uncertainty as they decide simultaneously. The risk-dominant decision in this situation is to transfer their entire endowment to the respective private account. To avoid this, the first mover can reduce strategic uncertainty to some extent by a decent contribution above the minimum required amount to one particular collective account which reduces the amount which has to be covered by followers to reach the corresponding threshold.

Overall, we do not observe substantial differences in contributions and earnings between the treatments with homogeneous and heterogeneous incentives, nor between first movers and followers. The latter result provides evidence that first movers are on average indeed willing to bear a fair share in the group's effort to coordinate on a threshold and, therefore, set a good example for followers. Averaging over all periods, two thirds of first movers in the homogeneous treatment and all first movers in the heterogeneous treatment make a contribution to the collective accounts above the minimum required amount. The analysis of followers' contribution behavior also emphasizes the impact of the role of the first mover. Not only contributions by first movers induce followers to act on her example but also first movers who use the semi-sequential move order. Making one collective account salient by contributing a decent amount to one particular collective account serves as signal for followers' allocations which eventually fosters coordination. Even if the first movers behaves selfishly and the salient collective account coincides with her preferred alternative, followers recognize the benefits of supporting the first mover's attempt to reach the corresponding threshold.

To conclude, the institution of a first-moving can foster coordination in a setting with several alternatives of which only one is feasible. Despite potential drawbacks due to the discretion of the first mover, there is no indication of first movers who systematically exploit their role. On the contrary, first movers have strong incentives to contribute a decent amount in the heterogeneous treatment if they choose their preferred collective accounts. Followers in this case do not object to the first mover's choice and strive to cover the remaining amount to reach the threshold. Thus, using a semi-sequential move order with a first-moving leader and second-moving followers can facilitate coordination in a situation with several alternatives—even if individuals face divergent incentives.

Notes

1. Usually, the term step return introduced by Croson and Marks (2000) is used in the context of threshold public good games. This concept, however, requires a pre-determined size of the public good. This is the case under a so-called no refund rule, i.e., if contributions above the threshold do not increase the value of the public good. In both, the study by Corazzini *et al.* (2015) and the present study, a linear rebate rule is employed. Hence, contributions above the threshold are treated as in a linear public good game (see also Section 2.3.1).

2. Note that $\gamma_i < 1$ ($i \neq 1$) since none of the second-moving followers can reach the threshold

$\widetilde{T}$ in the subgame alone.

3. The game is depicted by the following matrix:

		Column player: Rhombus	Column player: Trapezoid
Row player	Square	7, 3	3, 2
	Rectangle	1, 2	4, 6

   The first number in each cell denotes the payoff of the row player and the second number the payoff of the column player. There are two equilibria in pure strategies, (Square, Rhombus) and (Rectangle, Trapezoid). When asked to indicate their decisions as colum player conditional on the row player's decision, all but three individuals chose "Rhombus" conditional on "Square" and "Trapezoid" conditional on "Rectangle" which indicates that individuals understood the underlying conflict.

4. The figures in Appendix A.1 depict total contributions to the groups' preferred collective accounts by period in the homogeneous and heterogeneous treatment, respectively. Note that individuals may contribute to more than one collective account.

5. In the present study, the threshold of 160 tokens corresponds to 80% of a group's total endowment in each period, whereas Corazzini *et al.* (2015) require a total contribution by the group to a particular collective account of 132 tokens which is equivalent to 60% of the group's endowment in their setting.

6. See Figure A.3 for a screenshot. Collective accounts are labeled with the colors red, blue, yellow, and green. The order of collective goods varies between periods.

7. 1 = strongly agree; 2 = agree; 3 = disagree; 4 = strongly disagree.

3 Cooperation in the Classroom[†]

In this paper, we study the effectiveness of an intervention which conveys social skills in a playful way during physical education lessons to foster cooperative behavior of children aged 10–13 years. Situations which require cooperation of individuals frequently suffer from free riding behavior if individual incentives are in contrast to collective benefits. Cooperative behavior as one aspect of social skills has been identified as major prerequisite for success in the labor market and at the workplace as companies increasingly rely on teamwork. To study the effectiveness of the intervention, individuals participate in an incentivized linear public good game at three different times during the course of the study. The results of the present study provide weak evidence that our intervention can foster cooperative behavior despite its limited duration. In particular, individuals from households with low socio-economic status benefit from the intervention.

3.1 Introduction

Social dilemmas are common in human societies, ranging from efforts to mitigate the effects of climate change (e.g., Milinski *et al.*, 2006) to collaboration in the workplace (e.g., Gneezy *et al.*, 2016). The main characteristic shared by these situations is the conflict between individual incentives and collective benefits, i.e., the socially optimal outcome is only achieved if all individuals are committed to fully cooperate, whereas it is individually optimal to free ride and not to cooperate at all and rely on contributions by others instead. With respect to the workplace, team production requires contributions of several individuals which cannot be perfectly observed, whereas gains from cooperation are afterwards shared equally among individuals. The lack of perfect observability of individual contributions while being entitled to a fixed share of the gains from cooperation induces individual incentives to free ride (Alchian and Demsetz, 1972; Kandel and Lazear, 1992). This leads eventually to a situation in which all individuals refrain from cooperation and efficiency gains due to collective actions cannot be realized (Kim and Walker, 1984; Andreoni, 1988; Ledyard, 1995; Nalbantian and Schotter, 1997). On the other hand, the willingness to cooperate becomes increasingly important in the workplace as the share of companies which rely on teamwork has increased over the past decades (Eckel and Grossman, 2005; Lazear and Shaw, 2007; Bandiera *et al.*, 2013). Moreover, Deming (2017) reports an increase in employment and wages in the United States for occupations which require a high degree of social skills[1] as prerequisite of efficient collaboration by reducing coordination costs (see also Heckman, 2000; Enste *et al.*, 2018; Acosta and Muller, 2018).[2] At the same time, employers are concerned with a lack of soft skills among applicants. While employability is often associated with hard skills in the first place which refer to formal qualifications, soft skills which comprise individual characteristics related to general readiness for employment, such as willingness to work in teams, reliability, problem solving, and attitude to work, are equally relevant (Kautz *et al.*,

[†]This section is based on joint work with Petra Nieken, Hagen Wäsche, Rita Wittelsberger, and Alexander Woll (Haas *et al.*, 2019b).

2014; Cobb-Clark, 2015; European Commission, 2017).[3]

One approach frequently proposed to close the gap between employers' expectations and applicants' actual sets soft skills is sports participation (European Commission, 2016, 2017). Although causal evidence is rare (Cabane and Clark, 2015; Schüttoff *et al.*, 2017), there seems to be broad agreement on the benefits of sports participation beyond direct effects on health and physical fitness (e.g., German Bundestag, 2014). The transfer of soft skills is argued to be facilitated by employing sports participation as means of informal learning. In particular, organized forms of sports participation, e.g., in sports clubs which is a popular leisure time activity of children and adolescents in Germany, are regarded favorable for the development of soft skills (Felfe *et al.*, 2016; German Bundestag, 2017). Effects which are commonly attributed to sports participation and increase employability at the same time include predominantly non-cognitive skills such as the willingness to cooperate and to work in teams, self-discipline, communication, and resilience, as well as "managerial" skills which are essential to carry out complex and extensive tasks (Cornelißen and Pfeifer, 2009; Lechner and Sari, 2015; Cabane and Lechner, 2016; Felfe *et al.*, 2016; European Commission, 2017). One possible explanation is that organized forms of sports participation require individuals to interact with other people who do not genuinely belong the individual's social network otherwise (Cabane and Clark, 2015; Felfe *et al.*, 2016). Thus, regular and active sports participation and being embedded in an organization fosters the capability to cooperate with others to achieve common goals which is also crucial for collaboration in the workplace as it, e.g., refrains individuals from shirking (Alchian and Demsetz, 1972; Long and Caudill, 1991; Celse *et al.*, 2017). Yet, contributions which study the effects of sports participation on the development of soft skills in young people are scarce (Pawlowski *et al.*, 2018).

To contribute to this field of research, we conduct an intervention study. The novel approach used here comprises the elicitation of preferences in incentivized standard economic games including a linear public good game and the transmission of social skills in playful way during purposefully designed physical education lessons. Hence, the goal of this paper is to study the impact of an intervention employing enhanced physical education lessons on cooperative behavior as well as its transmission to situations outside the gymnasium which require balancing off individual incentives against collective benefits. To complement previous findings from the literature (see Section 3.2), we study whether *(i)* individuals who participate in organized sports activities, e.g., in a sports club, are more cooperative and *(ii)* the intervention which aims at conveying social skills in a playful way enhances cooperative behavior compared to individuals in the Control group who do not receive the intervention.[4]

3.2 Related Literature

There is broad evidence on health-related benefits of sports participation and physical activity in general (e.g., Physical Activity Guidelines Advisory Committee, 2018). While physical inactivity is a major cause for non-communicable diseases which account for almost half of the overall

burden of diseases and 6% of deaths worldwide (Warburton *et al.*, 2006; WHO, 2010), physical activity has been recognized as effective treatment for a wide range of diseases (see, e.g., Pedersen and Saltin, 2015, for a comprehensive overview; see also Reiner *et al.*, 2013; Warburton and Bredin, 2016).

At the same time, positive externalities of sports participation can also be observed in other domains not directly related to health. An exhausting overview in this regard is provided by Breuer *et al.* (2016) who identify non-monetary outcomes of sports participation such as increased social capital, improvements in educational attainments, and success in the labor market. Furthermore, these effects translate into monetary returns from sports participation which result, e.g., in higher wages on an individual level. While the relation between sports participation and labor market outcomes is widely acknowledged (e.g., Cornelißen and Pfeifer, 2009; Lechner, 2009, 2015; Lechner and Sari, 2015; Cabane and Clark, 2015; Cabane and Lechner, 2016), the underlying mechanism is not yet fully understood. Lechner (2009) proposes three explanations that can potentially establish this link: *(i)* As sports participation is an important determinant for physical health, individuals who participate in sports are generally expected to be healthier which results in higher individual productivity in the labor market and, thereby, results in higher levels labor market participation and higher wages. *(ii)* As sports participation is often associated with a high degree of social interactions, individuals who participate in sports are also likely to exhibit skills which are valuable in the workplace when assignments require joint endeavors of several individuals. *(iii)* Finally, sports participation can be interpreted by potential employers as signal of an individual's motivation to exert oneself for achieving good results also in domains related to the workplace. Moreover, Cabane and Lechner (2016) emphasize the impact of sports participation on non-cognitive skills which include self-discipline, the ability to perform well even under stressful circumstances, and the willingness to cooperate with other individuals in a team.

As noted before, it is not straightforward to draw causal conclusions on the relation between sports participation and outcomes in other domains as there may be unobserved causes which influence both, sports participation and the outcome of interest. There are, however, some studies which employ sophisticated econometric approaches to disentangle potentially confounding effects to establish a causal relation. Felfe *et al.* (2016) study health, performance at school, and behavioral development of children aged 3–10 years in Germany conditional on their sports participation. Data is primarily provided by the 2003–2006 wave of the German Health Interview and Examination for Children and Adolescents (KiGSS; see Kurth, 2007, for a description). Additionally, they access the German Child Panel and use its structure to avoid issues due to selection effects of sports participation. As a result, Felfe *et al.* (2016) find children who participate in sports to score better in all aspects under study compared to their peers who do not participate in sports. Namely, sports participation improves children's health status and performance at school and reduces antisocial behavior. To investigate the effect of sports participation during childhood on labor market outcomes as adults, Cabane and Clark (2015) use data of the National Longitudinal Study of Adolescent and Adult Health (Add Health) from

the United States. Individuals participated for the first time in this study in the years 1994–1995 when they were in grades 7–12. The third and final follow-up survey was conducted in 2008 when individuals were 24–32 years old. Information about labor market status and other control variables, e.g., on health, lifestyle, and education from the final survey were compared to different types of sports participation (individual vs. team) and other leisure activities during childhood. Cabane and Clark (2015) consider different aspects of adult individuals' labor market outcomes, in particular *(i)* having a paid job with at least ten working hours per week, *(ii)* job satisfaction, *(iii)* being in a position with managerial responsibilities, *(iv)* opportunity to decide independently on material issues in the workplace, and *(v)* annual income. Indeed, significant correlations between sports participation during childhood and adult labor market outcomes 13 years later are found. The link between participation in individual sports and managerial responsibility is particularly pronounced, whereas participation in team sports is strongly associated with discretion over important decision in the workplace. Schüttoff *et al.* (2017) study the impact of sports participation on social capital formation during adolescence which is considered crucial for success in the labor market. They use answers from the German Socio-Economic Panel (SOEP) at two different points in time. First, individuals born between 1986–1995 answer the youth questionnaire which contains questions about the frequency of sports participation and the organizational format at age 17. The regular personal questionnaire is answered at age 18–19. This time, the questionnaire contains items which allow to derive conclusions on individuals' social engagement reflected by the frequency of voluntary work, e.g., in clubs or social services or the frequency of civic engagement. Interpersonal networks as additional aspect of social capital is measured by the frequency of helping friends, neighbors, or relatives. Sports participation is classified by type (team vs. individual) and organizational format (sports club vs. other organizational format). The results show that especially the formation of interpersonal networks is fostered by sports participation during adolescence, whereas only small effects on civic engagement can be observed. With respect to the organizational format, sports clubs are apparently more relevant for social capital formation during adolescence than other organizational formats. Whereas the studies discussed so far are limited to developed countries, Pawlowski *et al.* (2018) report results of a study conducted with data from Peru. Data were collected as part of the Young Lives study program. Beginning in the year 2002 at age 12, individuals answered the questionnaire which contained items also on other domains of individuals' lives beyond sports participation three times in total. Additional information were elicited by a household questionnaire which was answered by the persons primarily responsible for the individuals' upbringing. Outcome variables of particular interest for the study by Pawlowski *et al.* (2018) include human capital as measured by the Peabody Vocabulary Picture Test, social capital which is further divided into neighborhood trust, friends support, respect by other children, and feeling safe outside the house, subjective well-being as satisfaction with life, and health capital as a subjective indicator assessed by individuals themselves. Whereas individuals who participate in sports report higher levels of subjective health status and exhibit higher measures for social capital, a relation between sport participation and educational outcomes which are likely to

influence labor market outcomes in later life is not observed which is in contrast in contrast to the previous studies focusing on developed countries. Another aspect studied by Knaus *et al.* (2018) is the impact of physical education in Germany on different outcome variables, including cognitive and non-cognitive skills. They argue that physical education is a mandatory element of the curriculum, whereas participation in sports clubs is voluntary, i.e., focusing on sports clubs only captures individuals who are already physically active. Based on data from the activity survey of the 2015–2017 wave of the *Motorik-Modul* longitudinal study (MoMo; see Woll *et al.*, 2017, for a description), they measure the impact of an additional lesson of physical education. The strongest effect is on cognitive skills as measured by improvements of grades in German and math; in particular, boys from low-income households benefit in this regard. Considering non-cognitive skills, however, the results are less favorable. Whereas girls benefit from an additional lesson of physical education as indicated by a decrease in emotional symptoms, the effect is reversed for boys. More specifically, they are more likely exhibit behavioral problems. Positive outcomes for motor skills and physical activity are only found for girls, whereas health-related effects cannot be observed at all.

To address the impact of sports participation as signal to potential employers during the hiring process, Rooth (2011) conducts a correspondence study. The goal here is to investigate callback rates of fictitious applications with varying information about type (team vs. individual[5]) and intensity (recreational vs. competitive) of leisure sports to real job openings in Sweden. He observes an overall increase in callback rates for those applications which contain information about leisure sports of about 2 percentage points. This effect is statistically significant and corresponds reportedly to two additional years of job experience. Results are heterogeneous with respect to the type of sports and across occupations. A direct comparison reveals that the increase in callback rates is higher for individual sports than for team sports. Moreover, applications to job openings in physically demanding occupations benefit exceptionally from mentioning sports participation with an increase in callbacks of 4.8 percentage points. In a related approach, Piopiunik *et al.* (2018) use a broader set of signals relating to cognitive skills, social skills, and maturity in a survey experiment with 579 HR managers in Germany who are presented two resumes simultaneously and are then asked to decide which applicant they would prefer to invite for a job interview at their company. The fictitious applicants are either graduates of a secondary school applying for an apprenticeship or college graduates. The domain of social skills is represented by information about volunteering such as neighborhood help, youth work, elderly work, and offering German language courses or participation in team sports. Academic performance is deliberately negatively correlated with other skill signals. Overall, HR managers take skill signals from all three domains used in the study into account which eventually results in an increase in the probability of being invited to a job interview. The type of signal which is most valuable, however, differs between both types of applicants: Whereas HR managers primarily focus on grades as one aspect of cognitive skills when hiring college graduates, applicants for an apprenticeship benefit from signaling social skills instead. Moreover, the specific type of engagement in the domain of social skills also differs between graduates of

secondary schools and colleges. Whereas volunteering is a valuable signal for applicants for an apprenticeship, college graduates benefit conversely from mentioning participation in team sports. The authors explain this observation by the fact that participation in team sports is quite common at the age when young people in Germany typically graduate from secondary school and apply for an apprenticeship, whereas volunteering of college graduates may be perceived as strategic act to mimic social skills. Heinz and Schumacher (2017) compare signals from individuals' resumes about their willingness to cooperate with their actual behavior in a linear public good game. Therefore, they ask individuals to bring their resumes to the experiment. Willingness to cooperate can be signaled by social engagement in terms of voluntary work with individuals who need support or engagement in an association such as sports clubs,[6] student associations, or political parties. To take different degrees of engagement into account, the relevant information in individuals' resumes are rated on a scale from 0 to 10 by other students in a separate part of the study. The first finding of this study is that students who are socially engaged contribute more to the collective account in the public good game than students whose resumes do not contain references about social engagement. Moreover, contributions in the public good game and the intensity of social engagement are positively associated. Contributions of students who are engaged in an association, however, are not different from their non-engaged peers' contributions. Additionally, Heinz and Schumacher (2017) ask HR managers to predict individuals' contributions in the public good game based on signals contained in individuals' resumes about their willingness to cooperate. Predictions are particularly accurate for individuals who are highly socially engaged which can be interpreted as HR managers' perception of high commitment to social engagement as convincing signal of applicants' willingness to cooperate in the workplace.

To investigate differences in cooperative behavior of individuals who exhibit a high intensity of sports participation compared to average individuals, Celse *et al.* (2017) exploit a special subject pool consisting of students enrolled in sports sciences and students of other fields of study. For the former group, sports participation is an integral part of their everyday lives as they pursue an academic degree which requires them to take mandatory sports classes and the majority of these students additionally participates in high-level competitions on a regular basis. Both group of students participate in a linear public good game following Fischbacher *et al.* (2001) where they make one unconditional and 21 conditional decisions. The pattern of conditional conditions allows to classify individuals as conditional cooperators, free riders, or triangle contributors. Regarding unconditional contributions, Celse *et al.* (2017) do not observe significant difference between both groups of students. The distribution of types according to conditional contributions, however, differs. The share of conditional cooperators among students enrolled in sports sciences is higher than among students of other fields of study. Yet, a closer inspection of the data reveals that athletic individuals who are classified as conditional cooperators are more likely to undercontribute. On the other hand, the proportion of free riders is higher among individuals whose intensity of sports participation is not exceptionally high. The authors note that success in sports requires both, the willingness to compete and the willingness

to cooperate. What appears contradictory at first glance can be illustrated by the fact that most types of sports require at least some degree of cooperative behavior, whereas other situations in sports promote pursuing one's own interests. Hence, Celse *et al.* (2017) argue that the higher inclination for conditionally cooperative behavior among students enrolled in sports sciences reflects the notion that they are indeed willing to cooperate as means to achieve their goals, whereas they are not generally more cooperative than their peers who participate in sports at lower intensity levels.

Another strand of literature relevant for the present paper discusses the effectiveness of interventions to foster social skills. Considering a broad range of studies on intervention programs for children and adolescents, Kautz *et al.* (2014) find approaches which focus on the development of non-cognitive skills of preschoolers and primary school children to be most promising. While both, cognitive and non-cognitive skills, are malleable at young ages, the gap in efforts necessary to achieve desirable outcomes widens over the years. The reason is that the prefrontal cortex which is the region of the brain associated with non-cognitive skills does not mature before the early 20s and, thus, later than other parts of the brain (Cunha *et al.*, 2006; Howard-Jones *et al.*, 2012; see also OECD, 2015). As a result, there are also interventions aiming at improvements of non-cognitive skills of older children and adolescents (e.g., Currie, 2001; Blau and Currie, 2006; Almond and Currie, 2011; Kautz *et al.*, 2014; Fryer, 2017; Almond *et al.*, 2018). However, most of the related studies evaluating these programs focus on academic attainments, but also on welfare dependence and criminal behavior later in life. In contrast to that, Kosse *et al.* (2018) conduct a large-scale study with primary school children in Germany. They investigate the impact of a mentoring program on the development of prosocial behavior of children from households with low socio-economic status. More specifically, children are assigned to a mentor whom they regularly meet for joint leisure activities over the course of one year. Additionally, prosocial behavior including altruism, trust, and other-regarding preferences of children under study is elicited in a series of incentivized economic games before and right after the intervention as well as two years after the end of the intervention. Self-reported information about the prosocial behavior of children's mothers are also taken into account. At the beginning of the study, prosocial behavior of children from households with low socio-economic status lags substantially behind their peers from households with high socio-economic status. Moreover, children's prosocial behavior is strongly linked to their mothers' attitudes. After the intervention, this gap does not occur any more. Furthermore, the effect is enduring as even two years after the end of the study no differences in prosocial behavior of children from households with low and high socio-economic status can be observed. According to Kosse *et al.* (2018) there are two main reasons which explain this result. As part of the intervention, children from households with low socio-economic status are assigned to a mentor who are usually university students and expend a notable share of their leisure time to volunteer in this program and, hence, can be considered as a role models who convey prosocial patterns as part of informal learning. Furthermore, the high intensity of social interaction between the child and the mentor is also expected to affects the development of prosocial behavior positively. The authors argue that about 60% of the overall effect can be

	Female	Male	Total
Control	27 (45%)	33 (55%)	60
Treatment	23 (50%)	23 (50%)	46
ESS	11 (46%)	13 (54%)	24
	61 (47%)	69 (53%)	130

Table 3.1: Number of observations by group and gender.

attributed to these factors, where the exposure to a role model accounts for about two thirds of this particular effect and the high intensity of social interactions between the children and their mentors explains remaining third.

3.3 Data Set

The study was conducted at five different upper secondary schools in Karlsruhe (Germany) between December 2016 and April 2017. At four schools, two classes of the sixth grade were randomly selected to participate in the study; at the fifth school, one randomly selected class of the sixth grade participated in the study. The specific characteristic of this school is its status as so-called elite school of sports (ESS) which offers conditions to reconcile compulsory school attendance with extensive practicing necessary to pursue a professional career in sports and participation in national and international competitions (e.g., Wartenberg *et al.*, 2014). The parents of pupils attending the classes in question received a letter via the respective schools before the beginning of the first part of the study. This letter contained information about the opportunity to let their children participate in a scientific study during regular lessons and a consent form which had to be returned to the school if they agreed on their children's participation in the study. Neither the teachers nor the parents or their children were aware of the purpose of this study. The study was approved by the superintendent of the local school district (ref.-no. 71 c2-6499.25) and the board of ethics of the Karlsruhe Institute of Technology (see Appendix B.3 and Appendix B.4).

In total, 199 individuals aged 10–13 years returned the consent form prior to the first part of the study (response rate: approx. 87%). At schools with two classes under study, one class was randomly assigned to the Treatment and the Control group, respectively. The single class at the ESS was not under treatment.

The study comprises three measurements and an intervention (see Figure 3.1). In all three measurements, individuals made decisions in a series of different standard economic experiments. The purpose was the elicitation of preferences and personality traits in different economic environments (see Haas *et al.*, 2019c, for more details). As the present paper focuses primarily on cooperative behavior, contributions in a linear public good game are of particular interest. Furthermore, competitiveness was elicited by making individuals choose between two payment schemes in a real-effort task. For a detailed description of the procedure, see Section 3.4.1. Additionally, individuals received questionnaires which had to be answered at home. Questions

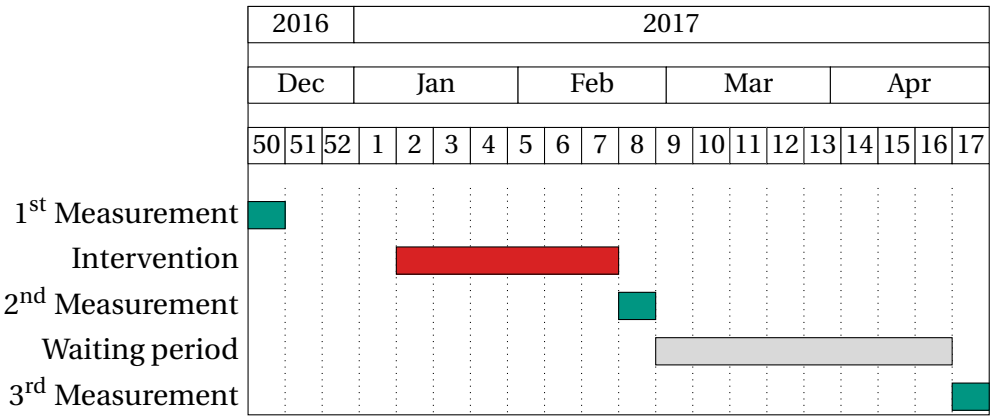


Figure 3.1: Study set-up.

covered primarily free time activities (DIW/SOEP, 2015c) and engagement in physical activities (Bös *et al.*, 2009) as well as several psychometric scales such as the Big Five inventory (Weinhardt and Schupp, 2011), social self-efficacy and self-efficacy of working in a team, perspective-taking, rivalry, and understanding of democracy (Jerusalem *et al.*, 2009; Kunter *et al.*, 2002; Eder, 1998; Abs *et al.*, 2007).[7] Along with the first measurement, parents filled out a questionnaire to provide additional information about their children, characteristics of the household, and their own educational backgrounds (based upon DIW/SOEP, 2015a, 2015b).

In the intervention, physical education lessons of classes assigned to the Treatment group were supplemented by a novel concept which had been developed to convey social skills in a playful way. Instructors were students of sports sciences who were pursuing a university degree which qualifies for teaching at upper secondary schools. They were not aware of the purpose of this study. Each lesson covered a certain topic (e.g., cooperation, fairness, or competitiveness). The sequence of a lesson followed a fixed protocol which included two blocks of ten minutes at the beginning and the end of each lesson in which individuals had the opportunity to discuss their opinions and experiences on the respective topic. The main part of each lesson involved physical activities which had been purposefully designed to put individuals into situations which reflect issues related to the underlying topic of the particular lesson. A detailed description of the intervention is available in Woll *et al.* (2018).[8]

The first measurement took place in December 2016. Subsequently, classes assigned to the Treatment group took part in the intervention for a period of six weeks in January and February 2017. The second measurement took place in February 2017 immediately after the end of the intervention in the Treatment group. The final measurement took place in April 2017 after a waiting period of eight weeks. The repeated measurements of decisions in different economic environments using incentivized standard economic experiments were necessary to study both, the immediate impact of the enhanced physical education lessons on individuals' behavior as well as potential long-term effects.

3.4 Set-Up

3.4.1 Experimental Design

To elicit individuals' cooperative behavior, we use a linear public good game. Competitiveness is measured by a binary choice between two compensation schemes a in a real-effort task which comprises a piece-rate based on individual performance and a winner-takes-all tournament based on relative performance. Both experiments are described below along with a brief overview of findings from related studies.

Public Good Game The public good game constitutes a social dilemma in groups with two or more individuals indexed i as they face the choice between individual incentives to free ride and collective benefits which arise from cooperation (Ledyard, 1995; Chaudhuri, 2011). Individuals are endowed with 120 tokens each which they can allocate to a private account or to a public account in increments of 24 tokens. If the experiment is selected for payment, four randomly selected individuals of a class form a group. The number of tokens an individual allocates to the private account is multiplied with a factor of two and returned to the individual. For the collective account, the number of tokens allocated by all individuals of the group is totaled. Each token allocated to the collective account by the group—no matter by whom—yields a payoff of one token for every individual of that group. Thus, individual i's payoff π_i is given by

$$\pi_i = 2 \cdot (120 - c_i) + \sum_{i=1}^{4} c_i,$$

where $c_i \in \{0, 24, 48, 72, 96, 120\}$ denotes the individual's contribution to the collective account. The marginal per-capita return (MPCR) which equals the ratio of an individual's return from a contribution to the collective account relative to a contribution to the private account is 0.5.

The dilemma arises in a public good game with $n \geq 2$ individuals if $\frac{1}{n} < \text{MPCR} < 1$ as individual costs for contributing to the funding of the public good exceed the associated individual benefits from its provision. A selfish individual who maximizes her own monetary payoff, therefore, transfers the entire endowment to her private account. This situation constitutes the Nash equilibrium in which no individual of a group contributes any tokens to the collective account. The social optimum, however, is reached when all individuals fully cooperate and contribute their entire endowment to the collective account (Andreoni, 1988). Empirical results of public good games frequently report contributions between 40–60% of individuals' endowments in one-shot settings and the first round of repeated games, respectively (Ledyard, 1995; Chaudhuri, 2011). Harbaugh and Krause (2000) find no difference in children's contribution patterns compared to adults. Similarly, in a study at an elementary school in the United States with children predominantly from Hispanic and African-American families, individuals contribute on average 55% of their endowment to the collective account (Cipriani *et al.*, 2013). In a study among children aged 10–16 years focusing on differences in cooperative behavior between different school tracks in Germany, John and Thomsen (2015) observe an average contribution in the

public good game of 43% of the endowment. 7% of individuals in their study are free riders and 10% contribute their entire endowment. Overall, there are no differences between school tracks. Hermes *et al.* (2019) study cooperative types according to Fischbacher *et al.* (2001) among first-graders (aged 6 years) in Germany. In the unconditional setting, individuals contribute on average 37.4% of their endowment. Regarding different types of cooperators, Hermes *et al.* (2019) find a distribution which approximately corresponds to previous observations among adult individuals. In particular, one third to one half of the first-graders under study can be identified as conditional cooperators.

Encryption Task The encryption task is a real-effort task in which individuals have to encrypt words by replacing letters in two subsequent steps according to given replacement tables (adopted from Erkal *et al.*, 2011). Output is measured in terms of the number of correctly encrypted letters in a two minute working period. Individuals can choose between two payment schemes—piece-rate and tournament—before they start working on the encryption (Niederle and Vesterlund, 2007; Booth and Nolen, 2012; Samak, 2013; Almås *et al.*, 2016). If this experiment is selected for payment, an individual who chose the piece-rate scheme receives 3 tokens for each correctly encrypted letter with a cap at 300 tokens. The tournament scheme yields a winner prize of 600 tokens and a loser prize of zero tokens. If this experiment is selected for payoff, the output of an individual who chose the tournament scheme is compared to the output of a randomly selected individual of the same class. If the individual in question has the higher output, she receives the winner prize and the loser prize otherwise; ties are broken randomly.

The choice between a piece-rate scheme based on individual performance and a tournament with relative performance evaluation and a winner-takes-all scheme appeals to individuals' competitiveness, i.e., individuals who like to compete opt for the latter alternative. Typical findings are that three quarter of men and only on third of women choose the tournament scheme (Niederle and Vesterlund, 2007; see also Booth, 2009, and the references cited therein). A similar pattern, albeit with a slightly smaller gender difference in competitiveness, is observed by Almås *et al.* (2016) among adolescents aged 14–15 years in Norway. Booth and Nolen (2012), however, find in a comparison of pupils at single-sex and coeducational schools of the same age evidence that the gender difference in competitiveness may be attributed to socialization. A comprehensive survey on findings from the related literature is provided by Sutter *et al.* (2019).

3.4.2 Procedures

The experiments described above are part of a series of standard economic experiments which was repeated at three times throughout the study (see Section 3.3). Each session was conducted during regular lessons in classrooms. Experimenters were students of different fields of study who were pursuing a university degree which qualifies for teaching at upper secondary schools; they were not aware of the purpose of this study.

At the beginning of each session, individuals received cards from their teacher with an ID number

which was the same for a given individual in all three sessions. This pseudonymization was necessary to link decisions with individuals while keeping their identities under disguise and, thereby, allowing for anonymity. Experimenters were not informed about the identity of an individual and the corresponding ID number. After that, individuals were allocated in the classroom such that communication was impeded and they were able to make their decisions individually and independently. Before the first experiment started, general instructions were read aloud by one of two experimenters who conducted a session. After clarifying questions had been answered, individuals received the answer sheet for the first experiment. This sheet contained the individual's ID number, instructions for the experiment such as their action space and the consequences thereof regarding the payoff, and space to indicate their decisions. Again, the experimenter read the instructions aloud and answered clarifying questions before individuals were asked to make their decisions individually and independently. After all individuals had made their decisions, sheets were collected by the second experimenter and the subsequent experiment started following the same protocol. For a detailed description of the experiments discussed in this paper, see Appendix B.2.

After the last experiment, the teacher was asked to open an envelope containing a card which stated the number of the experiment which was relevant for payment. For each individual, the amount of tokens earned in this particular experiment was calculated and converted into Euro at an exchange rate of 75 tokens = 1.00 Euro. Additionally, each individual received a participation fee of 2.00 Euro. Payoffs in the experiments were designed to ensure that individuals' maximum payoff from a session was 10.00 Euro which corresponds to approximately 50% of the recommended monthly amount of pocket money for children aged 11–12 years in Germany (Langmeyer and Winklhofer, 2014). The teacher handed out sealed envelopes which contained the payoff of each individual in Euro in exchange for the ID card. Individuals were explicitly told not to open the envelopes in class to prevent potential tensions. On average, individuals received 6.37 Euro per session.

3.5 Results

3.5.1 Overview

Prior to the intervention in the Treatment group, all groups participated in a series of standard economic games, including the linear public good game and a real-effort task to elicit competitiveness. Moreover, individuals and their parents answered a questionnaire to gather additional information—*inter alia*—about membership in sports clubs, individual levels of physical activity, and socio-economic status.[9] Before analyzing contributions to the collective account in the public good game, an overview of individuals' characteristics on group level is provided.

As can be seen in Table 3.2, the ratio of female and male individuals is almost balanced in all three groups with no statistically significant differences between the three groups ($p = 0.871$, Pearson's χ^2-test). Socio-economic status,[10] however, differs between the groups ($p = 0.044$, Pearson's χ^2-test) as the share of individuals from households with low socio-economic status

	Control	Treatment	ESS
Share of female individuals [%]	45.00	50.00	45.83
Share of individuals from households with low socio-economic status [%]	23.33	26.09	50.00
Share of sports club members [%]	86.67	65.22	91.67
Time spent at sports club [mins./week]*	163.34	133.97	271.81
Total physical activity, undifferentiated [days]	4.53	4.39	5.45
Share of tournament choice [%]	38.33	39.13	75.00

* Truncated at 600 minutes per week.

Table 3.2: Overview of control variables.

in the ESS group is statistically significantly higher than in the Treatment and Control group (ESS vs. Treatment: $p = 0.090$, ESS vs. Control: $p = 0.051$, pairwise Pearson's χ^2-test with Holm-Bonferroni adjustment). The shares of individuals who are member of at least one sports club is in all three groups at least as high as the share among the corresponding age cohort (66%) of the 2009–2012 wave in the representative MoMo study in Germany (Schmidt *et al.*, 2017). The shares of sports club members in the present study are not equal among the individual groups ($p = 0.007$, Pearson's χ^2-test) due to the statistically significantly lower share in the Treatment group (Treatment vs. Control: $p = 0.027$, Treatment vs. ESS: $p = 0.032$, pairwise Pearson's χ^2-test with Holm-Bonferroni adjustment). The amount of time spent at the sports club per week is adjusted to account for seasonal fluctuations and truncated at 600 minutes per week to account for outliers and plausibility concerns.[11] Individuals in the Treatment group spend on average approximately as much time at the sports club as the reference group from the MoMo study with 136.17 minutes per week (Schmidt *et al.*, 2017). Again, the null hypothesis of equality between the three groups has to be rejected ($p = 0.0072$, Kruskal-Wallis H-test) due to individuals of the ESS group who spend on average substantially more time at the sports club than individuals in the Treatment and Control group (ESS vs. Treatment: $p = 0.0027$, ESS vs. Control: $p = 0.0208$, pairwise Dunn test with Holm-Bonferroni adjustment). Total physical activity reflects the average number of days an individual is physically active for at least 60 minutes irrespective of the type of activity (WHO, 2010) in both, the last week and in an average week (Schmidt *et al.*, 2016). For the 2003–2006 wave of the MoMo study Bös *et al.* (2009) report a mean in the cohort of individuals aged 11 years of 3.9 days for boys and 3.4 days for girls, respectively. The corresponding numbers for individuals in the present study are higher as reported in Table 3.2. In line with the previous observations, we find individuals in the ESS group to be physically more active than individuals in the Treatment and Control group (equality of groups: $p = 0.0562$, Kruskal-Wallis H-test; ESS vs. Treatment: $p = 0.0294$, ESS vs. Control: $p = 0.0432$, pairwise Dunn test with Holm-Bonferroni adjustment). Competitiveness is measured by a binary choice between a piece-rate and a winner-takes-all tournament as compensation schemes in the encryption task (see Section 3.4.1). Whereas the shares of individuals who choose the tournament scheme and are, hence, deemed competitive does not differ between the Treatment and Control group ($p = 0.520$) and the Treatment and EES group ($p = 0.104$), respectively, the share of competitive individuals in the

	Female	Male	p-value*
Control	20%	62%	0.000
Treatment	42%	60%	0.041
ESS	58%	90%	0.002

* Pearson's χ^2-test.

Table 3.3: Share of competitive individuals by gender and group.

ESS group is statistically significantly higher than in the Control group ($p = 0.078$, pairwise Pearson's χ^2-test with Holm-Bonferroni adjustment). The difference in competitiveness between female and male individuals (see Table 3.3) within the Treatment and Control group is by and large in line with the results observed by Almås *et al.* (2016) for individuals aged 14–15 years in Norway. A difference occurs also in the ESS group, albeit on a higher level which supports the notion that competitiveness is also influenced by the environment (Booth and Nolen, 2012).

3.5.2 First Measurement: Cross-Section

We start by analyzing contributions to the collective account in the public good game in the first measurement to study the relation between sports participation and the willingness to cooperate prior to the intervention. As mentioned in Section 3.5.1, the intensity of sports participation of individuals in the ESS group exceeds the intensity of sports participation of individuals in the Treatment and Control group, respectively. It is, therefore, to consider whether more intensive sports participation also yields different outcomes in cooperative behavior. Figure 3.2 depicts mean contributions to the collective account in the public good game by groups. Contributions are not equal between the three groups ($p = 0.0038$, Kruskal-Wallis H-test); in particular, contributions of the ESS group are statistically significantly different from those in both, the Treatment and Control group (ESS vs. Treatment: $p = 0.0024$, ESS vs. Control: $p = 0.0030$, pairwise Dunn test with Holm-Bonferroni adjustment). Yet, it has to be noted that the design of the current study does not allow to derive causal conclusions on the relation between the intensity of sports participation and cooperative behavior due to potential selection effects, but correlations.

The results reported in Table 3.4 confirm the difference between the ESS and Control group, whereas cooperative behavior does not differ between the Treatment and Control group. Along with the findings reported in column (2) for the control variables discussed in Section 3.5.1, this result indicates that the randomization of individuals in the Treatment and Control group was successful. The difference in contributions in the ESS group remains statistically significant also when controlling for other potential influences related to sports participation, such as the intensity of undifferentiated physical activity and sports participation, membership in sports clubs, and competitiveness. Regarding the first part of our research question, we find that neither membership in a sports club nor the level of undifferentiated physical activity or the time spent at the sports club per week with adjustments for seasonal fluctuations can

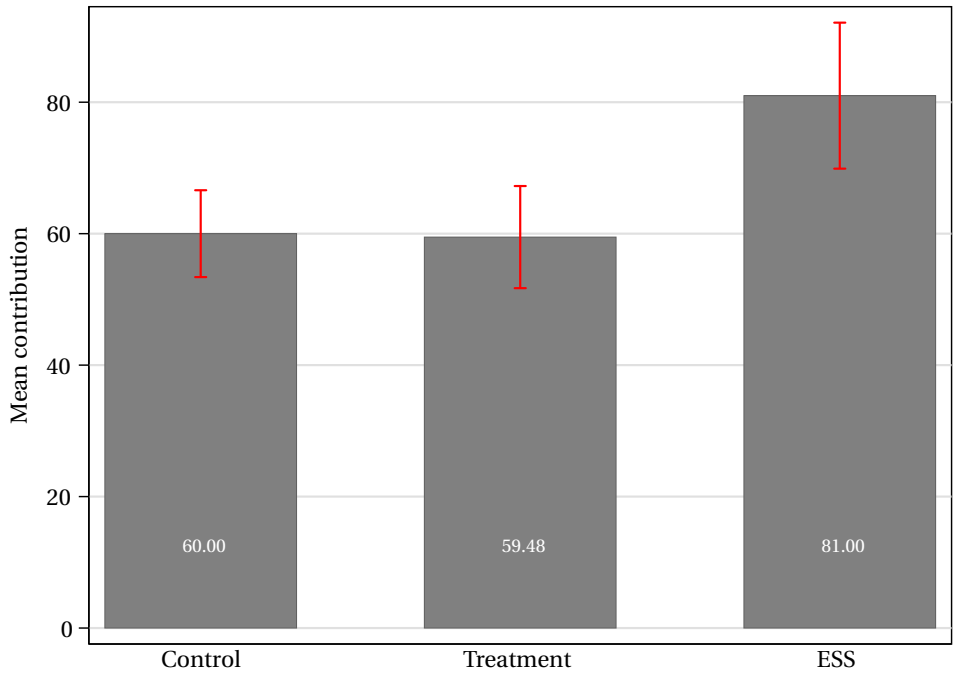


Figure 3.2: Mean contributions to the collective account in the first measurement by groups. Error bars indicate the 95% confidence intervals.

account for the different levels of contributions to the collective account of the ESS group. In particular, the coefficient for the time spent at the sports club in column (4) is very low in magnitude and, hence, not meaningful in economic terms. Although competitiveness is also more pronounced for individuals in the ESS group as compared to the Treatment and Control group, the variable cannot account for the differences in contributions to the collective account in the regression either. Another salient difference in the ESS group refers to the share of individuals from households with low socio-economic status. Yet, controlling for socio-economic status does not explain the differences in in contributions to the collective account in the public good game either. The robustness check reported in Table B.1 confirms these results as using ordered Logit estimates instead of OLS yields qualitatively similar results regarding significance and sign[12] of the coefficients.

As there is broad evidence on achievement gaps among children by parental income (see, e.g., Caucutt *et al.*, 2017, and the references cited therein), Figure 3.3 depicts contributions to the collective account in the public good game in the first measurement by groups and socio-economic status. Although the corresponding coefficient in the last column of Table 3.4 is statistically insignificant, contributions to the collective account in the public good game reveal heterogeneous patterns by groups and socio-economic status. While contributions of individuals from households with low and high socio-economic status do not differ in the Treatment and Control group (Control: $p = 0.1206$, Treatment: $p = 0.1140$), the difference is statistically significant in the ESS group ($p = 0.0033$, two-sided Mann-Whitney U-test). Furthermore, the direction of these differences varies between the groups. The results of OLS estimates of contributions to the collective account in the public good game reported in Table 3.5 are separated by the socio-

	(1)	(2)	(3)	(4)	(5)
Treatment	–0.522	0.625	0.192	–0.078	–0.831
	(2.775)	(2.744)	(2.558)	(2.565)	(3.464)
ESS	21.000***	20.047***	23.585***	22.836***	19.069***
	(1.352)	(2.601)	(2.517)	(2.567)	(2.570)
Total Activity		–1.471	–1.631		
		(1.705)	(1.654)		
Minutes Club				–0.005	
				(0.015)	
Member		1.511	2.901	2.674	
		(9.299)	(7.587)	(6.883)	
Competitive		–3.024	–2.890	–3.762	
		(4.397)	(4.233)	(4.216)	
Female					2.241
					(3.311)
Low SES					7.170
					(8.371)
Constant	60.000***	66.767***	65.902***	60.013***	57.319***
	(1.352)	(6.887)	(5.391)	(5.935)	(3.166)
Obs.	130	118	130	130	130
No. Groups	9	9	9	9	9
Adj. R^2	0.080	0.035	0.073	0.064	0.081

Note: Missing values of the variable Total Activity in column (3) are replaced with means by measurement, class, gender, and socio-economic status.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.4: Contributions to the collective account in the first measurement. Ordinary least squares estimates.

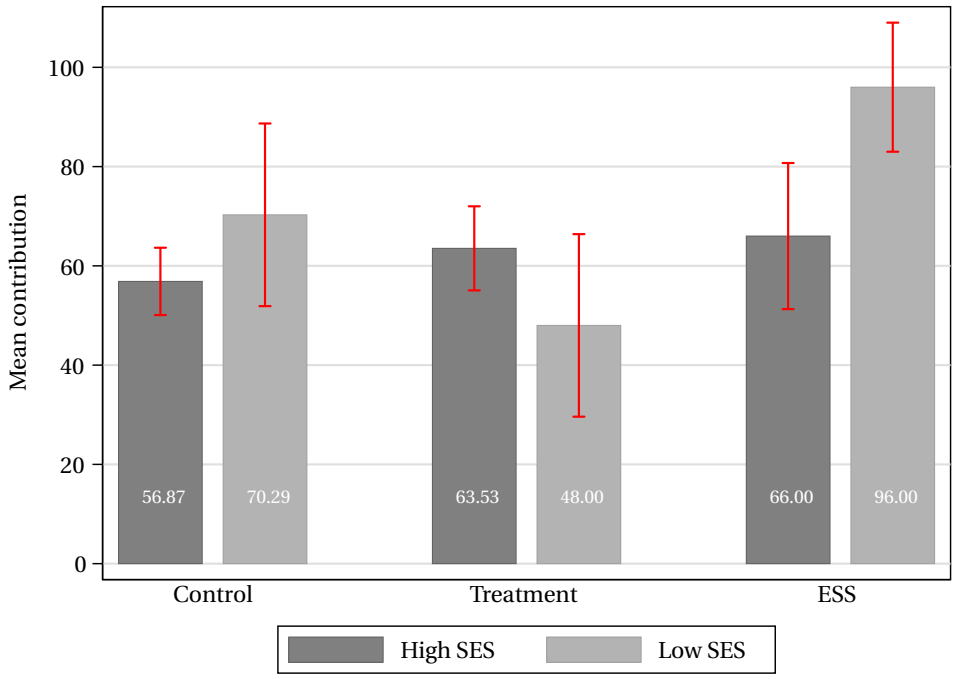


Figure 3.3: Mean contribution to the collective account in the first measurement by group and socio-economic status. Error bars indicate the 95% confidence intervals.

economic status of the households individuals live in. In columns (1)–(3) focusing on individuals from households with high socio-economic status, individuals from the ESS group contribute statistically (highly) significantly more to the collective account than individuals in the Control group, whereas the difference between the Treatment and the Control group is statistically not significant. It must be noted, however, that the adjusted R^2 of these specifications is close to zero and that the results, thus, have to be treated with caution as the models are not constructive for individuals from households with high socio-economic status. Focusing on contributions by individuals from households with low socio-economic status in column (4), both, the Treatment and ESS group, are statistically (highly) significantly different from the Control group. Moreover, the signs of the groups' deviations differ. Whereas individuals in the Treatment group contribute less than their counterparts from the Control group, individuals in the ESS group contribute more. Adding additional control variables related to sports participation in columns (5) and (6) of Table 3.5 does not provide additional insights regarding individual contributions to the collective account in the public good game as the corresponding coefficients are statistically insignificant. Furthermore, the significance of the coefficient for the dummy variable indicating the ESS group remains unchanged with these additional control variables, whereas the difference in contributions between the Treatment and Control group is statistically only marginally significant.

3.5.3 Panel Regressions

For the subsequent analyses of changes in cooperative behavior over the course of the study, we exclude the ESS group and consider the effects of the intervention by comparing the Treatment

	High SES			Low SES		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	6.660	9.104	7.532	–22.286**	–25.206*	–24.618*
	(5.446)	(6.275)	(5.492)	(7.965)	(12.165)	(11.884)
ESS	9.130***	10.494**	10.753**	25.714***	28.123***	32.715***
	(2.623)	(4.312)	(4.287)	(3.984)	(7.577)	(6.857)
Total Activity		–1.822	–2.104		–3.030	–3.090
		(1.476)	(1.405)		(4.222)	(4.211)
Member		–0.640	1.293		1.953	3.453
		(9.281)	(6.868)		(14.518)	(14.686)
Competitive		–1.532	–0.726		–9.804	–10.722
		(5.010)	(5.592)		(11.586)	(10.092)
Constant	56.870***	66.214***	65.457***	70.286***	86.934***	86.370***
	(2.623)	(8.391)	(5.432)	(3.984)	(16.246)	(16.250)
Obs.	92	82	92	38	36	38
No. Groups	9	9	9	9	9	9
Adj. R^2	0.004	-0.009	-0.007	0.301	0.227	0.296

Note: Missing values of the variable Total Activity in columns (3) and (6) are replaced with means by measurement, class, gender, and socio-economic status.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.5: Contributions to the collective account in the first measurement by socio-economic status. Ordinary least squares estimates.

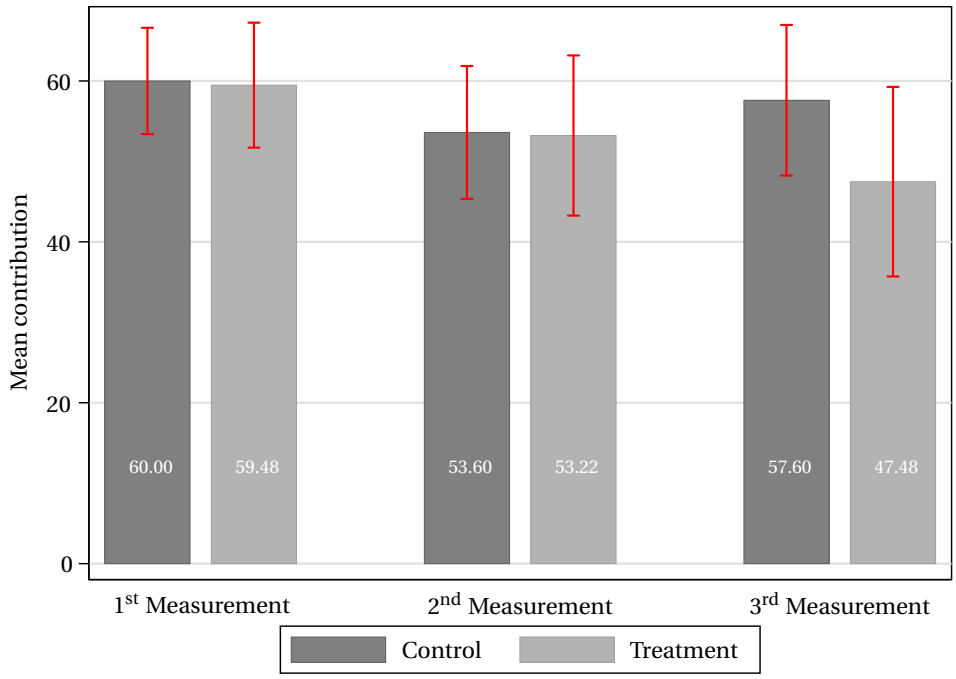


Figure 3.4: Mean contributions to the collective account by group and measurement. Error bars indicate the 95% confidence intervals.

and Control group.

An overview of mean contributions to the collective account in each measurement is provided in Figure 3.4. As discussed in Section 3.5.2, mean contributions to the collective account are statistically not significantly different between both groups in the first measurement ($p = 0.9064$, two-sided Mann-Whitney U-test). This applies also for differences between both groups in the subsequent measurements (second measurement: $p = 0.8538$, third measurement: $p = 0.1712$, two-sided Mann-Whitney U-test). Over the course of the study, mean contributions to the collective account decline slightly in the Treatment group, while the differences from measurement to measurement are statistically not significant (first vs. second measurement: $p = 0.1940$, second vs. third measurement: $p = 0.6139$, two-sided Wilcoxon signed-rank test). In the Control group a slight decline occurs from the first to the second measurement ($p = 0.1387$, two-sided Wilcoxon signed-rank test), whereas mean contributions to the collective account increase in the third measurement (second vs. third measurement: $p = 0.3280$, two-sided Wilcoxon signed-rank test) which is an untypical pattern compared to the vast majority of other studies involving repeated public good games.

As Figure 3.3 already pointed at different contribution behavior in the first measurement due to both, group and socio-economic status, the corresponding overview by socio-economic status and measurement in the Treatment and Control group is depicted in Figure 3.5. The patterns are again heterogeneous between the respective sub-groups over the course of the study. Hence, results of random-effects regressions are reported to examine the second part of the research question regarding the effectiveness of the intervention. Starting with the first spec-

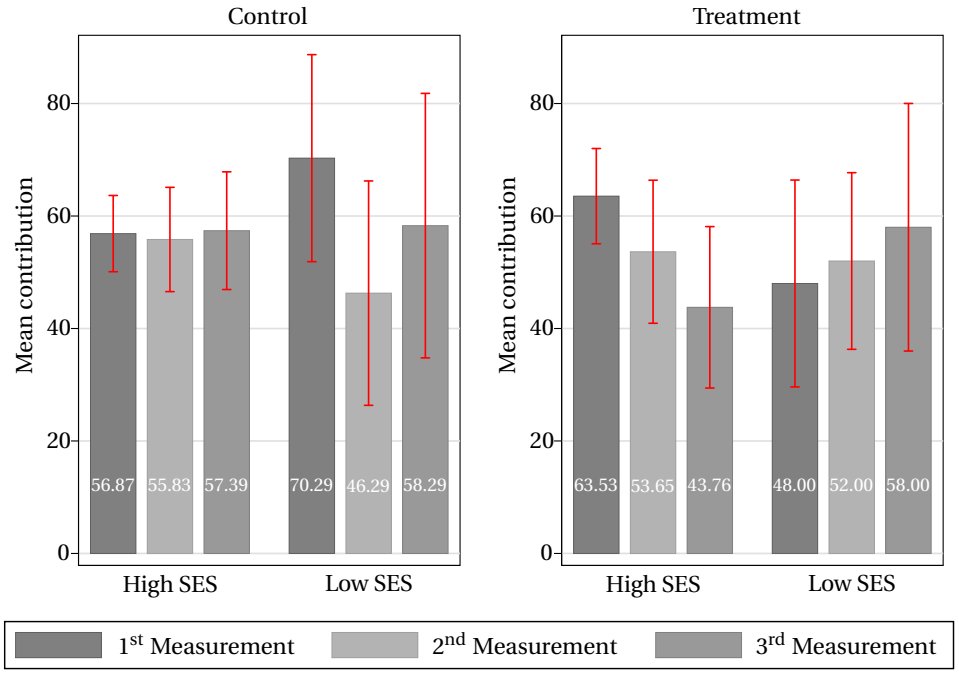


Figure 3.5: Mean contributions to the collective account by group, measurement, and socio-economic status. Error bars indicate the 95% confidence intervals.

ification of Table 3.6, the main effect on differences between the Treatment and Control group as well as differences of the second and third measurement relative to the first measurement are statistically insignificant and so are the coefficients of the corresponding interaction terms. Controlling for socio-economic status and simultaneously adding three-way interaction terms to account for the interdependence between the intervention in the Treatment group, the time of measurement, and individuals' socio-economic status reveals the influence of socio-economic status on different patterns of contribution behavior. The main effect addressing differences between the Treatment and Control group which is already included in the previous specification remains statistically insignificant, whereas the coefficient of the newly added dummy variable for individuals from households with low socio-economic status is positive and statistically highly significant which indicates differences between contributions of individuals from households with high and low socio-economic status in the Control group during the first measurement. Contrary to this observation, the direction of the difference in contributions to the collective account is reversed when considering individuals from households with low socio-economic status in the Treatment group during the first measurement. In both, the second and third measurement, individuals from households with low socio-economic status in the Control group reduce their contributions as denoted by negative and statistically significant coefficients of the corresponding interaction terms. Looking at the impact of the intervention on contributions to the collective account by individuals from households with low socio-economic status in the Treatment group during the second and third measurement, the positive and statistically significant coefficients fortify the effect of the intervention applying an enhanced concept of physical education on cooperative behavior for this particular sub-group in both, the short and the long

run. Controlling additionally for gender and factors related to sports participation only reveals a gap in cooperative behavior by gender due to statistically significantly higher contributions by female individuals. Robustness checks for these results are reported in Table B.2; estimates of ordered Logit models with random effects yield qualitatively similar results.

Focusing on individuals from households with low socio-economic status in the Treatment and Control group in Table 3.7 allows for additional insights into changes in cooperative behavior of this particular sub-group over the course of this study. Even before the intervention in the Treatment group, contributions to the collective account differed between both groups. While contributions in the Control group in the second and third measurement are statistically significantly lower than in the first measurement, the coefficient of the interaction effect between the Treatment group and the second measurement is positive and statistically significant. Although the overall effects is still negative, the positive coefficients of the interaction term provide additional support to the second part of the research question regarding the effectiveness of the intervention as the enhanced physical education lessons seem to counterbalance the overall decline in contributions to the collective account to some extent. Moreover, when controlling for additional variables as in the previous regressions, the effect of the intervention on individuals from households with low socio-economic status is positive and statistically marginally significant which suggests that the change in cooperative behavior is present even some time after the end of the intervention.

To summarize the results on the impact of the intervention on cooperative behavior, contributions to the collective account in general do not differ between the Treatment and Control group in a specific measurement. However, results differ depending on the socio-economic status of households individuals live in. Whereas individuals from households with high socio-economic status do not increase their contributions to the collective account over time in both, the Treatment and Control group, individuals from households with low socio-economic status in the Treatment group become more cooperative over time which translates into higher contributions to the collective account in the public good game. As reported in Table 3.6, the coefficients of both three-way interaction terms between the dummy variables for the Treatment group, the time of measurement, and individuals from households with low socio-economic status are statistically significant which emphasizes the effectiveness of our intervention using enhanced physical education lessons to foster cooperative behavior for this particular sub-group both, directly after the end of the intervention and also nine weeks later in the final measurement.

3.6 Conclusion

The results of this study indicate the effectiveness of our intervention using a novel concept employing structured physical education lessons on cooperative behavior outside the gymnasium.

Comparisons of individuals attending an elite school of sports who integrate a high intensity of sports participation in their everyday lives and individuals attending other schools at which physical education and sports participation are less pronounced in the curriculum support the

	(1)	(2)	(3)	(4)
Treatment	−0.522	6.660	9.208	8.039
	(2.797)	(5.525)	(7.437)	(6.828)
2nd Measurement	−6.400	−1.043	−0.506	−0.689
	(3.961)	(2.890)	(3.259)	(3.353)
3rd Measurement	−2.400	0.522	−1.372	1.082
	(5.165)	(6.087)	(6.542)	(6.146)
Treatment × 2nd Measurement	0.139	−8.839	−11.886**	−10.855*
	(7.460)	(6.903)	(5.060)	(6.076)
Treatment × 3rd Measurement	−9.600	−20.286*	−21.958*	−21.971*
	(11.778)	(12.111)	(12.464)	(11.510)
Low SES		13.416**	14.224**	14.164**
		(5.350)	(6.159)	(6.272)
Treatment × Low SES		−28.946**	−34.969**	−32.176**
		(12.780)	(14.511)	(14.232)
2nd Measurement × Low SES		−22.957**	−21.359*	−21.530*
		(10.609)	(11.783)	(12.363)
3rd Measurement × Low SES		−12.522**	−9.394	−11.093*
		(5.541)	(6.845)	(6.382)
Treatment × 2nd Measurement × Low SES		36.839**	37.781**	37.348**
		(15.216)	(16.532)	(17.064)
Treatment × 3rd Measurement × Low SES		42.286***	43.100***	42.476***
		(14.795)	(15.741)	(14.547)
Female			9.796***	9.214***
			(3.327)	(2.688)
Total Activity			−1.728	−2.116**
			(1.248)	(1.021)
Member			0.866	4.401
			(7.365)	(6.746)
Competitive			2.522	1.845
			(3.781)	(3.772)
Constant	60.000***	56.870***	58.213***	57.506***
	(1.363)	(2.661)	(6.426)	(5.398)
Obs.	318	318	296	318
No. Individuals	106	106	106	106
No. Groups	8	8	8	8
Overall R^2	0.017	0.037	0.076	0.068

Note: Missing values of the variable Total Activity in column (4) are replaced with means by measurement, class, gender, and socio-economic status.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.6: Contributions to the collective account in all three measurements. Generalized least squares estimates with random effects.

	(1)	(2)	(3)
Treatment	–22.286***	–34.680***	–33.557***
	(8.075)	(10.092)	(9.194)
2nd Measurement	–24.000**	–23.312**	–22.864**
	(10.381)	(11.020)	(10.468)
3rd Measurement	–12.000**	–12.498*	–11.586*
	(4.832)	(6.899)	(6.313)
Treatment × 2nd Measurement	28.000**	32.249**	32.889**
	(14.203)	(15.946)	(14.322)
Treatment × 3rd Measurement	22.000*	26.054**	25.720**
	(12.294)	(11.135)	(10.344)
Female		16.998***	18.888***
		(5.475)	(5.900)
Total Activity		–1.276	–1.641
		(3.864)	(3.586)
Member		–9.228	–4.490
		(12.609)	(11.115)
Competitive		–7.960	–10.937
		(6.521)	(7.286)
Constant	70.286***	81.756***	80.126***
	(4.039)	(13.405)	(13.453)
Obs.	78	75	78
No. Individuals	26	26	26
No. Groups	8	8	8
Overall R^2	0.062	0.188	0.204

Note: Missing values of the variable Total Activity in column (3) are replaced with means by measurement, class, gender, and socio-economic status.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3.7: Contributions to the collective account by individuals from households with low socio-economic status in all three measurements. Generalized least squares estimates with random effects.

notion that individuals in the former group who report on average substantially higher levels of sports participation are indeed associated with higher levels of cooperative behavior. This result is in contrast to Celse *et al.* (2017) who find that university students of sports sciences which requires a considerable level of sports participation as part of the curriculum are more inclined to conditional cooperation than their peers, albeit not generally more cooperative. Our results from the first measurement, however, do not allow to derive causal conclusions about the effects of sports participation on the formation of social skills since we do not explicitly address the issue of potential selection effects. Moreover, the results of the corresponding non-parametric and parametric statistical analyses indicate only a significant difference between individuals in the ESS group and individuals at other schools, whereas different levels of sports participation cannot account for different levels of contributions to the collective account in the public good game.

Contributions in the public good game of individuals of the Treatment and Control group in the first measurement do not substantially differ which implies that the randomization was successful. This applies to the largest extent also for the control variables used in the regressions. At first glance, there are hardly any changes in average contributions to the collective account in the public good game between individuals in the Treatment and Control group over the different measurements. Taking a closer look and considering contributions of individuals from households with low and high socio-economic status separately, however, provides more diverse results. More specifically, individuals from households with low socio-economic status in the Treatment group who are initially the least cooperative sub-group respond positively to the intervention and increase their contributions in the public good game slightly over all three measurements and eventually catch up with their classmates whose contributions decline over time as commonly observed in repeated public good games. Results of a regression comparing only individuals from households with low socio-economic status in the Treatment and Control group provide additional weak support for the effectiveness of the intervention on cooperative behavior.

Overall, the results presented above indicate the effectiveness of our intervention as part of regular physical education lessons. While it is generally not straightforward to evaluate benefits of sports participation due to concerns about selection effects, our experimental design with a randomized controlled trial avoids this issue. Moreover, the lack of significance for explanatory variables related to sports participation indicate that the results of the intervention are not biased due to prior sports participation. The rather weak effects can be explained by the fact that individuals are already quite cooperative at the beginning of the study as average contributions in the Treatment and Control group amount to half of the endowment which is in line with common findings reported in the literature (e.g., Ledyard, 1995; Chaudhuri, 2011). Hence, the scope for improvements in cooperative behavior is limited. Furthermore, duration and intensity of the intervention are moderate. In contrast to, e.g., Kosse *et al.* (2018) whose intervention lasts one year, the time span of six weeks in this study is quite short. Additionally, individuals had mandatory physical education classes before; the intervention was only a slight adjustment

whose main contribution are two blocks of ten minutes each at the beginning and end of each lesson in which individuals were encouraged to share and discuss their experiences about the topic covered in a specific lecture. The nature of cooperation as conveyed in the intervention and in the public good game is related but not perfectly congruent. In a comparison of individuals' cooperative behavior inside and outside the laboratory, Englmaier and Gebhardt (2016) observe complementary behavioral patterns in situations which provide similar incentives for cooperation using a within-subject design. Galizzi and Navarro-Martinez (2018), in contrast, cannot confirm this finding as situations inside and outside the laboratory are quite different in their setting. In line with this, the magnitude of the effects as measured by contributions to the collective account in the public good game in the present study are quite moderate. Finally, the share of individuals from households with low socio-economic status ranges around one quarter in the Treatment and Control group which consequently results in a low number of observation. Results in this regard should, therefore, be treated with caution.

Despite some limitations, we find pieces of evidence for the effectiveness of an intervention as part of regular physical education lessons to foster cooperative behavior which is an integral determinant for success in various domains of life, including individual labor market outcomes. As physical education is mandatory at schools in Germany, slight adjustments of the existing curriculum can be used to foster cooperation even with a limited input of additional resources.

Notes

1. In the literature, soft skills and non-cognitive skills are frequently used as synonyms; either term comprises social competencies as a subset (Heckman and Kautz, 2012). Social competencies are concerned with aspects which foster successful interaction with other individuals such as leadership skills and social skills. The latter category describes a set of acquired capabilities in the interpersonal domain, including cooperation, interpersonal skills, empathy, assertion, and responsibility (Farrington *et al.*, 2012; Gutman and Schoon, 2013).

2. Although Enste *et al.* (2018) describe a linear public good game, they use in fact a binary threshold public good game which is frequently used to study coordination instead of cooperation (see Section 1 and Section 2).

3. It is important to note that cognitive and non-cognitive skills are no strictly separated domains; in particular, outcomes can only very rarely be attributed to exactly one of these realms (Borghans *et al.*, 2008; Farrington *et al.*, 2012).

4. Due to fairness concerns, we offered teachers of the classes in the Control group the opportunity to conduct the intervention in their classes after the end of the study.

5. Team sports: football or basketball; individual sports: jogging, swimming, golf, or tennis.

6. This does not necessarily imply physical activity. Felfe *et al.* (2016) note that the acquisition of social skills in a sports club may not be achieved by physical activity in the first place.

Rather, the fact of being involved in an organization is assumed to foster social skills (see also German Bundestag, 2017).

7. Similar to Becker *et al.* (2012), we find no cogent relation between decisions in the economic experiments and the psychometric scales. Thus, these data are not included in the analysis.

8. Note that the intervention is not designed to replicate the properties of a public good game. In contrast, Fan (2000) uses an approach which explicitly informs individuals about the consequences of different actions in a prisoner's dilemma for individuals themselves, the other individual, and the group. Positive effects of this kind of intervention are only found within a few minutes after the provision of this information.

9. Unlike their parents, individuals also answered the questionnaire after the second and third measurement. For sake of clarity, however, only answers from the first measurement are used in this overview.

10. In line with Kosse *et al.* (2018) the socio-economic status of a household is deemed low if one or more of the following criteria apply: *(i)* low education: both parents left school without a university entrance certificate; *(ii)* low income: equivalence income below 1,033 Euro, i.e., 60% of the median net equivalent income in Germany in 2015 (Destatis, 2017); *(iii)* single parent.

11. Non-parametric and parametric results are robust with respect to this truncation.

12. Note that the coefficients reported in Table B.1 are exponentiated.

4 Guilt and Shame[‡]

This paper studies the impact of private and social image concerns on sabotage decisions in a tournament. More specifically, we use different layers of information disclosure to either individuals themselves or their opponent to study the effect of visibility of the implication of one's sabotage activities on individuals' private and social image, respectively. Furthermore, we provide individuals with the opportunity to pay a price to prevent information disclosure, i.e., individuals can prevent that—depending on the treatment—they or their opponent are informed about the implication of the individual's sabotage activities. The results show that the decision to inflict sabotage on the opponent *per se* is not affected by different layers of information disclosure. Having the opportunity to prevent information disclosure, however, individuals in the treatment which makes the implication of their sabotage activities visible for the opponent and, hence, addresses social image concerns exhibit a significantly higher willingness to pay.

4.1 Introduction

Rank-order tournaments are commonly used in companies, e.g., when trying to find the suitable candidate for a promotion (Prendergast, 1999; Lazear and Shaw, 2007). One major advantage of this approach is that it only requires an ordinal ranking of candidates based on their relative performance. Moreover, Lazear and Rosen (1981) show in their seminal contribution that the first-best solution of a principal-agent setting with verifiable effort can be attained in a tournament even under asymmetric information. Besides increasing one's probability to win the tournament by exerting more effort, decreasing the opponent's output by means of sabotage yields the same effect (Lazear, 1989).[1] Ample empirical evidence—primarily from studies conducted in the laboratory—suggests that sabotage is indeed a serious issue when only relative performance is evaluated (e.g., Harbring and Irlenbusch, 2004, 2005, 2008, 2011; Harbring *et al.*, 2007; Carpenter *et al.*, 2010; Vandegrift and Yavas, 2010; Balafoutas *et al.*, 2012; Dato and Nieken, 2014; see also Chowdhury and Gürtler, 2015, for an overview).

The present paper is to study how an individual's decision to inflict sabotage on her opponent is influenced by potential impairment of her private and social image. Even if the opponent is not aware that the individual inflicted sabotage on him, the individual herself may find it difficult to maintain a positive private image if sabotage is generally considered illegitimate. Furthermore, the individual's social image can suffer if the opponent is informed that he has been subject of sabotage by the individual. There is evidence that individuals do not only care about material outcomes but also have concerns for both, their private and social image. The former refers to the notion that an individual attempts to act upon her own standards which allows to think of herself as a good person (Mazar *et al.*, 2008; Matthey and Regner, 2011). On the other hand, actions taken by the individual convey a signal of her attitudes which induces the individual to present herself as, e.g., prosocial in order to be perceived pleasantly by others

[‡]This section is based on joint work with Sandra Ludwig and Petra Nieken (Haas *et al.*, 2019a).

(Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009; Ariely *et al.*, 2009). The corresponding terms used in the psychological literature are guilt and shame. While the former occurs as a result of inappropriate behavior towards someone else even if the sufferer is not aware of any harm inflicted on him, the latter involves disapproval by the affected individual (Savikhin Samek and Sheremeta, 2014; see also de Hooge *et al.*, 2007).

In certain situations, the desire for a positive social image is at odds with an individual's actual preferences. Experimental studies on dictator games suggest that a substantial fraction of individuals who share their endowment with another individual in an anonymous laboratory setting prefer not to be put in a situation in which they have to decide on the allocation of the endowment between themselves and someone else. This observation indicates that the individual may suffer from disutility even under anonymity due to fears to violate expectations of another individual in a setting with common knowledge about the action space of the individual (Geanakoplos *et al.*, 1989; Charness and Dufwenberg, 2006; Battigalli and Dufwenberg, 2009). Hence, when offered the opportunity to opt out of a situation in which they can share an endowment between themselves and another individual without letting the potential recipient know about the fact that he could have received money as part of a game, a substantial fraction of those reluctant sharers indeed chooses the alternative to opt out (Dana *et al.*, 2006; Lazear *et al.*, 2012; see also DellaVigna *et al.*, 2012, for a related field study). Moreover, individuals in the role of dictators frequently engage in strategic ignorance by exploiting a "moral wiggle room" to ensure a high payoff for themselves when their accountability for adverse results of the recipient is not unambiguous (Dana *et al.*, 2007; Larson and Capra, 2009; Golman *et al.*, 2017).

To study the impact of private and social image concerns on sabotage when the individual's accountability is not obvious we run an experiment employing a tournament with a real-effort task. Unlike other studies on tournaments in which the individual with the higher output is directly awarded the winner prize, in the present study the spread of both individuals' outputs translates into winning probabilities which determine which player receives the winner and loser prize, respectively. Sabotage does not reduce the opponent's winning probability for sure which provides a moral wiggle room for the individual. Additionally, we offer individuals the opportunity to prevent the disclosure of information about the effectiveness of their sabotage either to themselves or to the other individual by paying a price.

4.2 Related Literature

The focus of the present study is on an individual's decision to improve her relative position in a real-effort tournament by inflicting sabotage on her opponent if information about this decision is or may be disclosed to the individual and—depending on the treatment—to the opponent. Hence, the decision to sabotage can affect not only the individual's private image, but also her social image.

In the domain of dishonest behavior, Mazar *et al.* (2008) observe that cheating is prevalent in their experiments—but predominantly on levels far below the maximum possible magnitude.

This does not change when monetary benefits from cheating increase. They corroborate their findings with the notion that the cost-benefit analysis of whether or not to behave honestly is not restricted to monetary terms. Moreover, individuals have a conception on patterns of behavior which are desirable and violations thereof decrease the individual's overall utility. Yet, low levels of dishonest behavior observed by Mazar *et al.* (2008) suggest that there is a certain range of deviation from the desirable behavior which individuals deem tolerable. This can be either explained by individuals who disguise situations in which they exhibit deviating behavior with a euphemistic reinterpretation or temporarily disregard their standards on desirable behavior.

To study the influence of private image concerns on giving in a dictator game, Dana *et al.* (2007) provide individuals in their treatments with different opportunities to remain uninformed about the payoff of the recipient. In one situation, the dictator can choose between two alternatives in a binary version of the dictator game. While she knows which alternative yields a higher payoff for herself, the individual is uninformed whether the recipient's payoff is aligned with hers or not. Although dictators can acquire this information without bearing any monetary costs, only slightly more than half of the dictators use this opportunity while the others remain uninformed which is one explanation for an increase in selfish behavior compared to the baseline treatment in which dictators know the implication of their choice for the recipient's payoff. Moreover, similar observations are obtained by Dana *et al.* (2007) in a treatment with multiple dictators in which the selfish alternative is only implemented if both dictators independently choose this option or in another treatment in which dictators can defer their decision until eventually a computer intervenes and chooses one of the alternatives with equal probability. To summarize, the results show that dictators deliberately exploit situations which provide a moral wiggle room, i.e., circumstances which do not allow to unambiguously map the decision of the dictator with the implication for the recipient's payoff as an excuse to themselves for selfish behavior to maintain a positive private image. A robustness check by Larson and Capra (2009) using a slight modification of the experimental design confirms this result and assures that the results are not driven by an omission bias which makes dictators perceive harm for the recipient caused by inaction less objectionable than actively inflicting harm on the recipient. While Grossman (2014) also finds evidence for dictators who prefer to stay uninformed and engage in selfish behavior, dictators in this study must actively choose to stay uninformed which decreases the share of dictators deciding to stay uninformed compared to the previous studies.

In a related study, Matthey and Regner (2011) use a within-subject design to distinguish dictators who are intrinsically motivated to share their endowment with the respective recipients from those dictators who share reluctantly to avoid adverse sentiments if they are confronted with the implication of their allocation for the recipient's payoff. Based on the work of Konow (2000), they test a model of cognitive dissonance which postulates psychological costs in case of an individual who finds herself confronted with a choice between two diametrically opposed alternatives, namely to comply with a norm which advocates sharing and their actual intention to keep the entire endowment. To resolve this problem, dictators who share reluctantly in a situation with full information prefer to remain ignorant about the implication of their choice if

possible. The results obtained by Matthey and Regner (2011) are in line with those of Dana *et al.* (2007). Furthermore, they provide evidence that dictators who are not intrinsically motivated to share their endowment with a recipient still do so in a situation with full information about the implication of their decision on the recipient to avoid disutility incurred by cognitive dissonance.

A different approach to study the impact of strategic ignorance in situations which contrast individual benefits and negative externalities is the experiment by Kajackaite (2015). Individuals can earn money by working on a real-effort task which yields a piece-rate for each correctly solved task. In some treatments, however, individuals' effort also generates a donation for the National Rifle Association (NRA) which represents a predominantly negatively perceived organization. Both, individuals who are not informed whether or not the NRA benefits from their effort and individuals who explicitly decide to stay ignorant about this fact, exert more effort than individuals who know that they are generating a donation for the NRA. While this result is by and large in line with findings of other studies on strategic ignorance cited above, a closer look reveals that the results in the treatment in which individuals can choose to remain ignorant is mainly driven by a sorting effect when individuals have the choice whether they want to be informed about the implication of their action for a third party or stay ignorant instead.

The studies outlined so far focus on the impact of selfish behavior on individuals' private image. In contrast to that, concerns for their social image can also influence individuals' decisions to engage in behavior which may affect others negatively. More precisely, the decision not to behave selfishly is likely to foster a benevolent perception of an individual by others. In the study by Ariely *et al.* (2009), individuals work on a tedious real-effort task which consists of pressing X and Z keys alternating on a keyboard. At the same time, a donation for a charity is generated which depends on the individual number of X-Z key strokes. Ariely *et al.* (2009) use a $2 \times 2 \times 2$ between-subject design in which the charity is either the Red Cross or the NRA, individuals work without or with private incentives in form of a piece-rate, and their performance is disclosed to an audience or not. When donations accrue to the Red Cross, individuals exert more effort when their result is visible for an audience to signal their prosocial attitude. If the recipient of the donation is the NRA, individuals still exert positive levels of effort but on a substantially lower magnitude compared to the case generating donations for the Red Cross. Monetary incentives for individuals induce higher levels of effort only in the private settings as the signal of high effort levels is perceived differently by the audience as they cannot disentangle the actual motivation behind individuals' effort. Therefore, exerting high levels of effort can have detrimental effects on individuals' social image in a public setting and individual monetary rewards which provides an explanation for the insensitivity to incentives when the result is disclosed to an audience.

To test the relevance of social image concerns for giving in dictator games, Dana *et al.* (2006) offer dictators the opportunity to quit the game. If the dictator chooses this option, the potential recipient is not informed about the game which was supposed to be played. The quitting dictator receives a fixed payment which is slightly below the endowment of the related dictator game. The rationale behind this set-up is that a dictator who is maximizing her monetary payoff decides to

play the dictator game—and keeps the entire endowment for herself. If the dictator is, however, concerned about her social image, violating possible expectations of the recipient by not sharing the endowment induces disutility which outweighs monetary gains from keeping the entire endowment compared to the fixed payment when quitting. Hence, giving in dictator games must not be confused with generosity. In particular, dictators who are concerned about their social image share their endowment with a recipient although they have no intrinsic motivation to do so. Merely the fact that they find themselves in a situation which imposes a risk of damaging their social image makes them imitate generous behavior. Therefore, a decent share of dictators in the experiment of Dana *et al.* (2006) forgo a small share of the maximum possible payoff to elude the dictator game without letting the potential recipient know in order to avoid an impairment of their social image.

A more detailed overview on a dictator's choice whether to remain in a situation which facilitates sharing the endowment with a recipient or not is provided by Lazear *et al.* (2012). Complementary to the approach by Dana *et al.* (2006), they distinguish individuals who have an intrinsic motivation to share from those who prefer not being asked to share but nevertheless share for the sake of maintaining a positive social image. When the payment for the outside option equals the endowment in the dictator game, the results are in line with Dana *et al.* (2006), i.e., a substantial fraction of dictators decides to quit the game. Additional support for this finding is provided by a related field experiment of DellaVigna *et al.* (2012). In a door-to-door fundraising campaign, the percentage of households which open the door decreases if the time when the solicitors come over is announced in advance. Similar to the laboratory experiments, individuals who prefer not to donate but do not refute a request when being asked by the solicitor face-to-face to maintain a positive social image can elude this situation by not opening the door. In another laboratory experiment, Klinowski (2018) observes that female individuals in a dictator game with an endowment of 10 tokens share a higher fraction than their male counterparts. Yet, when given the opportunity to retract their previous choice by assigning a higher probability on an outcome which yields 9 tokens for themselves and 0 tokens for the recipient, female individuals choose this alternative to a larger extent than males which suggests that female dictators only appear more generous, although they may not generally be more intrinsically motivated to share than male dictators.

To shed light on the interaction of both, private and social image concerns, in an individual's decision to share her endowment with a recipient in a dictator game, Cappelen *et al.* (2017) use an experimental design which systematically varies both dimensions. To address the individuals' private image concerns, there are two treatment manipulations which either emphasize the entitlement or neediness of the recipient. Orthogonal to this variation, a distinction whether the potential recipient is informed about the procedure of the dictator game or not covers the second domain focusing on social image concerns. Compared to a baseline condition in which no information about the recipient is provided, both manipulations which address the individuals' private image concerns result in higher shares of the endowment sent to the recipients. Regarding social image concerns, Cappelen *et al.* (2017) observe that the disclosure

of information to the recipients *per se* does not increase dictators' generosity. However, when aligned with information about entitlement or neediness of the recipient affecting private image concerns in the first place, disclosure of information to the recipients additionally increases the share of the endowment sent to the recipients. The authors explain this finding by a crowding-in effect evoked by a convincing argument for sharing. In a slightly modified version of this experiment, individuals have the opportunity to switch to the opposite information policy after they have made their decision which amount to send to the recipient. In general terms, individuals who send a small share of their endowment to the recipient prefer a situation in which no information is disclosed, whereas generous dictators do not object to visibility of their decision. In line with evidence from the studies cited above, this supports the notion that social image concerns are indeed a crucial determinant for giving in dictator games.

To promote truth-telling in a cheap talk sender-receiver game, Greenberg *et al.* (2015) introduce *ex-post* disclosure in one of their treatments. The sender sees the result of a die roll on his screen and is then asked to communicate this result to a receiver. Payoffs depend on the sum of the pips. If the receiver states the correct number, both players receive the same payoff, whereas an incorrect number yields a higher payoff for the sender at the expense of the receiver. Although the setting of the experiment is anonymous and concerns for reputation can be ruled out due to its one-shot nature, *ex-post* disclosure of the actual outcome of the die roll which is presented to the sender before passing a message to the receiver is an appropriate device to prevent the sender from exploiting her edge on information to attain a monetary gain by deceiving the receiver. In particular, if being exposed as an impostor even at the very end of the experiment after all decisions have been made and payoffs have been determined to an unknown other participant who is not aware of the sender's identity causes negative sentiments with the sender, the institution of *ex-post* disclosure can be expected to foster truth-telling compared to a situation which lacks this institution. Indeed, Greenberg *et al.* (2015) observe a substantial increase in the share of senders who communicate the actual outcome of the die roll in the treatment using *ex-post* disclosure. Notably, this effect is more pronounced for male senders than for females.

The study most closely related to the present paper is by Harbring and Wilhelm (2016). Individuals work on a real-effort task and the individual with the higher output receives a winner prize. To enhance their output relative to the opponent, individuals can either exert more effort in the real-effort task or inflict sabotage on the opponent. Besides a differentiation between active and passive sabotage in tournament by means of commission and omission, Harbring and Wilhelm (2016) also vary whether sabotage is detected and consequentially disclosed to the sufferer or not. More specifically, one of their treatment variations reveals sabotage with a probability of 70% and the information about the amount of sabotage inflicted on the other individual is supplemented by a photo of the delinquent. As a partial unveiling of anonymity by showing the respective individual's photo erodes the social image in case of detection, sabotage occurs less frequently under a positive detection probability irrespective of the type of sabotage. Yet, results of a questionnaire presented to delinquents immediately after their sabotage activity

	INFO SELF$_{100}$ OPP$_0$	INFO SELF$_{70}$ OPP$_0$	INFO SELF$_{100}$ OPP$_{70}$
Period 1	individual always learns whether sabotage inflicted on opponent was effective	individual learns with probability of 70% whether sabotage inflicted on opponent was effective	individual always learns whether sabotage inflicted on opponent was effective (same as in INFO SELF$_{100}$ OPP$_0$)
	opponent never learns whether individual inflicted sabotage on him nor whether individual's sabotage was effective	opponent never learns whether individual inflicted sabotage on him nor whether individual's sabotage was effective (same as in INFO SELF$_{100}$ OPP$_0$)	opponent learns with probability of 70% whether individual inflicted sabotage on him and whether individual's sabotage was effective, i.e., individual's sabotage may become visible for victim (opponent)
Period 2	*difference to Period 1:* individual can prevent information disclosure to herself about effectiveness of sabotage inflicted on opponent by herself by paying up to 10 tokens	*difference to Period 1:* individual can prevent information disclosure to herself about effectiveness of sabotage inflicted on opponent by herself by paying up to 10 tokens (same as in INFO SELF$_{100}$ OPP$_0$)	*difference to Period 1:* individual can prevent information disclosure to opponent about effectiveness of sabotage inflicted on opponent by herself by paying up to 10 tokens

Note: Individuals can prevent information disclosure in period 2 for sure only if they are willing to pay the maximum possible price of 10 tokens; otherwise, information disclosure is only prevented if their willingness to pay is at least as high as the randomly determined price.

Table 4.1: Differences between treatments in part 2 of the experiment.

is exposed indicate that these individuals by the majority do not regret their choice and, hence, report only a low willingness to pay for the opportunity to prevent detection.

4.3 Experimental Design

4.3.1 Set-Up

We implement three treatments to study the impact of private and social image concerns on sabotage behavior in competitions. Sabotage, however, is not effective for sure to provide a moral wiggle room for individuals. The treatments only differ with respect to the information individuals receive about the implication of their sabotage decisions. More precisely, individuals learn or may learn whether their sabotage inflicted on the opponent is effective and reduces his winning probability and whether the opponent may learn the extent and implication of the sabotage inflicted on him.

All treatments comprise three parts, beginning with three practice periods which give individuals the opportunity to familiarize with the payoff mechanism (see Section 4.3.2). In these rounds,

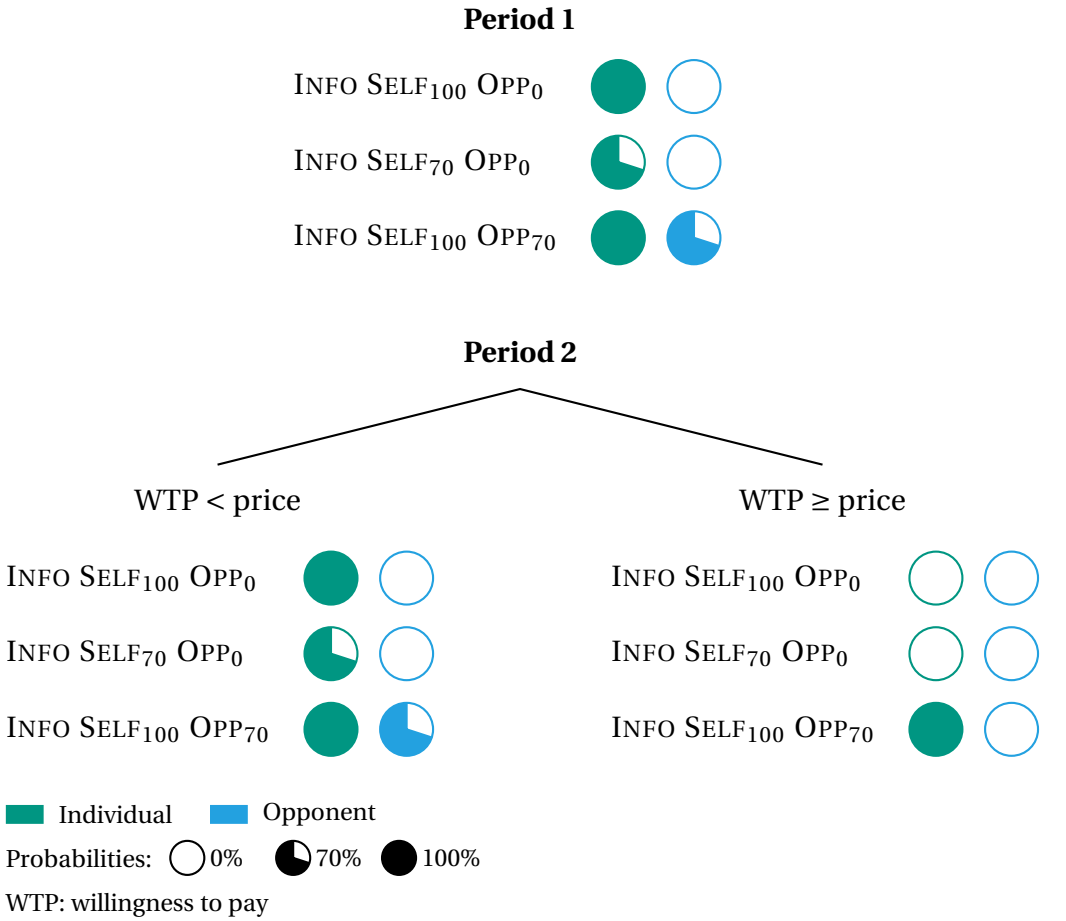


Figure 4.1: Probabilities of information disclosure about individuals' sabotage activities in part 2 of the experiment.

individuals play the role of both contestants and can simulate different strategies to observe the outcomes.

In the INFO SELF$_{100}$ OPP$_0$ treatment, individuals are in the first period of part 2 always informed about the implication of their sabotage inflicted in the opponent. In the second period, individuals can prevent receiving this information by paying a price up to 10 tokens. The opponent, however, never receives information about sabotage inflicted on him in the INFO SELF$_{100}$ OPP$_0$ treatment. The INFO SELF$_{70}$ OPP$_0$ treatment is identical to the INFO SELF$_{100}$ OPP$_0$ treatment, except for the fact that individuals receive information about the effectiveness of their sabotage inflicted on the opponent only with a probability of 70%. As in the INFO SELF$_{100}$ OPP$_0$ treatment, individuals in the INFO SELF$_{100}$ OPP$_{70}$ treatment are always informed about the effectiveness of their sabotage inflicted on the opponent. Moreover, the opponent is also informed about the sabotage inflicted on him by the individual and its implication with a probability of 70%. In the second period of part 2, individuals can prevent the disclosure of this information by paying a price up to 10 tokens. Individuals in the INFO SELF$_{100}$ OPP$_0$ and INFO SELF$_{70}$ OPP$_0$ do not receive information about the implication of sabotage inflicted on the opponent if the stated willingness to pay is sufficiently high. In the INFO SELF$_{100}$ OPP$_{70}$ treatment, however, individuals can prevent the disclosure of information to the opponent. Please note that individuals themselves are still

	Winning probability	
Absolute performance difference (points)	Leader	Trailing
0–20	50%	50%
20.5–45	60%	40%
45.5–70	70%	20%
70.5–100	80%	20%
100.5 or more	90%	10%

Table 4.2: Translation of performance differences into winning probabilities.

informed in any case about the effectiveness of their sabotage inflicted on the opponent in the second period of part 2 of the INFO SELF_{100} OPP_{70} treatment. Table 4.1 summarizes the differences between our treatments. A visual representation of different probabilities for information disclosure about individuals' sabotage activities to individuals themselves and their opponent is provided in Figure 4.1.

In both, the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment, the opponent never learns about an individual's sabotage decision. Hence, only private image concerns can matter in an individual's decision whether or not to choose a positive level of sabotage. As individuals always receive information about the effectiveness of their sabotage inflicted on the opponent in the INFO SELF_{100} OPP_0 treatment, but only with a probability of 70% in the INFO SELF_{70} OPP_0 treatment, the latter treatment provides a moral wiggle room which alleviates the impairment of an individual's private image due to a positive sabotage decision. The effect on an individual's private image in the INFO SELF_{100} OPP_{70} treatment is identical to the INFO SELF_{100} OPP_0 treatment as information to the individual herself is always disclosed in either treatment. The difference between both treatments is attributed to the information the opponent may receive in the INFO SELF_{100} OPP_{70} treatment. In contrast to the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment which at most affect the individual's private image, the INFO SELF_{100} OPP_{70} treatment can also refer to social image concerns as the decision to sabotage and its implication becomes visible for the opponent with a probability of 70%. If these concerns affect individuals' decisions and cause disutility either due to an impairment of the private or social image (or both) when deciding whether or not to choose positive levels of sabotage, individuals can make use of our mechanism to prevent information disclosure. More precisely, individuals can state their willingness to pay between 0 and 10 tokens. If their willingness to pay is at least as high as the randomly determined price, individuals have to pay this price and—depending on the treatment—no information is disclosed either to the individual herself or to the opponent.

4.3.2 Procedures

After the initial practice periods in which individuals have the opportunity to familiarize with the payoff mechanism, the second part of the experiment begins with a real-effort task followed by two periods with a dyadic tournament each. Individuals have to code five-letter words into numbers for ten minutes. Each letter of the alphabet corresponds to a two-digit number (Erkal

et al., 2011; Dato and Nieken, 2014). Each correctly coded letter is rewarded with one point, whereas each wrongly coded letter leads to a deduction of 0.5 points. All individuals get the same words in the same order. After the end of the ten-minute working period, each individual is informed about her performance, i.e., the number of achieved points. Subsequently, all individuals are matched in groups of two. Each group enters a two-player competition. The winner of the competition receives the winner prize of 75 tokens while the loser gets 25 tokens.[2] In order to determine the winner, the performance of both individuals in a group is compared. The performance difference translates into a winning probability for both individuals. If the absolute performance difference is rather low (between zero and 20 points), both contestants have a 50% chance of winning. If the difference is larger than 20 but lower than 45.5 points, the leading individual has a 60% and the trailing individual a 40% winning probability (see Table 4.2 for all combinations of winning probabilities and performance differences). Each individual is informed whether she obtained the higher or lower score within the group and the corresponding winning probability. Subsequently, all contestants have the opportunity to inflict sabotage on the opponent by potentially lowering his winning probability by zero to 9 percentage points. A higher amount of sabotage leads to higher costs, which have to be paid by the sabotaging individual in any case. All individuals know the cost function and that each contestant has the opportunity to sabotage. The chosen sabotage, however, may not be effective in reducing the winning probability of the opponent. Each session consists of 18 groups of two individuals:[3]

- For 12 randomly determined groups, the sabotage decisions of both individuals are implemented.
- For 5 randomly determined groups, only one of the sabotage decisions of the group is implemented.
- For one randomly determined group, the sabotage decisions of both individuals are ineffective.

When making their sabotage decisions, individuals do not know whether or not their own or the opponent's sabotage decision is effective. They only know the possible scenarios and their frequencies. For the second period, individuals are matched with a new opponent. As before, the performance difference between the two contestants determines their winning probabilities. Both individuals can then again select their desired sabotage levels. In contrast to the previous period, they can now affect the information disclosure concerning the sabotage decisions and their implication. Depending on the treatment, individuals can either prevent receiving information about the implication of their sabotage on their opponent or prevent the opponent from being informed about the amount and implication of sabotage inflicted on him by paying a price (see Section 4.3.1). We elicit the willingness to pay following Becker *et al.* (1964). In all treatments, individuals can spend between 1 to 10 tokens to prevent information disclosure. The computer randomly determines the price to prevent information disclosure at the end of the second part. If the price is equal to or lower than the amount an individual is willing to pay, information is not disclosed and the individual has to pay the price. If the price is higher than

the individual's willingness to pay, the information is disclosed. At the end of the experiment, all individuals are informed about the outcomes of both tournament periods of part 2.

In part three, we elicit additional controls on the social value orientation of each individual using the Equality Equivalence Test (Kerschbamer, 2015). In the final questionnaire, we also elicit risk attitudes using the risk aversion scale from the German socio-economic panel (Dohmen *et al.*, 2011), measures for private and public self-consciousness (Heinemann, 1979; based upon Fenigstein *et al.*, 1975), competitiveness via the competitiveness index (Smither and Houston, 1992), and demographics such as gender, age, and field of study.

We conducted 14 sessions with 36 individuals each between July 2017 and January 2018 at Karlsruhe Decision & Design Lab.[4] Each individual was only allowed to enroll in one session and we used a between-subject design. The data set contains observations of 469 individuals (181 women, 288 men).[5] We recruited individuals via hroot (Bock *et al.*, 2014) and programmed the experiment in z-Tree (Fischbacher, 2007). Each session lasted about 90 minutes. At the end of each session, earnings were converted into Euro at a rate of 1 token = 0.20 Euro (5 tokens = 1.00 Euro); the average payment was 15.36 Euro.

4.3.3 Behavioral Conjectures

Whereas monetary incentives remain unchanged, our treatments differ with respect to the information about sabotage activities which are disclosed or not to either individuals themselves or their opponent and the specific probabilities. Hence, any differences in individuals' behavior between the treatments can be attributed to different information policies. To derive predictions for the empirical analysis, we compare the impact of different information policies on individuals' private and social image based on the results of the experimental studies discussed in Section 4.2 which provide evidence on the impact of private and social image concerns on individuals' behavior.

The upper half of Table 4.1 describes the differences in information disclosure between our treatments. If only private image concerns are decisive for the individual's choice of a sabotage level in period 1, differences between the INFO SELF_{100} OPP_0 and INFO SELF_{100} OPP_{70} treatment do not occur as individuals are always informed about the implication of sabotage inflicted on the opponent in both treatments. In the INFO SELF_{70} OPP_0 treatment, however, this information is only disclosed with a probability of 70% which is expected to result in higher levels of sabotage as the individual may not learn whether or not her sabotage decision in fact harmed the opponent:

$$\text{sabotage}_{\text{INFO SELF}_{70}\ \text{OPP}_0} > \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_0} = \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_{70}}.$$

Conversely, if only social image concerns matter for the individual's decision to inflict sabotage on the opponent in period 1, the sabotage levels in the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment do not differ as the opponent is not informed in either treatment. This changes in the INFO SELF_{100} OPP_{70} treatment in which the opponent is informed in 70% of the cases

	Score	Sabotage period 1	Sabotage period 2	WTP	positive WTP [%]
INFO SELF_{100} OPP_0	202.96	3.27	3.04	0.44	15.22
INFO SELF_{70} OPP_0	196.60	3.00	2.67	0.41	18.13
INFO SELF_{100} OPP_{70}	202.41	3.09	2.77	1.31	44.00

Table 4.3: Descriptive statistics.

about sabotage inflicted on him which is expected to decrease sabotage activities:

$$\text{sabotage}_{\text{INFO SELF}_{70}\ \text{OPP}_0} = \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_0} > \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_{70}}.$$

Finally, if both, private and social image concerns, influence the individual's decision to inflict sabotage on the opponent, the lowest level of sabotage in period 1 prevails in the INFO SELF_{70} OPP_0 treatment as only individuals themselves are informed about the implication of their sabotage activities with a probability of 70%. In the INFO SELF_{100} OPP_0 treatment, individuals are always informed and, thus, choose lower levels of sabotage. As the opponent is additionally informed in 70% of the cases in the INFO SELF_{100} OPP_{70} treatment, this is expected to result in a further decrease in sabotage activities of the individual:

$$\text{sabotage}_{\text{INFO SELF}_{70}\ \text{OPP}_0} > \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_0} > \text{sabotage}_{\text{INFO SELF}_{100}\ \text{OPP}_{70}}.$$

In period 2, the interdependence between sabotage activities and the opportunity to prevent information disclosure by paying a price is not straightforward to disentangle. The newly introduced opportunity to prevent information disclosure may affect individuals' considerations in different ways: Individuals can either choose a higher level of sabotage and simultaneously indicate a positive willingness to pay to keep the impact on their private and/or social image constant, or they do not change their sabotage activity and use the opportunity to prevent information disclosure to restore their private and/or social image. A positive willingness to pay to prevent information disclosure can, hence, be interpreted as evidence regarding the influence of—depending on the treatment—private and social image concerns on individuals' decisions to inflict sabotage on their opponent. Yet, predictions about individuals' behavior in period 2 are not unambiguous due to this interdependence and remain an empirical question.

4.4 Results

4.4.1 Overview

An overview of the outcome variables is provided in Table 4.3. According to a non-parametric test, the scores achieved in the real-effort task do statistically not significantly differ between the treatments ($p = 0.2245$, Kruskal-Wallis H-test). The allocation of groups to categories according to both individuals' winning probabilities in period 1 differs slightly between treatments ($p = 0.095$,

Pearson's χ^2-test). Pairwise comparisons between the treatments using the Holm-Bonferroni adjustment, however, reveal no statistically significant differences. In period 2, the differences between treatments are statistically not significant ($p = 0.105$, Pearson's χ^2-test). Within treatments, the allocation of groups according to both individuals' winning probabilities does not shift into one specific direction from period 1 to period 2 (INFO SELF_{100} OPP_0: $p = 0.7707$, INFO SELF_{70} OPP_0: $p = 0.8958$, INFO SELF_{100} OPP_{70}: $p = 0.1229$, Bowker test of symmetry).

Sabotage levels in period 1 and period 2 by treatment are depicted in Figure 4.2. There are no statistically significant differences between treatments in period 1 ($p = 0.8105$) and period 2 ($p = 0.3979$, Kruskal-Wallis H-test). The difference between leading and trailing individuals within groups is statistically not significant in the INFO SELF_{100} OPP_0 (period 1: $p = 0.3004$, period 2: $p = 0.4653$) and the INFO SELF_{70} OPP_0 treatment (period 1: $p = 0.1185$, period 2: $p = 0.4859$). In the INFO SELF_{100} OPP_{70} treatment, leading individuals inflict statistically significantly higher levels of sabotage on their opponent (period 1: $p = 0.0990$, period 2: $p = 0.0039$, two-sided Wilcoxon signed-rank test). Sabotage levels between treatments in period 1 are statistically not signficantly different for leading ($p = 0.1740$) and trailing individuals ($p = 0.5026$, Kruskal-Wallis H-test). Sabotage levels between treatments in period 2 are statistically not significantly different for leading individuals ($p = 0.1646$), whereas they are for trailing individuals ($p = 0.0474$, Kruskal-Wallis H-test). In partiular, trailing individuals in the INFO SELF_{100} OPP_{70} choose statistically significantly lower levels of sabotage than in the other treatments (INFO SELF_{100} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0339$, INFO SELF_{70} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0534$, pairwise Dunn test with Holm-Bonferroni adjustment).[6]

Overall, there is a slight decline in sabotage levels from period 1 to period 2. This is indicated in Figure 4.3[7] as the slope of the solid red line representing the linear fit is flatter than the dashed 45°-line. Within treatments, sabotage levels decrease statistically significantly from period 1 to period 2 in the INFO SELF_{70} OPP_0 ($p = 0.0509$) and INFO SELF_{100} OPP_{70} treatment ($p = 0.0259$), but not in the INFO SELF_{100} OPP_0 treatment ($p = 0.3140$, two-sided Wilcoxon signed-rank test).

The willingness to pay to prevent information disclosure in period 2 is notably higher in the INFO SELF_{100} OPP_{70} treatment than in the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment; this difference is statistically highly significant (equality of groups: $p = 0.0001$, Kruskal-Wallis H-test; INFO SELF_{100} OPP_0 vs. INFO SELF_{70} OPP_0: $p = 0.3557$, INFO SELF_{100} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0000$, INFO SELF_{70} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0000$, pairwise Dunn test with Holm-Bonferroni adjustment). Moreover, the share of individuals who indicate a positive willingness to pay differs statistically significantly between treatments ($p = 0.0000$, Pearson's χ^2-test). More specifically, there are statistically significantly more individuals in the INFO SELF_{100} OPP_{70} treatment who are willing to pay a positive price to prevent information disclosure than in the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment (INFO SELF_{100} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0000$, INFO SELF_{70} OPP_0 vs. INFO SELF_{100} OPP_{70}: $p = 0.0000$, pairwise Pearson's χ^2-test with Holm-Bonferroni adjustment). In the INFO SELF_{70} OPP_0 treatment, trailing individuals indicate a statistically marginally higher willingness to pay than leading individuals ($p = 0.0781$),

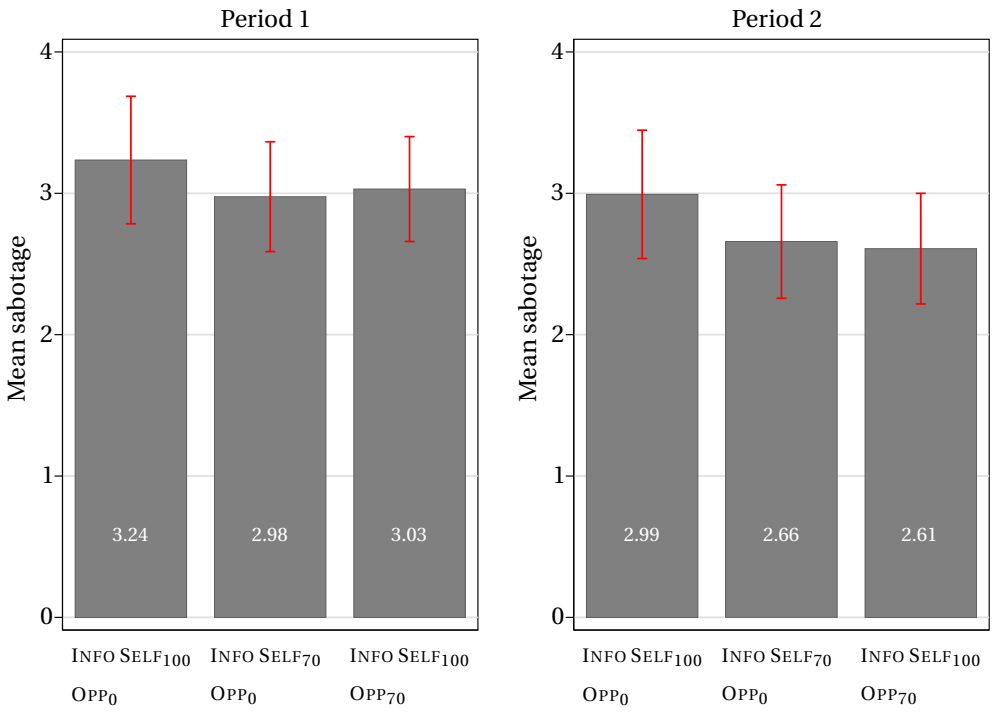


Figure 4.2: Sabotage in period 1 and period 2 by treatment. Error bars indicate the 95% confidence intervals.

whereas the difference between leading and trailing individuals is statistically not significant in the other treatments (INFO SELF_{100} OPP_0: $p = 0.5651$, INFO SELF_{100} OPP_{70} $p = 0.2219$, two-sided Wilcoxon signed-rank test).

4.4.2 Sabotage Behavior

The results of regressions analyses on sabotage behavior in period 1 and period 2 are reported in Tables 4.4 and 4.5, respectively. Starting with period 1, there are no differences between the INFO SELF_{70} OPP_0 and the INFO SELF_{100} OPP_{70} treatment and the INFO SELF_{100} OPP_0 treatment. The coefficients of the dummy variables for either treatment are statistically insignificantly different from zero. Despite this insignificance it has to be noted that the negative sign of the coefficients of the dummy variable for the INFO SELF_{70} OPP_0 treatment is not inline with our conjecture according to which a lower probability of information disclosure than in the INFO SELF_{100} OPP_0 treatment is expected to result in higher levels of sabotage if the individuals' decision is influenced by self image concerns. The negative and statistically highly significant coefficients for the squared distance from the 50/50-category which corresponds to equal winning probabilities for both individuals of a given group indicates that sabotage activities decrease substantially for both individuals the further their group is away from the from the 50/50-category which corresponds to increasingly unequal winning probabilities within these groups. Put differently, sabotage is used more intensively when both individuals of a group are head-to head. An illustration of this relation is provided in Figure 4.4. The variable Distance indicates individuals' winning probabilities relative to the reference category with a balanced competition. The

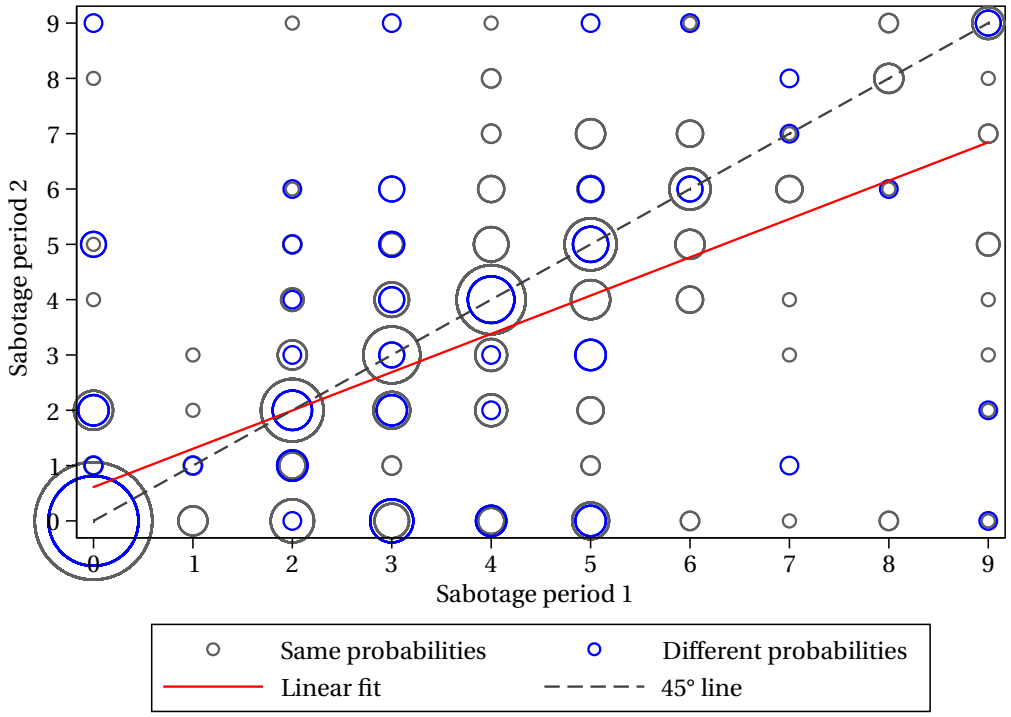


Figure 4.3: Sabotage in period 1 and period 2. The size of the bubbles indicates the frequency of distinct observations.

values of this variable range from −4 for individuals with a winning probability of 10% to 4 for individuals with a winning probability of 90%, respectively. Note that the absolute distance for both individuals within a group is the same. The red lines indicating an inverted U-shape represent the quadratic fit of sabotage levels by distance from the 50/50-category. Adding additional control variables in columns (3) and (4) does not change these results. Individuals who are prepared to take risks inflict statistically significantly higher levels of sabotage on their opponent. The same applies for females in specification (4), although the effect is statistically not significant any more when adding interaction terms for females and the treatment variables. Additional control variables are the three sub-scales emotion, argument, and games from the competitiveness index by Smither and Houston (1992), measures for private and public self-consciousness (Fenigstein *et al.*, 1975; Heinemann, 1979), and dummy variables for different types according to the social value orientation (Kerschbamer, 2015).[8]

Looking at the regression results on sabotage in period 2 reported Table 4.5, results remain by and large unchanged. One notable exception, however, are the coefficients for the dummy variable $\textsc{Info Self}_{70}\ \textsc{Opp}_0$ in columns (2)–(4). Whereas this coefficient is statistically not different from zero for conventional significance levels in column (1) containing only dummy variables for the treatments, the coefficient is negative and statistically marginally significant when adding additional control variables due to more precise estimates.

Robustness checks with separate regressions for leading and trailing individuals in both periods are reported in Tables C.4 and C.5. These results are largely in line with the results for the pooled

	(1)	(2)	(3)	(4)
INFO SELF_{70} OPP_0	−0.430	−0.576	−0.569	−0.684
	(0.433)	(0.430)	(0.415)	(0.508)
INFO SELF_{100} OPP_{70}	−0.339	−0.333	−0.358	−0.684
	(0.423)	(0.413)	(0.400)	(0.491)
Distance		−0.030	−0.051	−0.053
		(0.088)	(0.087)	(0.087)
Squared distance		−0.213***	−0.224***	−0.222***
		(0.048)	(0.048)	(0.048)
Risk			0.294***	0.300***
			(0.086)	(0.086)
Econ			0.121	0.116
			(0.335)	(0.335)
Female			1.035***	0.584
			(0.372)	(0.688)
INFO SELF_{70} OPP_0 × Female				0.359
				(0.862)
INFO SELF_{100} OPP_{70} × Female				0.872
				(0.840)
Constant	2.843***	3.709***	2.836*	2.911**
	(0.324)	(0.375)	(1.463)	(1.466)
Competitiveness	No	No	Yes	Yes
Self-consciousness	No	No	Yes	Yes
Social value orientation	No	No	Yes	Yes
Obs.	469	469	469	469
Pseudo R^2	0.001	0.012	0.026	0.027

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4.4: Sabotage in period 1. Tobit estimates (lower limit: 0, upper limit: 9; 128 observations left-censored, 20 observations right-censored).

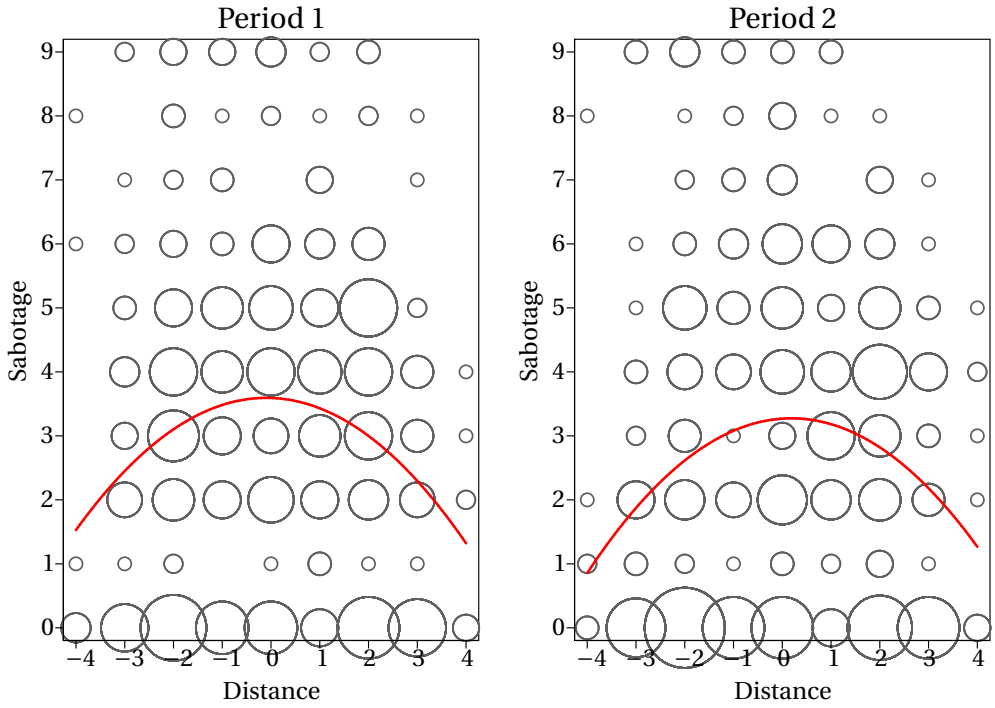


Figure 4.4: Individual winning probabilities and sabotage in period 1 and period 2. The size of the bubbles indicates the frequency of distinct observations. Positive and negative values of the variable Distance correspond to leading and trailing individuals, respectively.

sample discussed above. Moreover, the robustness checks do not reveal major differences in sabotage behavior between leading and trailing individuals which complements the results from the non-parametric tests in Section 4.4.1. Treatment differences in period 1 only occur between the INFO SELF_{100} OPP_0 and the INFO SELF_{70} OPP_0 treatment for leading individuals when controlling for a large set of individuals' characteristics. The negative sign of the coefficients for the dummy variable representing the INFO SELF_{70} OPP_0 treatment which indicate a lower propensity to inflict sabotage on the opponent is not in line with our behavioral conjectures as the individual's probability of being informed about the implication of her sabotage activities is lower than in the INFO SELF_{100} OPP_0 treatment. In period 2, differences occur between the INFO SELF_{100} OPP_0 and the INFO SELF_{100} OPP_{70} treatment for trailing individuals. This observation corroborates the notion that the possibility of information disclosure to the opponent induces individuals to choose lower levels of sabotage due to social image concerns. This effect is, however, not particularly robust.

4.4.3 Preventing Information Disclosure

Besides the decision whether and how much sabotage to inflict on their opponent, individuals have in period 2 additionally the opportunity to prevent information disclosure about their sabotage activities, depending on the treatment either to themselves or to the opponent by paying a price. More specifically, individuals indicate their willingness to pay a price between 1 and 10 tokens; if their willingness to pay is at least as high as the randomly determined price, no

	(1)	(2)	(3)	(4)
INFO SELF_{70} OPP_0	–0.667	–0.811*	–0.763*	–1.054*
	(0.474)	(0.468)	(0.448)	(0.556)
INFO SELF_{100} OPP_{70}	–0.709	–0.626	–0.568	–0.727
	(0.474)	(0.464)	(0.449)	(0.563)
Distance		0.107	0.091	0.084
		(0.100)	(0.099)	(0.100)
Squared distance		–0.222***	–0.239***	–0.240***
		(0.052)	(0.052)	(0.052)
Risk			0.389***	0.390***
			(0.094)	(0.094)
Econ			0.050	0.039
			(0.375)	(0.375)
Female			0.916**	0.457
			(0.407)	(0.719)
INFO SELF_{70} OPP_0 × Female				0.791
				(0.950)
INFO SELF_{100} OPP_{70} × Female				0.473
				(0.943)
Constant	2.405***	3.277***	0.802	0.938
	(0.345)	(0.387)	(1.648)	(1.655)
Competitiveness	No	No	Yes	Yes
Self-consciousness	No	No	Yes	Yes
Social value orientation	No	No	Yes	Yes
Obs.	469	469	469	469
Pseudo R^2	0.001	0.012	0.029	0.029

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4.5: Sabotage in period 2. Tobit estimates (lower limit: 0, upper limit: 9; 162 observations left-censored, 17 observations right-censored).

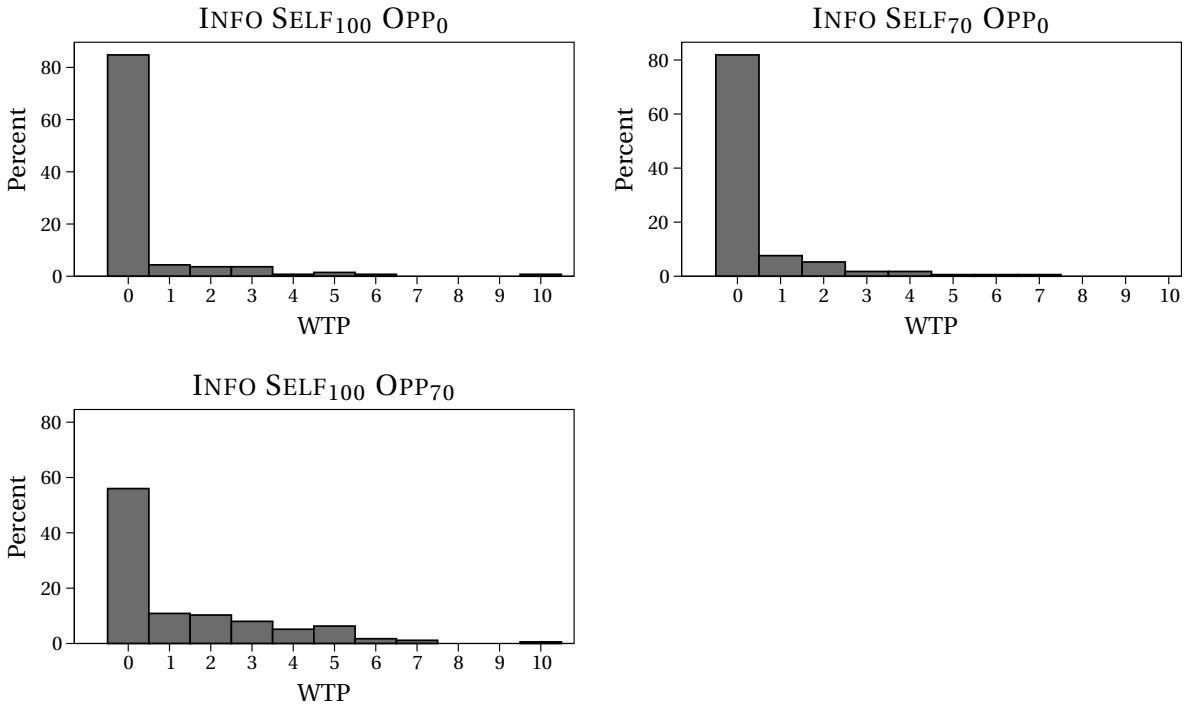


Figure 4.5: Willingness to pay to prevent information disclosure in period 2 by treatment.

information is disclosed. In the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment, the opponent never receives information about sabotage inflicted on him and the implication thereof; thus, individuals only have the opportunity to prevent themselves from receiving information about the implication of their sabotage activities by paying a price. If the willingness to pay is lower than the randomly determined price, individuals in the INFO SELF_{100} OPP_0 treatment are informed for sure and individuals in the INFO SELF_{70} OPP_0 treatment are informed with a probability of 70%. In contrast, individuals in the INFO SELF_{100} OPP_{70} treatment always receive this information. In this treatment, they can prevent the opponent from being informed about sabotage inflicted on him and the corresponding implication. If an individual's willingness to pay is not sufficiently high, the opponent receives this information with a probability of 70%.

In line with non-parametric tests reported in Section 4.4.1, Figure 4.5 shows that a higher share of individuals in the INFO SELF_{100} OPP_{70} treatment indicates a positive willingness to pay than in the INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0 treatment. Moreover, the average willingness to pay indicated in the INFO SELF_{100} OPP_{70} treatment is higher than in the other treatments as the corresponding panel for the INFO SELF_{100} OPP_{70} treatment exhibits more observations for strictly positive values on the horizontal axis (see also Table 4.3). This is confirmed by the regression results reported in Table 4.6 as all coefficients for the dummy variable representing the INFO SELF_{100} OPP_{70} treatment are positive and statistically highly significant. Individuals in the 50/50-category indicate a statistically (marginally) lower willingness to pay. One possible interpretation may be that individuals who face close competition tend to consider sabotage a legitimate means to surpass their opponent and reduce the impact of chance on receiving the winner prize. Similarly, leading individuals indicate a statistically (marginally) significantly

	(1)	(2)	(3)	(4)	(5)	(6)
INFO SELF_{70} OPP_0	0.207	0.115	0.199	0.166	0.239	0.498
	(0.709)	(0.712)	(0.678)	(0.590)	(0.556)	(0.697)
INFO SELF_{100} OPP_{70}	3.104***	2.981***	2.986***	1.675***	1.556***	1.657***
	(0.638)	(0.641)	(0.601)	(0.578)	(0.544)	(0.623)
Tie		−1.338*	−1.808**	−1.587**	−1.309**	−1.316**
		(0.736)	(0.714)	(0.667)	(0.622)	(0.620)
Leading individual		−1.001*	−1.116**	−0.760*	−0.748*	−0.738
		(0.556)	(0.503)	(0.458)	(0.451)	(0.451)
Sabotage			0.617***	0.468***	0.443***	0.445***
			(0.101)	(0.102)	(0.098)	(0.098)
Belief				0.063***	0.065***	0.064***
				(0.007)	(0.006)	(0.006)
Risk					−0.106	−0.106
					(0.105)	(0.103)
Econ					0.001	0.010
					(0.410)	(0.409)
Female					0.965**	1.254
					(0.443)	(0.860)
INFO SELF_{70} OPP_0 × Female						−0.591
						(1.143)
INFO SELF_{100} OPP_{70} × Female						−0.234
						(1.014)
Constant	−4.032***	−3.280***	−4.725***	−5.618***	−4.084**	−4.187**
	(0.623)	(0.660)	(0.702)	(0.693)	(1.775)	(1.821)
Competitiveness	No	No	No	No	Yes	Yes
Self-consciousness	No	No	No	No	Yes	Yes
Social value orientation	No	No	No	No	Yes	Yes
Obs.	469	469	469	469	469	469
Pseudo R^2	0.037	0.042	0.093	0.171	0.188	0.189

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4.6: Willingness to pay to prevent information disclosure in period 2. Tobit estimates (lower limit: 0, upper limit: 10; 351 observations left-censored, 2 observations right-censored).

lower willingness to pay to prevent information disclosure except in column (6). This can be interpreted evidence supporting the notion that individuals who obtain a higher score than their opponent perceive sabotage a legitimate means to assure receiving the winner prize which they may feel entitled to. Higher levels of sabotage inflicted on the opponent also increase statistically highly significantly the willingness to pay to prevent information disclosure. One potential explanation for this relation may be that the disutility incurred in case of information disclosure either to the individuals themselves or to the opponent due to an impairment of the private and/or social image increases in the level of sabotage inflicted on the opponent. If this expected disutility is larger than the monetary reduction of the final payoff due to the costs to prevent information disclosure, individuals' willingness to pay increases with their sabotage activities. The coefficients of individuals' beliefs regarding the share of other individuals who indicate a positive willingness to pay to prevent information disclosure are also positive and statistically highly significant. As the level of sabotage inflicted on the opponent and the belief about other individuals' willingness to pay are positively correlated (Spearman's $\rho = 0.1841$, $p = 0.0000$), this may be indirectly related to *ex-post* rationalization of individuals who assume others to behave similar to themselves. Finally, females are willing to pay a statistically significantly higher amount to prevent information disclosure pointing at different levels of disutility incurred by gender if information about one's own sabotage activities is revealed. This effects disappears, however, when adding interaction terms between the dummy variables for the treatments and female individuals.

Table C.6 reports robustness checks using a truncated normal hurdle model to disentangle the decisions of *(i)* whether or not to pay to prevent information disclosure and *(ii)* how much to pay if the previous decision was positive. In the first stage of the hurdle model, the coefficients of the dummy variable for the INFO SELF_{100} OPP_{70} treatment are positive and statistically highly significant which confirms a substantially higher propensity of individuals in this treatment to indicate a positive willingness to pay as compared to the reference group in the INFO SELF_{100} OPP_{0} treatment. For those individuals who report a positive willingness to pay, no treatment differences occur except between the INFO SELF_{100} OPP_{0} and INFO SELF_{100} OPP_{70} treatment in column (6) when controlling for a large set of individual characteristics. Individuals facing close competition in the 50/50-category exhibit a statistically (marginally) lower willingness to pay, whereas the effect for leading individuals with winning probabilities of 60–90% is no longer statistically significant. The coefficients for the level of sabotage inflicted on the opponent in period 2, however, is still positive and statistically highly significant which supports the notion that the disutility in case of disclosure of an individual's sabotage activities increases in the level of sabotage inflicted on the opponent.

4.5 Conclusion

The present paper studies the influence of private and social image concerns on individuals' decision to inflict sabotage on their opponent in a tournament. Whereas sabotage potentially increases the probability of winning the tournament and receiving the winner prize, the disclo-

sure of information about the implication of sabotage activities for the opponent either to the individuals themselves or to the opponent can induce disutility due to an impairment of the private and/or social image.

The results of our experiment show no differences in sabotage levels between our treatment groups establishing different levels of information disclosure. Thus, individuals' decision to inflict sabotage on their opponent is neither influenced by the fact they (may) obtain information about the implication of their sabotage activities nor that their opponent may receive this information. This observation is in line with Harbring and Wilhelm (2016) who find a substantial share of individuals whose sabotage has been exposed not to indicate any signs of regret as they state their intention to behave similarly again if the scenario is repeated. One explanation for this phenomenon may be the nature of the underlying experiment. As the opportunity to inflict sabotage on the opponent as well as the implication for the opponent and its probability of occurrence are explicitly mentioned in the instructions, individuals may consider sabotage as legitimate means to increase their winning probability in the tournament as these rules are common knowledge—unlike in situations outside the laboratory.

Offering the option to pay a price to prevent information disclosure about the implication of sabotage either to individuals themselves or the opponent reveals differences between the treatment groups. In particular, individuals in the INFO SELF_{100} OPP_{70} treatment indicate a statistically significantly higher willingness to pay to prevent information disclosure than individuals in the INFO SELF_{100} OPP_{0} and INFO SELF_{70} OPP_{0} treatment. Whereas only individuals themselves are informed about the implication of their sabotage activities for the opponent in the latter treatments with probabilities of 100% and 70%, respectively, the INFO SELF_{100} OPP_{70} treatment comprises the possibility of additionally disclosing this information to the opponent with a probability of 70%. As individuals themselves are informed in any case about the implication of their sabotage activities just as in the INFO SELF_{100} OPP_{0} treatment, the possibility of being exposed to the sufferer induces individuals in the INFO SELF_{100} OPP_{70} treatment to pay a substantially higher amount of tokens to prevent information disclosure to the opponent. This provides evidence that even in an anonymous setting in the laboratory the peril of one's sabotage activity becoming visible for another individual induces disutility in terms of an impairment of individuals' social image which is reflected by a higher willingness to pay to prevent information disclosure to the opponent. Furthermore, leading individuals with winning probabilities of 60–90% and individuals in the 50/50-category facing close competition indicate a lower willingness to pay to prevent information disclosure than trailing individuals with winning probabilities of 10–40%. Possible explanations for these observations are that individuals who already obtained a higher winning probability due to a better performance in the real-effort task than their opponent may feel entitled to receiving the winner prize and use sabotage, hence, to assure their claims. For individuals in the 50/50-category, the impact of chance on receiving the winner prize is substantial as both individuals of a group have equal winning probabilities. To enhance one's prospect of winning the tournament, the use of sabotage may be considered legitimate which decreases their willingness to pay to prevent information

disclosure. Furthermore, the level of sabotage inflicted on the opponent and the belief regarding the share of other individuals who indicate a positive willingness to pay to prevent information disclosure increase the willingness to pay. Considering the extensive and intensive margin of the willingness to pay to prevent information disclosure, however, yields that the participation decision is substantially more pronounced for individuals in the INFO SELF_{100} OPP_{70} treatment, whereas the remaining control variables can hardly provide a clear-cut picture of the indicated amounts except the level of sabotage inflicted on the opponent.

The results of this study show that neither private nor social image concerns fully prevent individuals from inflicting sabotage on their opponent in a tournament as part of controlled laboratory study. Moreover, we do not observe notable differences in sabotage levels between the respective treatments. Yet, when offered the opportunity to prevent information disclosure about their sabotage activities, individuals in the INFO SELF_{100} OPP_{70} treatment who can avoid being exposed to their opponent indicate a higher willingness to pay than individuals who can only prevent the disclosure of this information to themselves, i.e., concerns for individuals' social image do matter in this context.

Notes

1. Another alternative is cheating, i.e., claiming a higher output than actually achieved (e.g., Schwieren and Weichselbaumer, 2010; Cartwright and Menezes, 2014).
2. The exchange rate is 5 tokens = 1.00 Euro.
3. The reason for this partition is to induce an environment in which an individual's decision to sabotage the opponent is effective with a probability of approximately 80%:

$$\begin{aligned}
&\Pr(\text{both decisions effective}) = 0.8 \cdot 0.8 = 0.64 \approx \tfrac{12}{18} \\
&\Pr(\text{one decision effective}) = (0.8 \cdot 0.2) + (0.2 \cdot 0.8) = 0.32 \approx \tfrac{5}{18} \\
&\Pr(\text{no decision effective}) = 0.2 \cdot 0.2 = 0.04 \approx \tfrac{1}{18}
\end{aligned}$$

4. We conducted three more sessions that are not part of the paper. Due to technical problems, the recorded data was incomplete.
5. The deviation from the theoretical number of individuals ($14 \cdot 36 = 504$) is due to the fact that in some sessions our student assistants had to replace participants who did not show up for the experiment. This was necessary to guarantee that that exactly 18 groups participated in each session (see Note 3 above). Our student assistants and participants who interacted with them were excluded from the data set. Of the remaining 484 individuals, 15 made inconsistent choices in the Equality Equivalence Test and were, thus, also excluded.
6. A detailed overview of sabotage levels by individual winning probabilities is reported in Appendix C.1.
7. Note that individuals with the same and different winning probabilities in period 1 and 2

are depicted separately. Overall, sabotage levels are statistically not significantly different in period 1 ($p = 0.1216$) and period 2 ($p = 0.2512$, two-sided Mann-Whitney U-test) for individuals who obtain the same or different winning probabilities.

8. The reference category is *Selfish* which accounts for 66.53% of all individuals. Figure C.4 indicates the frequencies of the respective types.

5 Conclusion

The research presented in this thesis relates to different aspects which are prevalent in the workplace if individuals are required to work together: coordination, cooperation, and sabotage in tournaments. As outlined in Section 1, companies make frequently use of organizational settings in which individuals have to collaborate. A major advantage of such instances is that individuals can complement each other in their skills and knowledge and, thereby, realize efficiency gains over individual work settings. It is, however, necessary that individuals refrain from free riding and coordinate their activities to utilize the advantages of team production. Furthermore, companies frequently use tournaments in which individuals compete for a prize such as a bonus or promotion. These settings are particularly beneficial if individual performance is difficult to measure in absolute terms. Hence, rank-order tournaments using a relative information can alleviate the measurement issue. The major drawback is, however, that individuals cannot only improve their prospects of winning by increasing their productive effort—but also by means of sabotage, i.e., destructive effort.

The study discussed in Section 2 is concerned with coordination in a threshold public good game. In particular, groups of several individuals are required to *(i)* choose one out of four collective accounts and *(ii)* coordinate their contributions to reach the corresponding threshold. To alleviate coordination in absence of communication, a semi-sequential structure is employed, i.e., the allocation of a first-moving leader is disclosed to the second-moving followers before they decide about their allocations simultaneously and independently. By design, all individuals of a group are required to contribute some share of their endowment to reach a threshold which refrains individuals from cheap riding. The institution of a first-moving leader whose allocation becomes visible before second-moving followers are asked to make their allocation can reduce strategic uncertainty. If the first mover does not contribute at all or makes only a negligibly small contribution to a collective account, none of the thresholds is feasible for the group. The alternative is a contribution by the first mover to one particular collective account which meets or exceeds some critical value such that the followers' endowment is sufficient to reach the corresponding threshold. This allocation solves one domain of the coordination problem, the choice of one out of four collective accounts. Followers are then left with a subset of the original coordination problem, namely to coordinate their contributions to reach the corresponding threshold. If first movers and followers have homogeneous incentives regarding the collective accounts, followers have no reason to object to the first mover's choice of a particular collective account. On the other hand, if individuals have different incentives regarding the choice of a collective account, the allocation by a first mover to the collective account which yields the highest return for herself may change followers' perceptions—despite the fact that they are better off supporting a selfish first mover as compared to an outcome in which no threshold is reached. The results of the corresponding experiment provide evidence that individuals do not behave differently under homogeneous and heterogeneous incentives. In fact, first movers contribute notable amounts to a collective account in the first stage to act as good example for

the second-moving followers and, thereby, alleviate coordination. Although first movers in the setting with heterogeneous incentives frequently choose the alternative which yields the highest return for themselves, followers do not object to selfish behavior by first movers and provide the remaining amount to reach the corresponding threshold instead. Put differently, even followers in the setting with heterogeneous incentives do realize the material advantage of supporting the first mover's choice which is different from their own preferred choice over not reaching a threshold at all.

An approach to foster cooperative behavior of children aged 10–13 years who attend upper secondary schools in Germany by using an intervention which provides enhanced physical education lessons is presented in Section 3. Although positive effects of physical activity and sports participation on labor market outcomes are widely acknowledged, it is difficult to establish a causal link due to concerns about selection effects. To account for this issue, the current study using a randomized controlled trial provides a comparison of a Treatment group of individuals who participate in the intervention and a Control group of individuals who attend regular physical education lessons during the same period. The intervention is designed to convey social skills including cooperativeness in a playful way. To measure the impact of the intervention, individuals in the Treatment and Control group participated in a series of different incentivized standard economic experiments including a linear public good game. These measurements were conducted before the start and right after the end of the intervention as well as nine weeks later. The comparison with the Control group allows to disentangle the effect of the intervention on cooperative behavior from effects which may occur due to the repeated measurements. A comparison of decisions in the first measurement before the beginning of the intervention does not reveal any differences between individuals in the Treatment and Control group which indicates that the randomization was successful. A comparison with individuals who attend a so-called elite school of sports which provides conditions to reconcile compulsory school attendance with high intensity levels of sports participation reveals a higher willingness to cooperate than among individuals in the Treatment and Control group. While this difference is striking, the design of the current study does not allow draw causal conclusions on the relation between the intensity of sports participation and cooperative behavior. A comparison of cooperative behavior of individuals in the Treatment and Control group as measured by contributions to the collective account in the linear public good game over the course of the study does not reveal substantial differences at first sight, neither between nor within both groups. When taking the socio-economic status of the households individuals live in into account, a more differentiated pattern of cooperative behavior over the course of the study emerges. In particular, individuals from households with low socio-economic status in the Treatment group become more cooperative over time which emphasizes the effectiveness of the intervention. Considering the nature of the intervention, this finding appears promising: Physical education lessons were only slightly adapted during the intervention. The most obvious changes are two periods of ten minutes at the beginning and end of each lessons in which individuals are encouraged to discuss their views and experiences on the topic covered in the respective lesson. Furthermore, cooperative behavior conveyed

in the intervention and measured in the linear public good game relate to the same domain of individuals' behavior but are not perfectly congruent to prevent potential demand effects. Finally, the intervention was conducted during a very limited time span of six weeks. Despite these constraints, a positive effect on cooperative behavior for individuals from households with low socio-economic status is observed which suggests that even cautious changes in the curriculum can induce improvements in prosocial behavior. Moreover, physical education at school provides a reasonable environment due to the possibility to promote approaches of informal learning which benefits all children irrespective of their background—which is usually not the case in, e.g., sports clubs.

To tackle issues which arise due to sabotage in situations with relative performance evaluation, it is essential to understand the underlying motives. Therefore, Section 4 reports the results of an experiment designed to assess the impact of different layers of information disclosure about the implication of sabotage inflicted on the opponent on sabotage behavior in tournaments. More specifically, three different treatments are implemented to vary whether the individual herself and the opponent receives or may receive information about the individual's sabotage activities. The rationale behind this approach to assess whether this kind information disclosure provides an effective countermeasure to reduce the extent of sabotage in tournaments if the individual incurs disutility when being exposed as saboteur due to impairment of her private and social image, respectively. Moreover, individuals can indicate their willingness to pay to prevent information disclosure—depending on the treatment—either to themselves or the opponent in the second stage of the experiment. Overall, sabotage levels do not substantially differ between the treatments using different layers of information disclosure. When looking at individual winning probabilities as defined by differences in the scores from a previous real-effort task between both individuals of a group, it turns out that individuals who face close competition, i.e., find themselves in a group with (almost) balanced winning probabilities for both individuals, decide to sabotage their opponent more severely. In both treatments in which only individuals themselves receive information about the implication of their sabotage activities, the opportunity to prevent information disclosure by paying a price is barely used. In contrast, if it is possible that the opponent receives this information, individuals are willing to pay a positive amount to prevent information disclosure. This observation suggests that individuals who have the opportunity to inflict sabotage on their opponent do not only trade off the direct costs of sabotage against the increased probability of receiving the winner prize in a payoff-maximizing fashion. In fact, the possibility that the opponent receives information about sabotage inflicted on him induces individuals to sacrifice some share of their payoff to keep this information under disguise. Whereas in the other two treatments only the individual herself—albeit with different probabilities—receives information about the implications of her sabotage activity, the third treatment features the possibility that the opponent is informed as well which indicates that individuals' willingness to pay is influenced by concerns about their social image. Although disclosure of information about the implications individuals' sabotage activities to the opponent does not reduce sabotage levels *per se*, the positive willingness to pay suggests that individuals do

not want their opponent to be informed about sabotage inflicted on him despite the anonymous environment in the laboratory which emphasizes the impact of non-monetary motives on unethical behavior.

Appendix A Supplementary Material for Section 2

A.1 Total Contributions

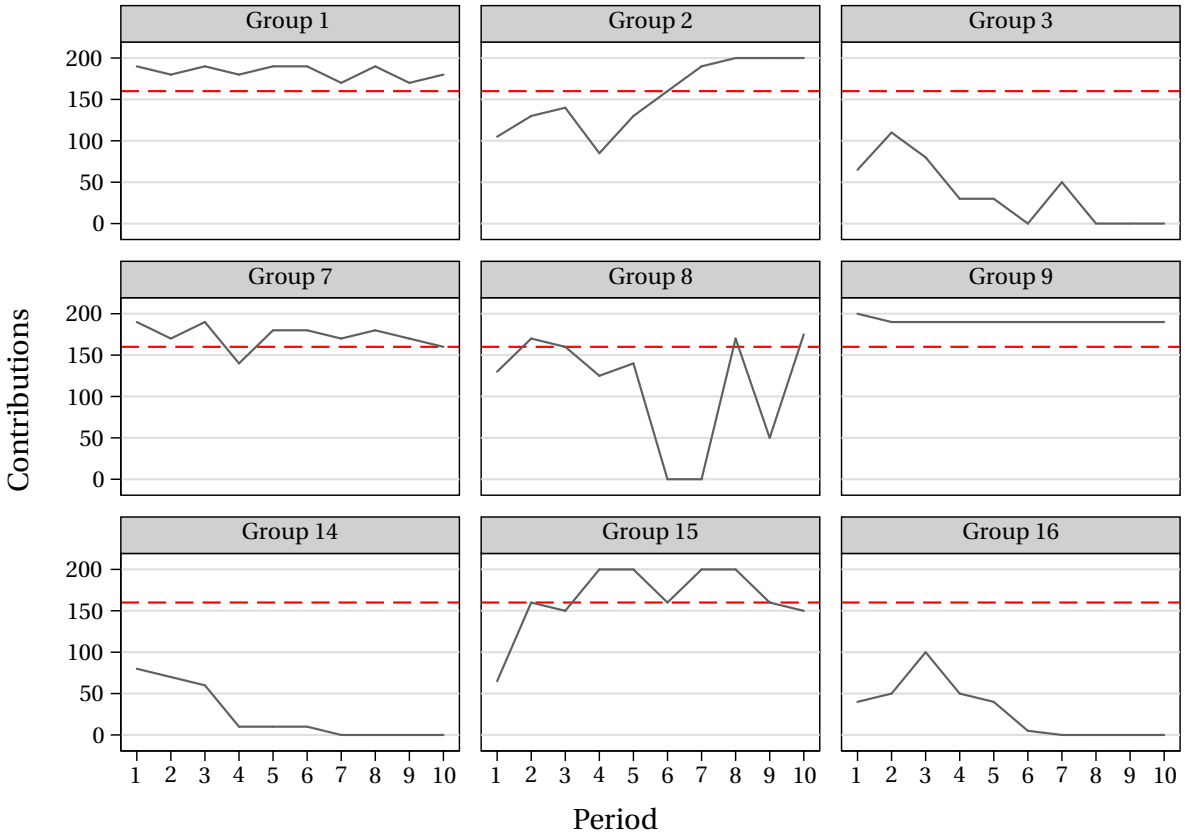


Figure A.1: Total contributions to the groups' preferred collective accounts by period in the homogeneous treatment. The dashed red line indicates the threshold of 160 tokens.

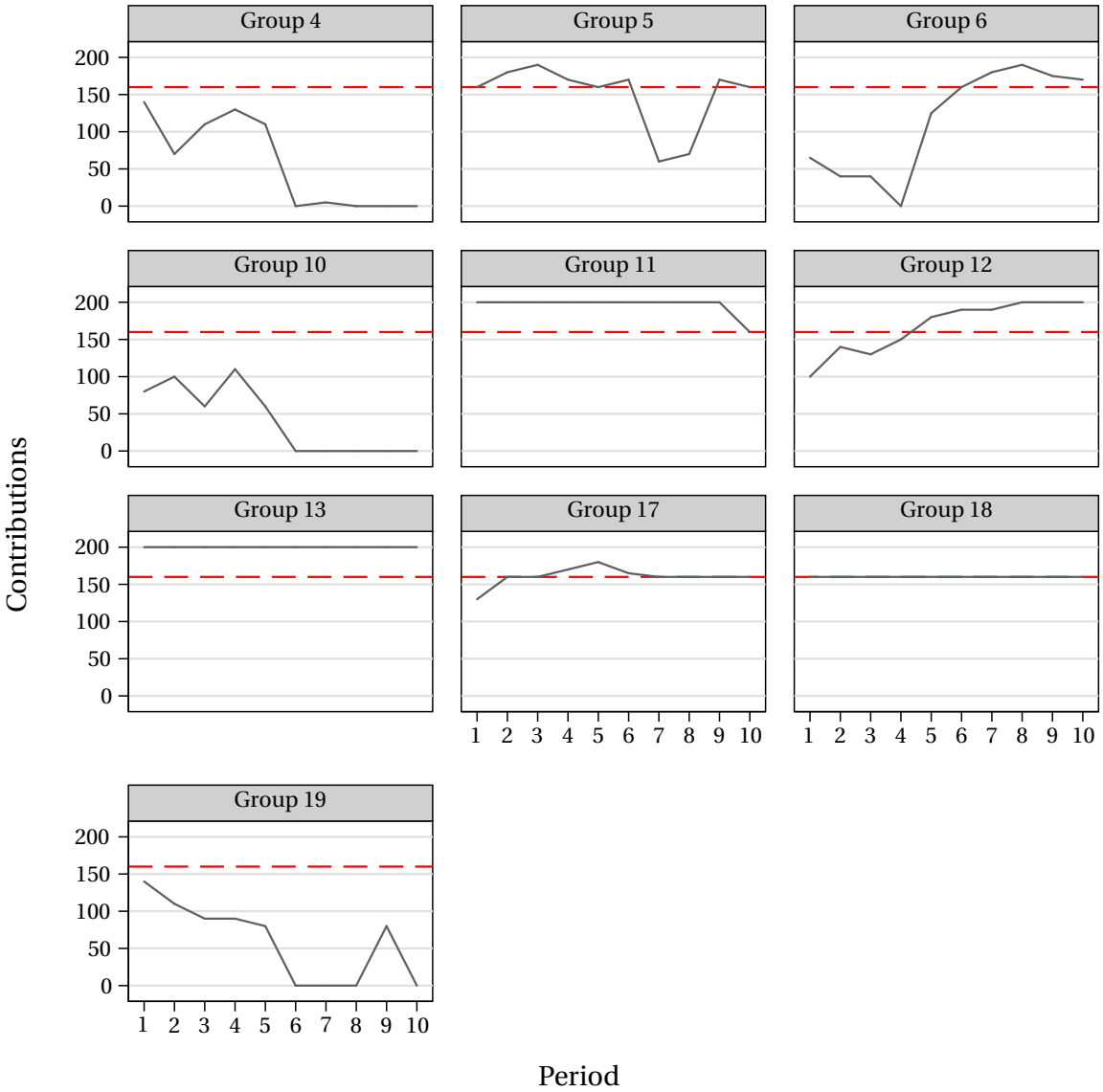


Figure A.2: Total contributions to the groups' preferred collective accounts by period in the heterogeneous treatment. The dashed red line indicates the threshold of 160 tokens.

A.2 Additional Parametric Specifications

	(1)	(2)
Het	2.975	1.267
	(7.308)	(4.256)
Period	–0.796	–1.445***
	(0.487)	(0.433)
LagThr		19.922***
		(3.980)
Constant	34.754***	29.018***
	(5.166)	(4.038)
Obs.	190	171
No. Groups	19	19
Wald-χ^2	2.845	25.96
Prob. > χ^2	0.241	$9.70 \cdot 10^{-6}$

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.1: Average contributions by group and period. Generalized least squares estimates with random effects.

	(1)	(2)	(3)	(4)	(5)
Het	0.046	–0.007	0.187	3.218	0.920
	(8.415)	(8.271)	(9.141)	(10.994)	(9.215)
Econ	–5.972	–6.109	–5.973	0.499	–5.878
	(12.353)	(12.759)	(13.320)	(16.401)	(13.008)
Female	–2.650	–2.818	–3.121	–7.107	–4.740
	(13.596)	(14.212)	(17.503)	(12.479)	(15.438)
Risk		–0.067	–0.149	–1.112	–0.144
		(1.923)	(3.103)	(2.141)	(2.100)
Q06			–0.353		
			(8.387)		
Q07				7.922	
				(8.881)	
Q08					3.289
					(10.879)
Constant	36.619**	37.217*	38.372	16.259	35.754
	(14.163)	(18.705)	(37.004)	(32.248)	(20.203)
Obs.	19	19	19	19	19
R^2	0.023	0.023	0.023	0.104	0.030

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.2: Average contributions of first movers by group over all periods. Ordinary least squares estimates.

	(1)	(2)	(3)	(4)	(5)	(6)
Het	3.157	1.216	1.348	3.926	3.585	7.407
	(7.555)	(5.195)	(4.673)	(5.638)	(5.528)	(4.793)
Period	−0.731	−0.080	−0.075	−0.081	−0.228	−0.525
	(0.458)	(0.302)	(0.270)	(0.278)	(0.316)	(0.391)
Salient		22.687***	13.383***	13.260***	14.769***	12.585*
		(3.049)	(4.654)	(4.888)	(5.232)	(6.426)
ContFM			0.274**	0.329**	0.295**	0.407**
			(0.109)	(0.149)	(0.149)	(0.177)
Het × ContFM				−0.081	−0.071	−0.207*
				(0.132)	(0.128)	(0.122)
Correct Belief					3.839	3.214
					(2.634)	(2.808)
LagThr						8.681**
						(3.479)
Constant	34.280***	13.809***	12.304***	10.729**	8.398*	4.546
	(4.798)	(3.778)	(3.711)	(4.767)	(4.838)	(4.794)
Obs.	190	190	190	190	190	171
No. Groups	19	19	19	19	19	19
Wald-χ^2	2.7130	81.40	96.06	110.9	152.7	328.9
Prob. > χ^2	0.258	0.000	0.000	0.000	0.000	0.000

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.3: Average contributions of followers by group over all periods. Generalized least squares estimates with random effects.

	(1)	(2)	(3)	(4)	(5)
Het	1.213	0.470	15.621	18.585*	19.836*
	(5.751)	(5.168)	(11.562)	(10.566)	(10.234)
Period	0.467	0.381	0.418	0.195	–0.202
	(0.415)	(0.361)	(0.381)	(0.363)	(0.307)
ContSalFM		0.380***	0.623***	0.608***	0.529***
		(0.139)	(0.164)	(0.160)	(0.146)
Het × ContSalFM			–0.391	–0.459*	–0.484**
			(0.254)	(0.234)	(0.234)
Correct Belief				6.438**	5.489*
				(3.115)	(2.959)
LagThr					8.742**
					(3.625)
Constant	32.964***	18.887***	9.479	6.291	7.000
	(4.655)	(6.278)	(7.499)	(6.839)	(5.283)
Obs.	150	150	150	150	133
No. Groups	18	18	18	18	18
Wald-χ^2	1.315	10.46	18.82	27.60	54.09
Prob. > χ^2	0.518	0.015	0.001	$4.35 \cdot 10^{-5}$	$7.08 \cdot 10^{-10}$

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.4: Average contributions of followers to the salient account by group over all periods. Generalized least squares estimates with random effects.

	(1)	(2)	(3)	(4)	(5)
Period	–0.342	0.010	0.079	–0.009	–0.435
	(0.507)	(0.315)	(0.372)	(0.458)	(0.281)
Selfish	21.416***	8.442*	18.216*	18.348*	16.924
	(4.625)	(4.811)	(11.060)	(10.041)	(13.201)
ContFM		0.393***	0.451***	0.471***	0.439***
		(0.068)	(0.061)	(0.069)	(0.070)
Selfish × ContFM			–0.244	–0.275	–0.347
			(0.252)	(0.220)	(0.302)
Correct Belief				2.146	0.096
				(3.602)	(3.792)
LagThr					12.464***
					(3.881)
Constant	19.448***	14.091***	11.684***	10.682***	12.011***
	(5.426)	(4.978)	(4.524)	(3.949)	(3.562)
Obs.	100	100	100	100	90
No. Groups	10	10	10	10	10
Wald-χ^2	40.20	41.27	94.86	85.51	677.6
Prob. > χ^2	$1.86 \cdot 10^{-9}$	$5.74 \cdot 10^{-9}$	0.000	0.000	0.000

Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A.5: Average contributions of followers in the heterogeneous treatment by group over all periods. Generalized least squares estimates with random effects.

A.3 Screenshot

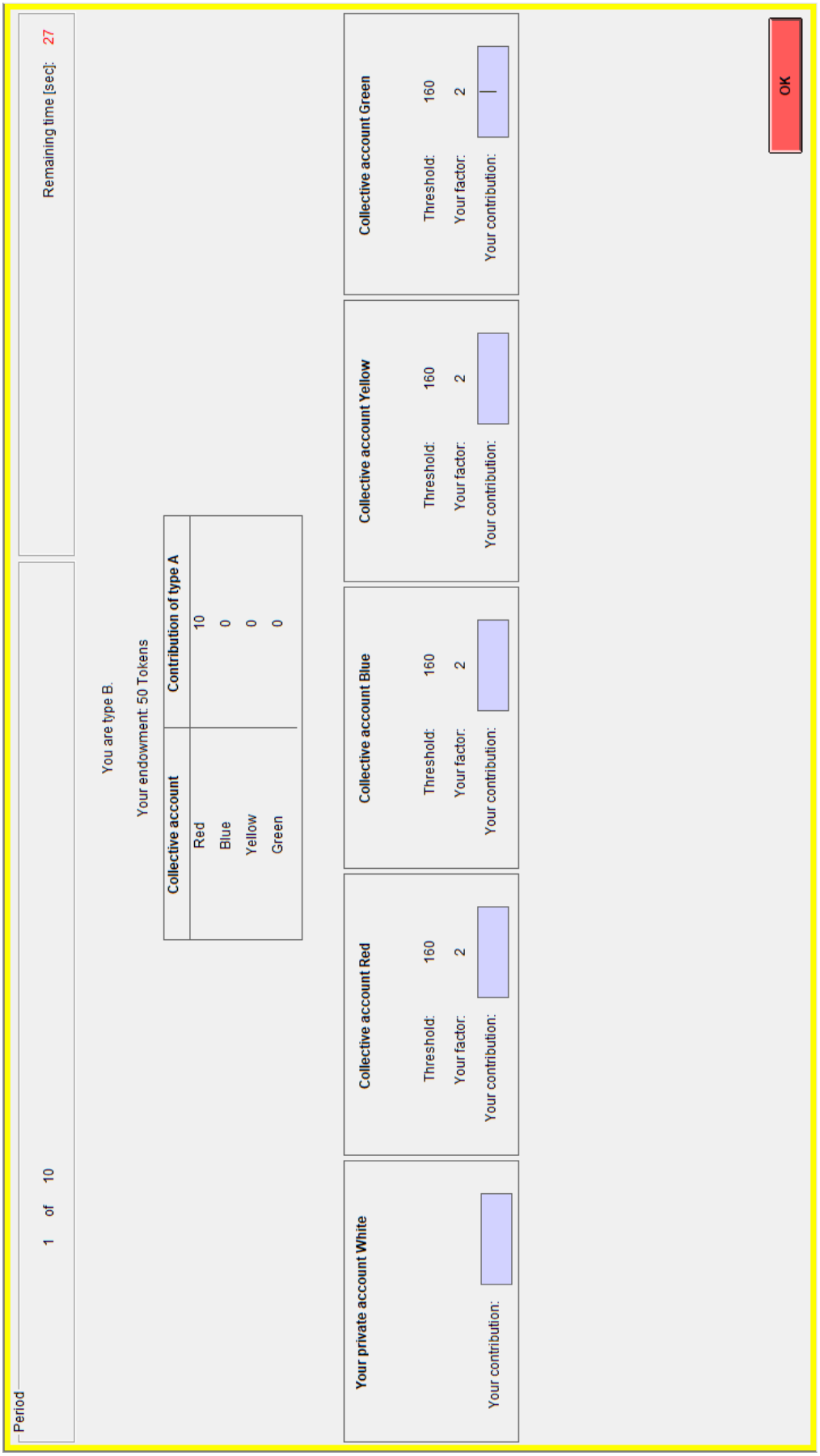


Figure A.3: Screenshot of followers' interface in the homogeneous treatment.

A.4 Instructions

Note: Instructions for the homogeneous and the heterogeneous treatment are identical for the largest part. Differences are highlighted.

Welcome to this Experiment!

You are participating in an economic experiment. All decisions are made privately, meaning that none of the other participants learns the identity of someone having made a certain decision. The payment is private information as well; none of the participants learns how much others have earned.

Please read the instructions carefully. If you have trouble understanding the instructions, please take a second look at it. If you still have questions, please give us a signal.

General Information

- This experiment consists of **three parts** and a **questionnaire**.
- You will receive instructions for each part before the respective part starts. You will receive longer instructions in paper form, whereas short instructions will appear on the screen. Please click OK only after you have read the instructions carefully and do not have any questions left.
- At the end of the experiment, you will get an overview of your results.
- Each part of the experiment affects your payoff. Therefore, please think carefully about every decision.
- In the experiment, we use the currency **"tokens"**. At the end of the experiment, your payoff will be converted into Euro; the conversion rate is 10 tokens = 1.00 Euro.
- Please stay at your seat at the end of the experiment and wait until we approach you to hand out your payoff.

Part I

The first part of the experiment comprises of **ten periods**.

You form a group with three other players. Groups are assigned randomly. You will not know the identity of the other players and none of the other players will be informed about your identity.

One player of your group is **type A** the other three players are **type B**. The assignment of the types is randomly. You will be informed about which player type you are prior to the beginning of the first period.

At the beginning of each period, you and the other players receive **50 tokens** each on an individual experimental account. You can split his amount between your **private account** WHITE

and **four collective accounts** RED, BLUE, YELLOW, and GREEN. For each token you allocate to your private account, you will receive one token at the end of the period. For each collective account your contribution and the contributions of the other players in your group will be added. If the total sum of contributions to of one community account reaches **at least 160 tokens**, the amount will be multiplied with a certain factor and distributed equally among the group members. **<HOM>** The table below provides an overview on the factors for each of the collective accounts.**</HOM>** **<HET>** Please note that the factors may differ for each player.**</HET>**

	Player 1 (type A)	Player 2 (type B)	Player 3 (type B)	Player 4 (type B)
$\text{Factor}_{\text{RED}}$	2	2	2	2
$\text{Factor}_{\text{BLUE}}$	2	2	2	2
$\text{Factor}_{\text{YELLOW}}$	2	2	2	2
$\text{Factor}_{\text{GREEN}}$	2	2	2	2

(*Note:* The table displays factors for the homogeneous treatment. In the heterogeneous treatment, factors on the diagonal are all equal to three.)

The factors which are relevant for you will be displayed on the screen. Throughout the experiment, the factors will not change. If there are less than 160 tokens on a collective account you receive no payoff from this collective account.

The payoff from your **private account** WHITE in a period (in tokens) is given by

$$\text{Payoff}_{\text{WHITE}} = \text{Contribution}_{\text{WHITE}}.$$

For the **collective accounts** your payoff in a period (in tokens) is calculated by:

- collective account RED

$$\text{Payoff}_{\text{RED}} = \begin{cases} \dfrac{\text{Factor}_{\text{RED}} \times \text{Sum}_{\text{RED}}}{4} & \text{if } \text{Sum}_{\text{RED}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{RED}} < 160 \end{cases}$$

- collective account BLUE

$$\text{Payoff}_{\text{BLUE}} = \begin{cases} \dfrac{\text{Factor}_{\text{BLUE}} \times \text{Sum}_{\text{BLUE}}}{4} & \text{if } \text{Sum}_{\text{BLUE}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{BLUE}} < 160 \end{cases}$$

- collective account YELLOW

$$\text{Payoff}_{\text{YELLOW}} = \begin{cases} \dfrac{\text{Factor}_{\text{YELLOW}} \times \text{Sum}_{\text{YELLOW}}}{4} & \text{if } \text{Sum}_{\text{YELLOW}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{YELLOW}} < 160 \end{cases}$$

- collective account GREEN

$$\text{Payoff}_{\text{GREEN}} = \begin{cases} \dfrac{\text{Factor}_{\text{GREEN}} \times \text{Sum}_{\text{GREEN}}}{4} & \text{if } \text{Sum}_{\text{GREEN}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{GREEN}} < 160 \end{cases}$$

Thus, your total payoff in a period (in tokens) is:

$$\begin{aligned} \text{Payoff}_{\text{TOTAL}} = \text{Contribution}_{\text{WHITE}} & \\ + & \begin{cases} \dfrac{\text{Factor}_{\text{RED}} \times \text{Sum}_{\text{RED}}}{4} & \text{if } \text{Sum}_{\text{RED}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{RED}} < 160 \end{cases} \\ + & \begin{cases} \dfrac{\text{Factor}_{\text{BLUE}} \times \text{Sum}_{\text{BLUE}}}{4} & \text{if } \text{Sum}_{\text{BLUE}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{BLUE}} < 160 \end{cases} \\ + & \begin{cases} \dfrac{\text{Factor}_{\text{YELLOW}} \times \text{Sum}_{\text{YELLOW}}}{4} & \text{if } \text{Sum}_{\text{YELLOW}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{YELLOW}} < 160 \end{cases} \\ + & \begin{cases} \dfrac{\text{Factor}_{\text{GREEN}} \times \text{Sum}_{\text{GREEN}}}{4} & \text{if } \text{Sum}_{\text{GREEN}} \geq 160 \\ 0 & \text{if } \text{Sum}_{\text{GREEN}} < 160 \end{cases} \end{aligned}$$

Sequence of a Period

Each period consists of **two stages**. At first, the type A player allocates his or her endowment to his or her private account and the collective accounts. At the same time, type B players state their beliefs as to which of the five accounts the type A player will contribute the largest share of his or her endowment. If the belief is correct, the corresponding type B player will receive 10 tokens. If the type A player contributes an equally large share of his or her endowment to two or more accounts, the corresponding type B player receives the payoff of 10 tokens if his or her belief matches any of these accounts.

Subsequently, type B players will be informed about the allocation of the type A player in the first stage and decide now individually upon the allocation of their endowment to their private account and the collective accounts. At the same time, the type A player states his or her belief as to which of the five accounts the type B players will on average contribute the largest share of their endowment. If the belief is correct, the type A player will receive 10 tokens. If type B players contribute on average an equally large share to two or more accounts, the type A player receives the payoff of 10 tokens if his or her belief matches any of these accounts.

Type A player As type A player, you have one minute to allocate your endowment to your private account and the collective accounts in the first stage. Please note that it is only possible to allocate your endowment to the individual account and the collective accounts in increments

of **5 tokens**. Please ensure also that the sum of your contributions to the private account and the collective accounts equals your endowment of **50 tokens**. In the second stage, you have one minute to state your belief as to which of the five accounts the type B players will on average contribute the largest share of their endowment. If your belief is correct, you will receive 10 tokens. If type B players contribute an equally large average share to two or more accounts, you receive a payoff of 10 tokens if your belief matches any of these accounts. Finally, you and the other players receive an overview on the result of the current period.

Type B player As type B player you have one minute to state your belief as to which of the five accounts the type A player will contribute the largest share of his or her endowment in the first stage. If your belief is correct, you will receive 10 tokens. If the the type A player contributions an equally large share of his or her endowment on two or more accounts, you will receive 10 tokens if one of these accounts matches your belief. In the second stage, you have one minute to allocate your endowment to your private account and the collective accounts. Please note that it is only possible to allocate your endowment on the private account and the collective accounts in increments of **5 tokens**. Please ensure also that the sum of your contributions to the private account and the collective accounts equals your endowment of **50 tokens**. Finally, you and the other players receive an overview on the result of the current period.

Determination of the Payment for Part I

At the end of the first part, i.e., after period 10, **one period will be randomly chosen** for each player. Only earnings from this particular round are relevant for the payoff from the first part of the experiment.

Control Questions and Practice Periods

Before the beginning of the payoff-relevant part, we would like to ask you to answer some control questions. You will then have the opportunity to familiarize with the procedure of the experiment during three practice periods.

If you still have any questions, please give us a sign, so the experiment leaders can come to you and answer your question. If there are no questions left, please click OK.

Part II

For the subsequent part, you will form a group with one randomly chosen player. You will not know about the identity of the other player and the other player will not know your identity.

In this part of the experiment, there are two types, type C and type D. Each type can choose between two options. Type C can choose between SQUARE and RECTANGLE; accordingly, type D can choose between RHOMBUS and TRAPEZOID.

For each type, the payoff depends on the own choice and the choice of the other type:

- If type C chooses SQUARE and type D chooses RHOMBUS, type C receives 7 tokens and type D receives 3 tokens.
- If type C chooses SQUARE and type D chooses TRAPEZOID, type C receives 3 tokens and type D receives 2 tokens.
- If type C chooses RECTANGLE and type D chooses RHOMBUS, type C receives 1 tokens and type D receives 2 tokens.
- If type C chooses RECTANGLE and type D chooses TRAPEZOID, type C receives 4 tokens and type D receives 6 tokens.

You and the other player decide on the choice of the options simultaneously. Your decision consists of two parts, an independent and a dependent decision.

- Independent decision: You decide as type C and choose from the options SQUARE and RECTANGLE. You have one minute for this decision. The other player makes this decision as well.
- Dependent decision: You make your decision as type D depending on the decision of the other type.
  - Assuming type C chooses SQUARE, do you choose RHOMBUS or TRAPEZOID?
  - Assuming type C chooses RECTANGLE, do you choose RHOMBUS or TRAPEZOID?

  You have 1 minute for both decisions. The other player makes this decision also.

After all decisions have been made, it is randomly determined whether you are type C or type D.

- For type C, the independent decision is implemented.
- For type D, the dependent decision is implemented.
  - If type C chose SQUARE, the corresponding decision of type D is implemented.
  - If type C chose RECTANGLE, the corresponding decision of type D is implemented.

If you still have any questions, please give us a sign, so the experiment leaders can come to you and answer your question. If there are no questions left, please click OK.

Part III

In this part of the experiment, you do not play with any of the other participants, so your decisions only influence your own payoff.

Please put yourself in the following Situation: You can choose between a safe payment and a lottery. The lottery generates a payment of 30 tokens with a probability of 50%. You receive nothing with the same probability.

Now assume you have to choose between the lottery and the safe payment. We will show you five different situations subsequently. The lottery is the same in each situation. The amount of the safe payment changes from situation to situation. For each decision you have 30 seconds.

After you have made all of your decisions, you will receive a payoff for this part of the experiment. If you chose the safe payment in the fifth situation, you receive this amount. If you choose the lottery in the fifth situation, then your payoff depends on the result of the lottery.

If you still have any questions, please give us a sign, so the experiment leaders can come to you and answer your question. If there are no questions left, please click OK.

Appendix B Supplementary Material for Section 3

B.1 Robustness Checks

	(1)	(2)	(3)	(4)	(5)
Treatment	0.957	1.020	0.968	0.964	0.926
	(0.254)	(0.278)	(0.231)	(0.230)	(0.312)
ESS	4.430***	3.864***	5.096***	5.007***	3.779***
	(0.251)	(0.931)	(1.111)	(1.170)	(0.676)
Total Activity		0.921	0.906		
		(0.114)	(0.111)		
Minutes Club				0.999	
				(0.001)	
Member		1.030	1.144	1.150	
		(0.731)	(0.678)	(0.617)	
Competitive		0.922	0.902	0.863	
		(0.337)	(0.320)	(0.320)	
Female					1.255
					(0.326)
Low SES					1.796
					(1.162)
Constant cut 1	0.056***	0.042***	0.038***	0.055***	0.068***
	(0.026)	(0.031)	(0.025)	(0.033)	(0.031)
Constant cut 2	0.164***	0.108***	0.112***	0.161***	0.201***
	(0.047)	(0.053)	(0.043)	(0.069)	(0.062)
Constant cut 3	0.877	0.603	0.612	0.867	1.092
	(0.092)	(0.326)	(0.259)	(0.429)	(0.222)
Constant cut 4	6.605***	4.309***	4.613***	6.520***	8.540***
	(0.887)	(2.406)	(2.138)	(2.813)	(2.586)
Constant cut 5	23.153***	18.051***	16.064***	22.889***	30.843***
	(7.048)	(8.492)	(6.054)	(8.630)	(9.598)
Obs.	130	118	130	130	130
No. Groups	9	9	9	9	9
Pseudo R^2	0.032	0.024	0.035	0.033	0.039

Note: Missing values of the variable Total Activity in column (3) are replaced with means by measurement, class, gender, and socio-economic status.
Exponentiated coefficients.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.1: Robustness check for contributions to the collective account in the first measurement. Ordered Logit estimates.

	(1)	(2)	(3)	(4)
Treatment	0.999	1.632	1.920	1.767
	(0.195)	(0.578)	(0.924)	(0.799)
2nd Measurement	0.624**	0.869	0.879	0.894
	(0.146)	(0.199)	(0.191)	(0.222)
3rd Measurement	0.812	0.961	0.804	0.983
	(0.263)	(0.345)	(0.297)	(0.342)
Treatment × 2nd Measurement	0.941	0.525	0.433**	0.466*
	(0.492)	(0.260)	(0.153)	(0.200)
Treatment × 3rd Measurement	0.517	0.250	0.223*	0.226*
	(0.427)	(0.213)	(0.190)	(0.180)
Low SES		2.921**	2.942**	3.088**
	(1.270)	(1.360)	(1.473)	
Treatment × Low SES		0.119**	0.080**	0.092**
		(0.110)	(0.081)	(0.093)
2nd Measurement × Low SES		0.189**	0.234*	0.203*
		(0.153)	(0.196)	(0.185)
3rd Measurment × Low SES		0.414**	0.587	0.456*
		(0.180)	(0.329)	(0.213)
Treatment × 2nd Measurement × Low SES		13.446**	13.580**	14.561**
		(14.656)	(15.802)	(17.777)
Treatment × 3rd Measurement × Low SES		16.978***	17.575***	18.469***
		(17.240)	(19.204)	(18.512)
Female			2.060***	1.975***
			(0.484)	(0.358)
Total Activity			0.894	0.866**
			(0.074)	(0.060)
Member			1.010	1.328
			(0.555)	(0.714)
Competitive			1.141	1.102
			(0.306)	(0.295)
Constant cut 1	0.067***	0.078***	0.070***	0.074***
	(0.021)	(0.028)	(0.036)	(0.033)
Constant cut 2	0.153***	0.182***	0.158***	0.171***
	(0.049)	(0.066)	(0.071)	(0.067)
Constant cut 3	0.952	1.190	1.111	1.138
	(0.099)	(0.222)	(0.496)	(0.425)
Constant cut 4	5.980***	7.759***	6.998***	7.509***
	(1.023)	(1.472)	(3.637)	(3.512)
Constant cut 5	19.083***	25.178***	21.203***	24.309***
	(5.669)	(5.512)	(12.587)	(13.100)
Obs.	318	318	296	318
No. Individuals	106	106	106	106
No. Groups	8	8	8	8

Note: Missing values of the variable Total Activity in column (4) are replaced with means by measurement, class, gender, and socio-economic status.
Exponentiated coefficients.
Clustered standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.2: Robustness check for contributions to the collective account in all three measurements. Ordered Logit estimates with random effects.

B.2 Instructions

General Instructions

- Dear participants, welcome to this session. My name is <name of experimenter 1> and together with my colleague <name of experimenter 2> and your teacher <name of teacher> I will guide you through the following 90 minutes.
- Today we want to play different games with you and the other children of your class. For your participation you will receive a payment, depending on the decisions you and the other children of your class will make.
- In each game, you will make decisions which either concern just yourself or other children of your class. The decision you will have to make will be explained in detail at the beginning of each game. In any case, you will receive a minimum amount of 2.00 Euro.
- At the end of the 90 minutes we will pay out one of the games. That means, we will convert all tokens you received in that game into Euro and pay them out in addition to the minimum amount. Therefore, it is worth it to think carefully about your decision in every game.
- The number of the game we will pay out is in this sealed envelope. After we have played the last game, we will ask your teacher to open this envelope.
- Your decisions are confidential, that means neither the other children of your class nor your teacher will learn about your decisions. You will not learn about the decisions of the other children of your class either. Therefore, it is important that you do not talk about your decisions with the person sitting next to you.
- The payment at the end of the games is also confidential. That means the other children of your class will not learn which payment you receive. You will not learn anything about the payments of the other children of your class either.
- In every game, we use the currency token. At the end, tokens will be converted into Euro. 75 tokens equal 1.00 Euro. Your teacher will do the payment. As stated before, you will receive a minimum amount of 2.00 Euro for your participation.
- At your table you find a card with a number. This number is important in order to ensure that you will receive the correct number of tokens in Euro. Please check on every sheet you receive that the number in the right upper corner corresponds to the number of your card. Only if the numbers are the same we are able to identify your sheet and make sure you receive your payment.
- Please keep this card until you have received your payment. After the last game we will put your payment in an envelope marked with your number. Your teacher will give you this envelope in exchange for the card with your number.

- Before starting a game, we will hand out a new sheet to you. On this sheet you will find a description of the game which we will also read out aloud. If you have any questions, please raise your hand.
- Do you have any questions? If not, we will begin.

Game 3

In this game you play together with three other children of your class who will be selected randomly after all sheets have been collected.

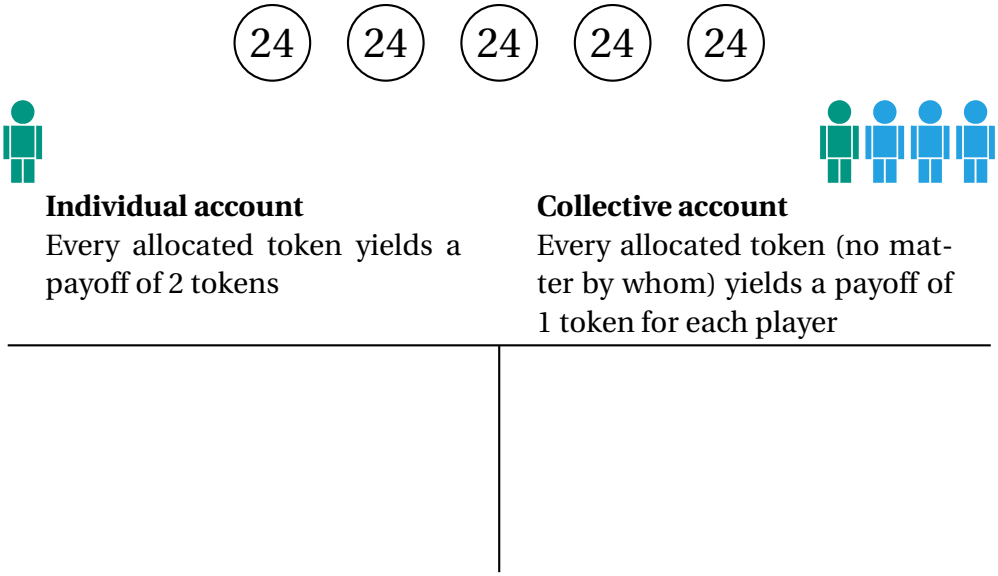
You and the other children of your class receive 120 tokens. You can allocate these 120 tokens to your own private account or to a collective account in intervals of 24 tokens. For every token you allocate to your own account, you receive 2 tokens in return.

For every token which is allocated to the collective account, you receive 1 token. At the end, we will sum up the total number of tokens which have been allocated to the collective account and distribute this amount to each player. Since everyone can contribute to the shared account, you benefit if other children of your group allocate tokens to the collective account and they benefit from your allocation.

Do you have any questions? If not, we will continue.

How many tokens do you want to allocate to your own account and how many tokens do you want to allocate to the collective account?

In the following illustration, one coin is equal to 24 tokens. Please draw the number of coins you want to allocate to your own account on the left side of the diagram. Likewise, you can draw the number of coins you want to allocate to the collective account on the right side of the diagram.



Please ensure that you allocated exactly 5 coins.

Game 5

Please do not turn the sheet until we made the start signal.

In this game you can encrypt words into a secret code. The letters have to be exchanged by other letters in two subsequent steps.

The tables below show the replacement for each letter:

1st step

Replace	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
with	K	L	W	U	B	F	N	M	O	E	P	J	A	I	S	D	V	X	R	G	Y	H	Z	C	T	Q

2nd step

Replace	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
with	P	T	F	E	S	D	U	O	G	V	B	Z	H	N	M	C	R	Q	Y	I	X	W	J	K	L	A

You have two minutes to encrypt the words on the back of the sheet. Before we start, you can choose one out of two payoff schemes:

1. For every correctly encoded letter you receive 3 tokens with a maximum of 300 tokens.
2. We compare your result with the result of another child of your class who will be selected randomly. The one with the higher number of correctly encoded letters receives 600 tokens. The other one receives zero tokens. In case of a tie we will determine randomly who receives 600 tokens and who receives zero tokens.

Do you have any questions? If not, we will continue.

Which payoff scheme do you choose? Please tick one box.

☐ 3 tokens for every correctly encoded letter

☐ 600 tokens for you if you encode more letters correctly than the other randomly-selected child of your class and zero tokens otherwise

On the back of this sheet you find more words which are to be encrypted into the secret code. Please do not turn the play bar until we made the start signal. After we made the start signal, you have 2 minutes to encrypt as many words two times as possible. The moment we make the end signal, please cease writing immediately.

B.3 Approval by the Superintendent

Eingegangen
21. [illegible]
KIT - IFSS

Baden-Württemberg
REGIERUNGSPRÄSIDIUM KARLSRUHE
ABTEILUNG 7 - SCHULE UND BILDUNG

Regierungspräsidium Karlsruhe · 76247 Karlsruhe

Herrn
Prof. Dr. Alexander Woll
Institut für Sport und Sportwissenschaft
Karlsruher Institut für Technologie
Engler-Bunte-Ring 15
76131 Karlsruhe

Karlsruhe 16.11.2016
Name Michael Huber
Durchwahl 0721 926-4162
Aktenzeichen 71c2-6499.25
(Bitte bei Antwort angeben)

Antrag auf Genehmigung einer wissenschaftlichen Erhebung an sechs Karlsruher Gymnasien im Rahmen der Studie MoviGen
Ihr Schreiben vom 08.11.2016

Sehr geehrter Herr Prof. Dr. Woll,

das Regierungspräsidium Karlsruhe hat keine Einwände gegen die Durchführung der geplanten wissenschaftlichen Erhebung im Rahmen der Studie MoviGen an sechs Karlsruher Gymnasien (Lessing-Gymnasium, Goethe-Gymnasium, Max-Planck-Gymnasium, Fichte-Gymnasium, Humboldt-Gymnasium, Otto-Hahn-Gymnasium).

Die formelle Genehmigung ist jedoch beim geschäftsführenden Schulleiter der Karlsruher Gymnasien einzuholen.
Gemäß Nr. 4.1 der Verwaltungsvorschrift des Kultusministerium vom 21.09.2002 (K.u.U. S. 309/2002) über Werbung, Wettbewerbe und Erhebungen in Schulen, entscheidet der geschäftsführende Schulleiter im Benehmen mit den betroffenen Schulleitern über die Genehmigung einer beantragten Erhebung an mehreren Schulen, wenn diese sich alle auf dem Gebiet eines Schulträgers befinden. Dies ist hier der Fall; alle sechs genannten Gymnasien befinden sich in Trägerschaft der Stadt Karlsruhe.

Dienstgebäude Hebelstraße 2 · 76133 Karlsruhe · Telefon 0721 926 0 · Fax 0721 93340270
abteilung7@rpk.bwl.de · www.rp.baden-wuerttemberg.de · www.service-bw.de
ÖPNV Haltestelle Marktplatz · Parkmöglichkeit Parkhaus Kreuzstraße · Behindertenparkplätze im Hof

Wir bitten darum, den Antrag auf Genehmigung der geplanten wissenschaftlichen Erhebung direkt dort zu stellen. Eine Weiterleitung Ihres Antrags durch uns kann leider nicht erfolgen.

Mit freundlichen Grüßen

Michael Huber

B.4 Approval by the Board of Ethics

Karlsruher Institut für Technologie

Vizepräsidentin für Personal und Recht

Dr. Elke Luise Barnstedt

Kaiserstraße 12
76131 Karlsruhe

Telefon: 0721-608-43900
Fax: 0721-608-42835
E-Mail: elke.barnstedt@kit.edu

Datum: 11.10.2016 / rd

Eingegangen
17. Okt. 2016
KIT - IFSS

Herrn
Prof. Dr. Alexander Woll
Institut für Sport und Sportwissenschaften
Karlsruher Institut für Technologie

Ihr Antrag an die Ethikkommission des KIT vom 15. September 2016
Forschungsvorhaben: „movigen: Kompetenzen im Schulsport"

Sehr geehrter Herr Professor Woll,

die Ethikkommission hat in ihrer Sitzung am 29. September 2016 einstimmig das folgende Votum zu Ihrem Antrag für das Forschungsvorhaben: „movigen: Kompetenzen im Schulsport" erteilt:

Es bestehen keine ethischen Bedenken an der Zulässigkeit des Forschungsvorhabens „movigen: Kompetenzen im Schulsport".

Die Ethikkommission begrüßt Ihren Vorschlag, eine Reflexionsphase in den Ablauf der Studie zu integrieren, um den Schülern und Schülerinnen die Möglichkeit zu bieten, über die Erfahrungen, welche sie insbesondere im Rahmen der Spiele gesammelt haben, mit den Interventionsleitern und Interventionsleiterinnen sowie Lehrern und Lehrerinnen zu sprechen.

Änderungen des Forschungsvorhabens sowie alle schwerwiegenden oder unerwarteten unerwünschten Ereignisse vor oder während der Durchführung des Forschungsvorhabens, die die Sicherheit der Teilnehmer/-innen oder die Durchführung des Forschungsvorhabens beeinträchtigen, sind der Ethikkommission unverzüglich mitzuteilen.

Sie werden ferner gebeten, die am KIT geltenden Regeln zur Sicherung guter wissenschaftlicher Praxis (Anlage) und hier insbesondere die unter Punkt 6 enthaltene Regelung zur Aufbewahrung und Nutzung von Primärdaten zu beachten.

Das Votum der Ethikkommission entbindet die für das Forschungsvorhaben verantwortliche/n Person/en nicht von der Verantwortung für die Durchführung des Forschungsvorhabens.

Mit freundlichen Grüßen

E. L. Barnstedt

Dr. Elke Luise Barnstedt
Vorsitzende der Ethikkommission

Anlage

Appendix C Supplementary Material for Section 4

C.1 Sabotage by Winning Probability

Winning probability	Equality of groups*	pairwise comparisons†		
		INFO SELF_{100} OPP_0 vs. INFO SELF_{70} OPP_0	INFO SELF_{100} OPP_0 vs. INFO SELF_{100} OPP_{70}	INFO SELF_{70} OPP_0 vs. INFO SELF_{100} OPP_{70}
10%	$p = 0.3875$	–	–	–
20%	$p = 0.6287$	–	–	–
30%	$p = 0.3759$	–	–	–
40%	$p = 0.3831$	–	–	–
50%	$p = 0.4942$	–	–	–
60%	$p = 0.0349$	$p = 0.0336$	$p = 0.4995$	$p = 0.0417$
70%	$p = 0.2866$	–	–	–
80%	$p = 0.2943$	–	–	–
90%	$p = 0.0563$	$p = 0.1417$	$p = 0.1502$	$p = 0.0246$

* Kruskal-Wallis H-test.
† Pairwise Dunn test with Holm-Bonferroni adjustment.

Table C.1: Comparison of sabotage between treatments by winning probability in period 1.

Winning probability	Equality of groups*	pairwise comparisons†		
		INFO SELF_{100} OPP_0 vs. INFO SELF_{70} OPP_0	INFO SELF_{100} OPP_0 vs. INFO SELF_{100} OPP_{70}	INFO SELF_{70} OPP_0 vs. INFO SELF_{100} OPP_{70}
10%	$p = 0.2585$	–	–	–
20%	$p = 0.3272$	–	–	–
30%	$p = 0.0350$	$p = 0.0788$	$p = 0.0144$	$p = 0.1469$
40%	$p = 0.9309$	–	–	–
50%	$p = 0.8677$	–	–	–
60%	$p = 0.3768$	–	–	–
70%	$p = 0.5804$	–	–	–
80%	$p = 0.4681$	–	–	–
90%	$p = 0.2525$	–	–	–

* Kruskal-Wallis H-test.
† Pairwise Dunn test with Holm-Bonferroni adjustment.

Table C.2: Comparison of sabotage between treatments by winning probability in period 2.

Winning probability	comparison within treatments: period 1 vs. period 2*		
	INFO SELF_{100} OPP_0	INFO SELF_{70} OPP_0	INFO SELF_{100} OPP_{70}
10%	—	$p = 0.3173$	–
20%	$p = 0.1536$	$p = 0.8527$	$p = 0.2498$
30%	$p = 0.0891$	$p = 0.2622$	$p = 0.0138$
40%	$p = 0.2922$	$p = 0.7455$	$p = 0.5442$
50%	$p = 0.2863$	$p = 0.2278$	$p = 0.0460$
60%	$p = 0.2495$	$p = 0.4995$	$p = 0.1636$
70%	$p = 0.3003$	$p = 0.1055$	$p = 0.2326$
80%	$p = 0.2775$	$p = 0.3173$	$p = 1.0000$
90%	$p = 0.3173$	–	–

* Two-sided Wilcoxon signed-rank test.

Table C.3: Comparison of sabotage within treatments by winning probability between period 1 to period 2 (only individuals with same winning probability in both periods).

The figures below provide an overview of sabotage by winning probability and treatment. Note that the variable Distance depicted on the horizontal axis indicates individuals' winning probabilities relative to the reference category with a balanced competition, i.e., a situation in which either individual within a group has a winning probability of 50%. The values of this variable range from −4 for individuals with a winning probability of 10% to 4 for individuals with a winning probability of 90%, respectively.

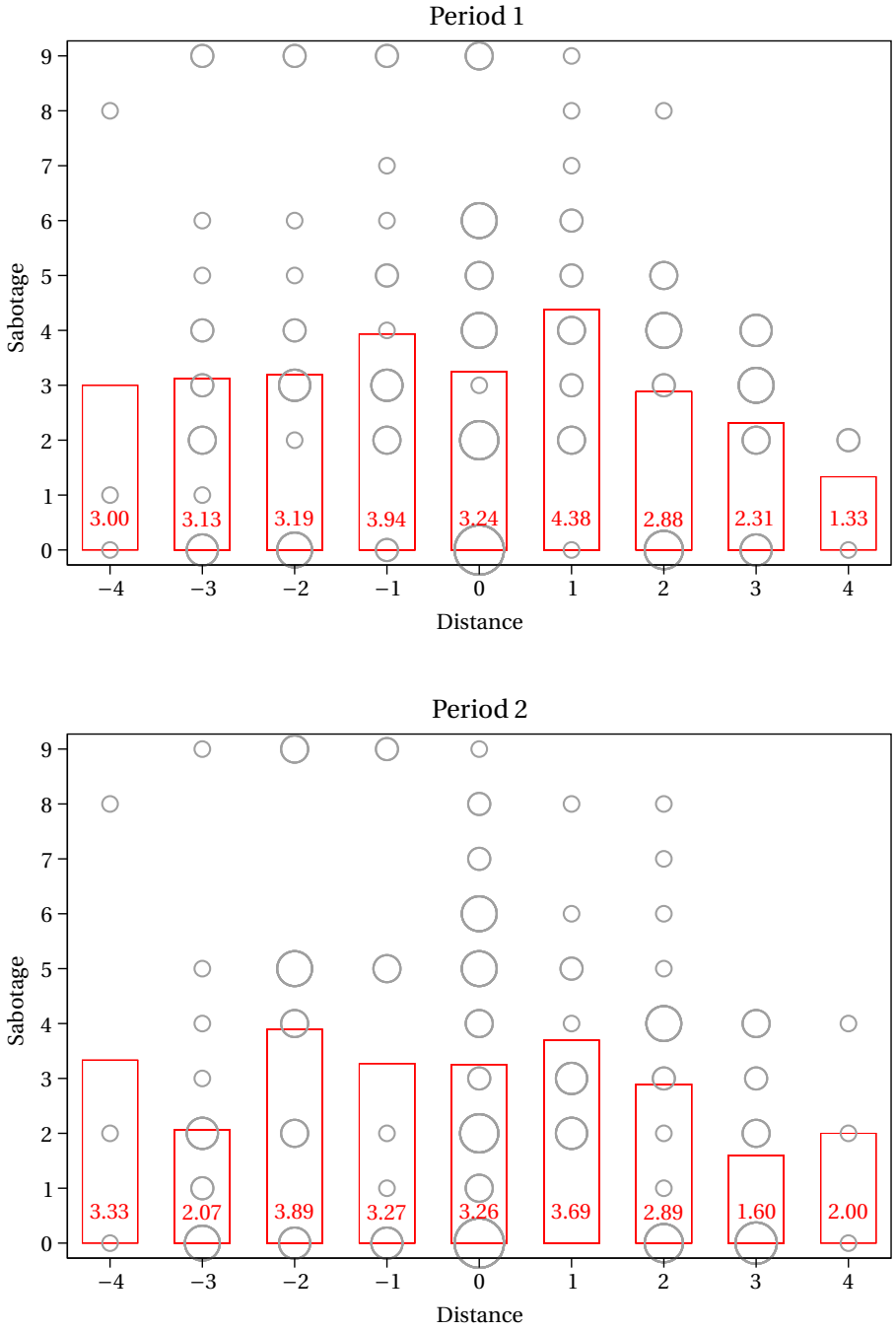


Figure C.1: Sabotage in the INFO SELF_{100} OPP_0 treatment. The size of the bubbles indicates the frequency of distinct observations. The bars indicate mean sabotage levels for the respective categories.

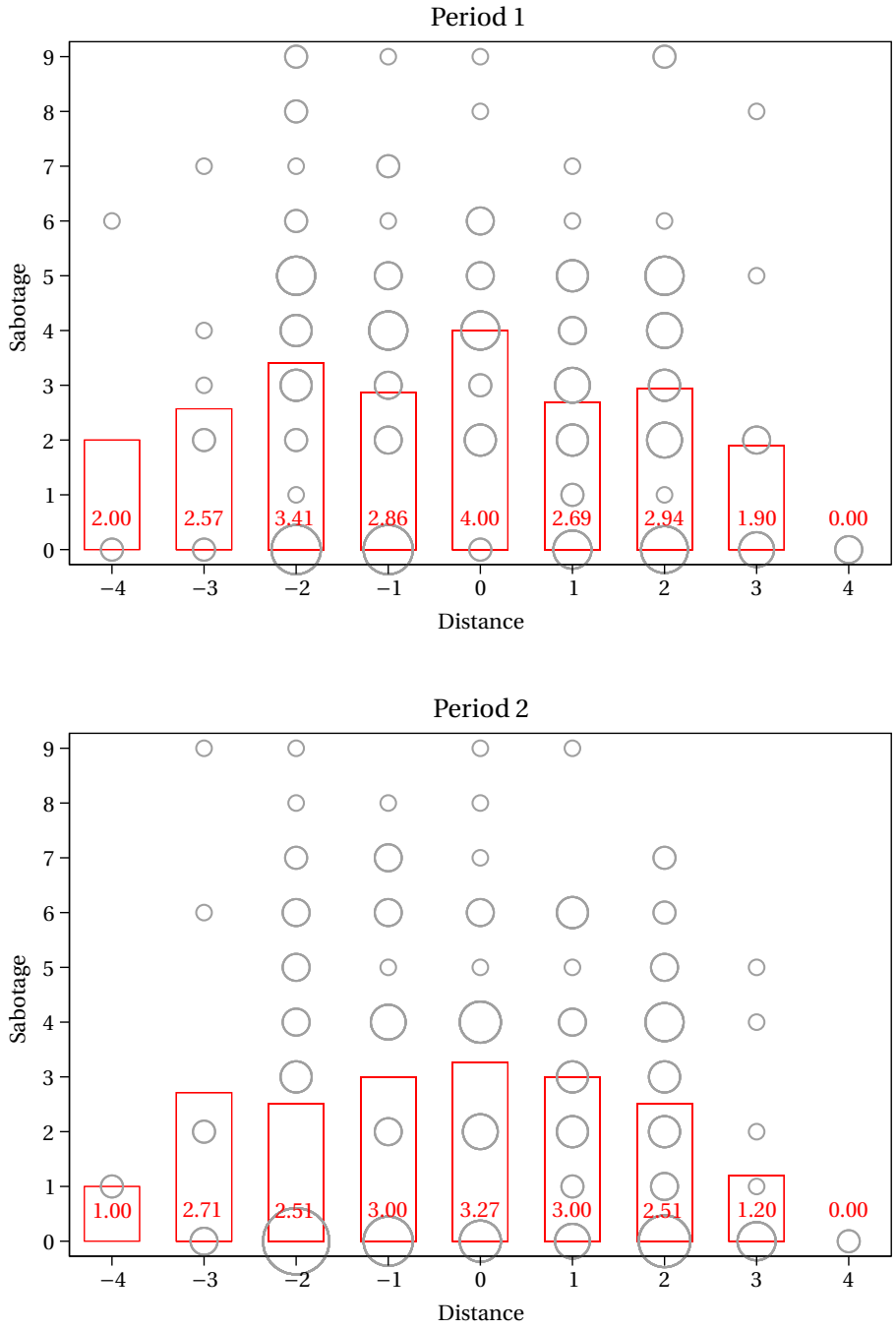


Figure C.2: Sabotage in the INFO SELF$_{70}$ OPP$_0$ treatment. The size of the bubbles indicates the frequency of distinct observations. The bars indicate mean sabotage levels for the respective categories.

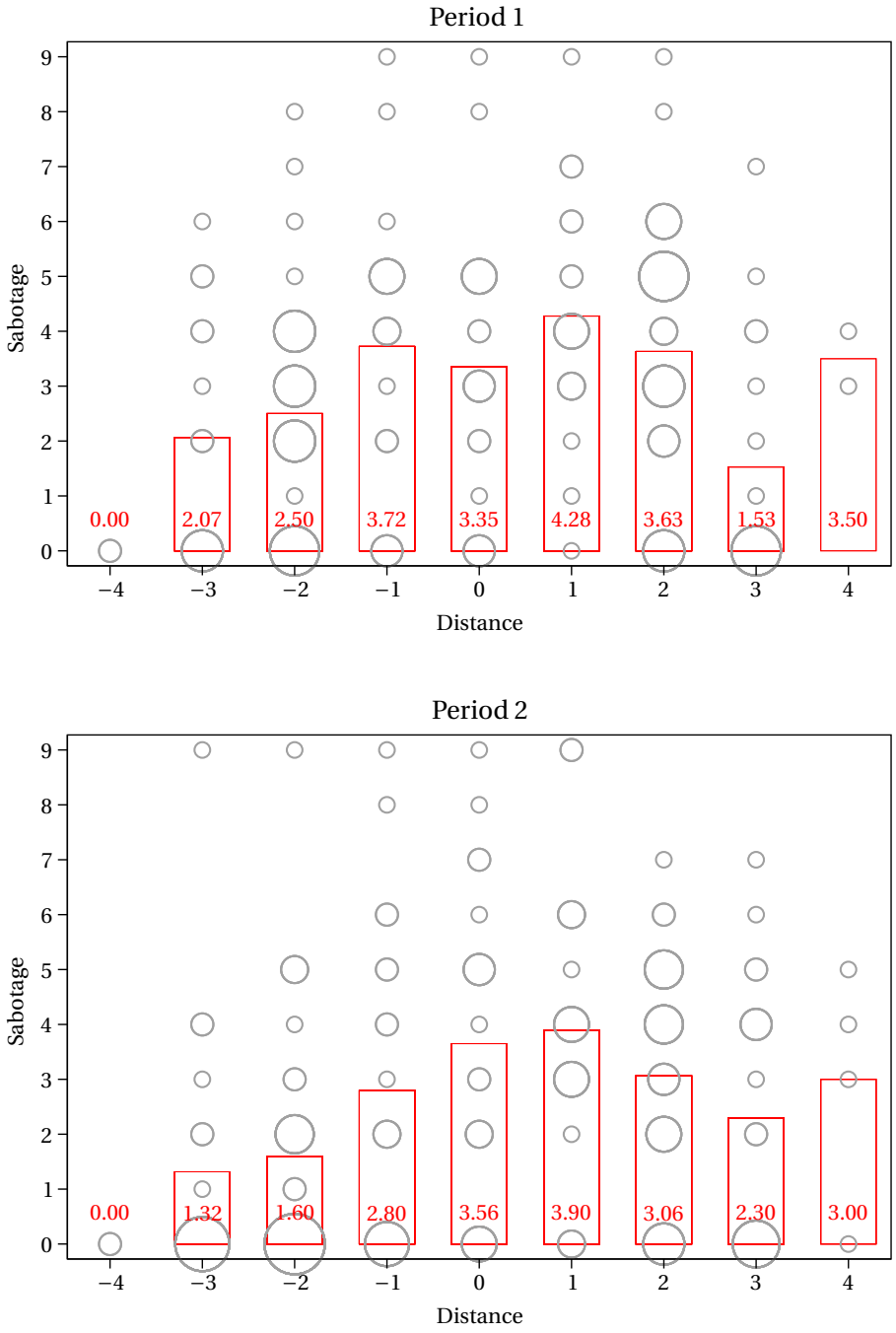


Figure C.3: Sabotage in the INFO SELF$_{100}$ OPP$_{70}$ treatment. The size of the bubbles indicates the frequency of distinct observations. The bars indicate mean sabotage levels for the respective categories.

C.2 Social Value Orientation

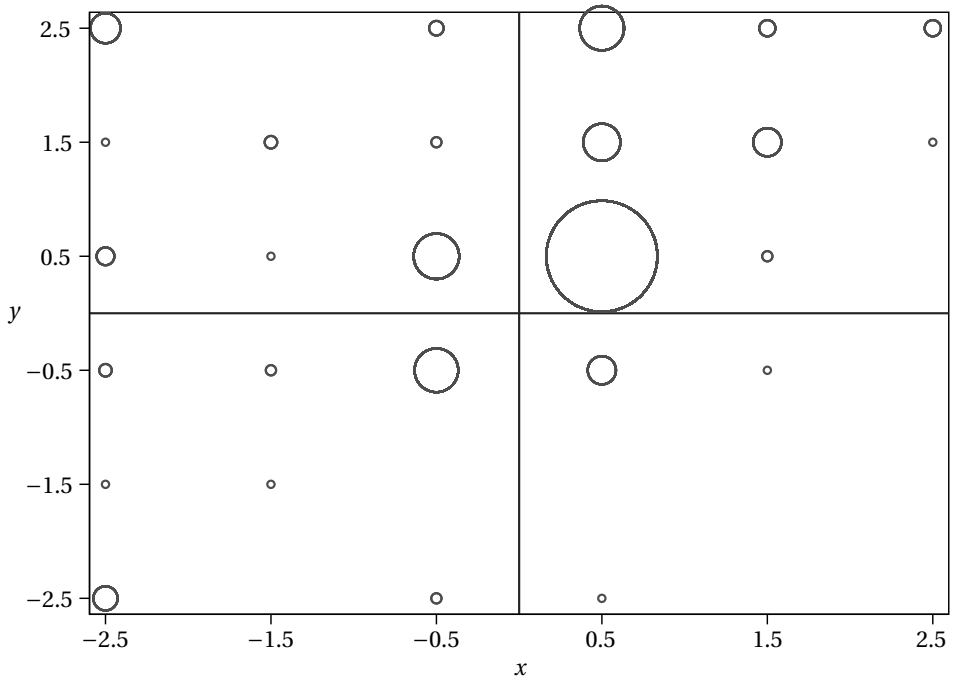


Figure C.4: Distributional types according to Kerschbamer (2015). Individuals with coordinates $-0.5 \leq x \leq 0.5$ and $-0.5 \leq y \leq 0.5$ are categorized as selfish. The remaining individuals in quadrants 1–4 are deemed altruistic, inequality averse, spiteful, and equality averse. The size of the bubbles indicates the frequency of distinct observations.

	Trailing individuals				Leading individuals			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INFO SELF_{70} OPP_0	−0.726	−0.969	−0.627	−0.653	−0.694	−1.033*	−1.136**	−1.491**
	(0.770)	(0.775)	(0.715)	(0.950)	(0.591)	(0.575)	(0.574)	(0.699)
INFO SELF_{100} OPP_{70}	−1.205	−1.289*	−1.033	−1.417	0.235	0.130	0.154	0.254
	(0.745)	(0.734)	(0.678)	(0.864)	(0.572)	(0.537)	(0.566)	(0.666)
Absolute distance		−0.746**	−0.725**	−0.728**		−1.255***	−1.293***	−1.314***
		(0.369)	(0.358)	(0.356)		(0.268)	(0.283)	(0.281)
Risk			0.328**	0.337**			0.445***	0.444***
			(0.149)	(0.151)			(0.111)	(0.110)
Econ			0.357	0.361			−0.552	−0.596
			(0.572)	(0.576)			(0.480)	(0.481)
Female			2.116***	1.797			0.832*	0.564
			(0.671)	(1.190)			(0.489)	(0.957)
INFO SELF_{70} OPP_0 × Female				0.036				0.840
				(1.411)				(1.182)
INFO SELF_{70} OPP_0 × Female				0.839				−0.120
				(1.381)				(1.133)
Constant	3.077***	4.651***	4.062*	4.216*	2.691***	5.351***	3.860*	4.029*
	(0.585)	(0.935)	(2.356)	(2.380)	(0.435)	(0.703)	(2.277)	(2.236)
Competitiveness	No	No	Yes	Yes	No	No	Yes	Yes
Self-consciousness	No	No	Yes	Yes	No	No	Yes	Yes
Social value orientation	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	195	195	195	195	199	199	199	199
Pseudo R^2	0.003	0.009	0.038	0.039	0.004	0.027	0.055	0.056

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.4: Sabotage in period 1 by trailing and leading individuals. Tobit estimates (lower limit: 0, upper limit: 9; trailing individuals: 59 observations left-censored, 10 observations right-censored; leading individuals: 53 observations left-censored, 5 observations right-censored).

	Trailing individuals				Leading individuals			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
INFO SELF_{70} OPP_0	−1.309	−1.694	−1.216	−1.629	−0.361	−0.707	−0.805	−1.073
	(1.008)	(1.038)	(0.947)	(1.319)	(0.595)	(0.564)	(0.545)	(0.682)
INFO SELF_{100} OPP_{70}	−2.657***	−2.842***	−2.351**	−2.035	0.507	0.464	0.714	0.522
	(1.014)	(1.008)	(0.953)	(1.301)	(0.576)	(0.551)	(0.557)	(0.685)
Absolute distance		−0.922*	−1.069**	−1.048**		−1.163***	−1.197***	−1.200***
		(0.474)	(0.453)	(0.451)		(0.285)	(0.286)	(0.290)
Risk			0.251	0.243			0.530***	0.528***
			(0.196)	(0.196)			(0.114)	(0.113)
Econ			0.708	0.682			−0.703	−0.739
			(0.745)	(0.749)			(0.472)	(0.474)
Female			3.045***	2.912**			0.354	−0.217
			(0.857)	(1.467)			(0.490)	(0.967)
INFO SELF_{70} OPP_0 × Female				0.960				0.789
				(1.845)				(1.161)
INFO SELF_{100} OPP_{70} × Female				−0.662				0.656
				(1.834)				(1.153)
Constant	2.487***	4.544***	3.673	3.810	2.137***	4.635***	0.965	1.057
	(0.751)	(1.347)	(3.407)	(3.525)	(0.449)	(0.676)	(2.103)	(2.114)
Competitiveness	No	No	Yes	Yes	No	No	Yes	Yes
Self-consciousness	No	No	Yes	Yes	No	No	Yes	Yes
Social value orientation	No	No	Yes	Yes	No	No	Yes	Yes
Obs.	188	188	188	188	196	196	196	196
Pseudo R^2	0.010	0.015	0.051	0.052	0.003	0.025	0.059	0.059

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.5: Sabotage in period 2 by trailing and leading individuals. Tobit estimates (lower limit: 0, upper limit: 9; trailing individuals: 82 observations left-censored, 11 observations right-censored; leading individuals: 58 observations left-censored, 3 observations right-censored).

	(1)	(2)	(3)	(4)	(5)	(6)
2nd stage						
INFO SELF_{70} OPP_0	−1.265	−1.193	−1.125	−1.082	−0.510	0.620
	(1.233)	(1.144)	(0.825)	(0.827)	(0.635)	(0.822)
INFO SELF_{100} OPP_{70}	−0.017	−0.012	0.022	0.020	0.263	1.407**
	(0.969)	(0.902)	(0.678)	(0.673)	(0.504)	(0.599)
Tie		−1.850*	−1.846***	−1.923***	−1.269**	−1.296**
		(0.965)	(0.691)	(0.682)	(0.598)	(0.551)
Leading individual		−0.868	−0.638	−0.639	−0.803*	−0.826*
		(0.697)	(0.481)	(0.481)	(0.444)	(0.440)
Sabotage			0.536***	0.527***	0.420***	0.419***
			(0.137)	(0.146)	(0.096)	(0.094)
Belief				0.007	0.011	0.009
				(0.009)	(0.007)	(0.007)
Risk					0.214**	0.179*
					(0.104)	(0.099)
Econ					−0.460	−0.322
					(0.429)	(0.408)
Female					0.817*	2.517***
					(0.428)	(0.879)
INFO SELF_{70} OPP_0 × Female						−2.056*
						(1.232)
INFO SELF_{100} OPP_{70} × Female						−2.142**
						(1.000)
Constant	1.841**	2.475***	0.681	0.337	1.577	1.168
	(0.838)	(0.849)	(0.847)	(0.871)	(1.644)	(1.653)
Competitiveness	No	No	No	No	Yes	Yes
Self-consciousness	No	No	No	No	Yes	Yes
Social value orientation	No	No	No	No	Yes	Yes
1st stage						
INFO SELF_{70} OPP_0	0.109	0.109	0.109	0.109	0.109	0.109
	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)	(0.175)
INFO SELF_{100} OPP_{70}	0.836***	0.836***	0.836***	0.836***	0.836***	0.836***
	(0.165)	(0.165)	(0.165)	(0.165)	(0.165)	(0.165)
Constant	−1.049***	−1.049***	−1.049***	−1.049***	−1.049***	−1.049***
	(0.132)	(0.132)	(0.132)	(0.132)	(0.132)	(0.132)
Obs.	469	469	469	469	469	469
Pseudo R^2	0.039	0.043	0.071	0.072	0.099	0.102

Robust standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table C.6: Willingness to pay to prevent information disclosure in period 2. Truncated normal hurdle model (threshold: 0).

C.4 Instructions

Welcome to this Experiment!

You are participating in an economic experiment. All decisions are made privately, meaning that none of the other participants learns the identity of someone having made a certain decision. The payment is private information as well; none of the participants learns how much others have earned. Please read the instructions carefully. If you have trouble understanding the instructions, please take a second look at it. If you still have questions, please give us a signal.

General Information

- This experiment consists of a **main part** with a working stage and two decision-making stages. A short second part and a questionnaire will follow. You will receive a compensation for these two parts as well. The details of the payoff in the second part are given directly on the screen. For the questionnaire, you receive an additional payoff of 2.00 Euro.
- At first, you will receive instructions for the working stage and the first decision-making stage. You will receive the instructions for the second decision-making stage after the first decision-making stage. You can read short instructions and the questionnaire directly on the screen. Please click "Next" after you have read the instructions completely and do not have any questions.
- At the end of the experiment, **one of the two decision-making stages will be drawn randomly**. This decision-making period determines your payoff from the main part.
- In each decision-making stage, **two players form a group**. In each decision-making stage, you will be randomly assigned to a **new player** whose identity will not be revealed.
- In each decision-making stage, one player from the group will receive a **high payoff of 75 tokens** while the other player will receive a **lower payoff of 25 tokens**. In the instructions below you can find details on how this payoff is determined.
- At the end of the experiment, you will get an overview of your results.
- Each of your decisions can affect your payoff, so please think carefully about every decision.
- In the experiment, we use the currency "token". At the end of the experiment, your payoff will be converted into Euro. The conversion rate is **5 tokens** = **1.00 Euro**.
- Please stay at your seat at the end of the experiment and wait until we approach you to hand out your payoff.

Working Stage

- The main part starts with a working stage in which each player works on their own.

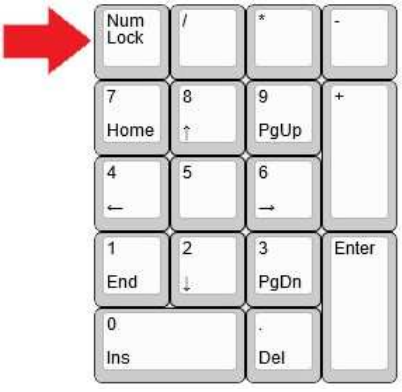
- Your task is to encode words to numbers. You have to replace each letter of a word with the numbers given in table at the end of the instructions.
  - Example: The word “camel” appears on the screen. According to the table, “C” = 80, “A” = 92, “M” = 35, “E” = 55, and “L” = 20, so the code for “camel” is: 8092355520.
  - For each letter, you have to enter the assigned number in a separate box. You can go from box to box by using the tab key. The tab key can be found on the keyboard left of the letter “Q”.

    Position of the tab key on the keyboard:

  - The easiest way to enter the numbers is by using the numeric keypad on the right side of the keyboard. The numeric keypad should be activated automatically. If the input does not work, press the “Num Lock” key, which is located in the upper left corner of the numeric keypad. Before you begin to code, please select the (first) input field with the mouse; otherwise your input will not be recorded.

    Position of the Num-Lock key on the keyboard:

- Five-digit words will appear. You will earn one point for each correctly encoded letter. Please note that incorrectly encoded letters will lead to a deduction of 0.5 points. The sum of the points is the **obtained score**.
- When you have entered the code and pressed “OK”, you will automatically receive a new word for encoding. You do not have the option to return to previous words.
- **The processing time is 10 minutes**. After that, the processing time stops automatically.
- At the end of the working stage, you will be informed about your score.

At first, each player encodes words into numbers for **10 minutes**. All players receive identical words in the same order.

Decision-Making Stage 1: Course of Action

- Below, you will find information about the first decision-making stage.
- After completing this stage, you will get the instruction for the second decision-making stage.
- In this decision-making stage, **two players form a group** and enter a competition.
- If this decision-making stage is drawn, one player of the group will receive a **high payoff of 75 tokens** while the other player will receive a **lower payoff of 25 tokens**. In the instructions below you can find details on how this payment is determined.
- First, the difference between the obtained scores of both players from the working stage is calculated in order to determine the **performance difference**.
- The performance difference determines the **probability** that a player wins the competition and thus receives the high payoff as shown in the following table:

Absolute performance difference	Probability (in %) for the player with the **higher score** to receive the high payment (Probability for the player with the **lower score** to receive the high payment)
greater or equal to 100.5	90 (10)
70.5–100	80 (20)
45.5–70	70 (30)
20.5–45	60 (40)
0–20	50 (50)

- Example: If you have achieved 50 points more than the other player in your group in the working stage, your probability of getting the high payoff is 70% (otherwise you will get the lower payoff). This means for the other player in your group that he will get the high payoff with a probability of 30% (and otherwise he will get the lower payoff). Keep in mind that only one player in your group will get the high payoff and the other one will get the lower payoff.
- We will inform you of the performance difference in your group and whether you have achieved the higher or lower score. The other player receives this information as well.

Decision-Making Stage 1: Choice of Number Z

- Now you can select a number Z which may influence the probability that you or the other player will win the competition and thus get the high payoff. For this purpose, you select a number Z between 0 and 9.
- Depending on the selected number Z, the other player's probability of winning decreases by the appropriate number of percentage points and your profit probability increases by the appropriate number: If you select zero, the probabilities of winning do not change. For **selecting 1**, the other player's probability of winning is **reduced by one percentage point**, while your percentage increases by one percentage point. For **selecting 2**, the other player's probability of winning is **reduced by two percentage points**, while your probability of winning increases by two percentage points, and so on.
- The higher the selected number Z, the higher the associated costs. The costs are listed in the following table. If this round is randomly drawn for payoff at the end, the costs for the number Z will be deducted from your payoff.

Number Z	0	1	2	3	4	5	6	7	8	9
Cost in token	0	1	2	4	6	9	12	16	20	25

- The other player in your group also selects a number Z that may reduce your probability of get-ting the high payoff and increase his probability of getting the high payoff, respectively. For the other player, the same rules for the selection of Z apply.
- Whether the number Z selected by you and/or the number Z selected by the other player in your group are implemented and have an effect on the probability of winning is determined as follows:
- The **18 attendant groups** are separated in:
  - **12** randomly determined **groups** in which the selected numbers Z of **both players** have an effect.
  - **1** randomly determined **group** in which the selected numbers Z of **both players have no effect**.
  - **5** randomly determined **groups** in which only the number Z of **one of the two players** has an effect.
- Please note that neither you nor the other player knows which of the conditions mentioned above applies to your group when choosing Z, and that the cost of selecting Z always incur.
- After you and the other player have selected a number Z, the computer will first check the condition which applies to your group, and then it will calculate the **final probabilities of winning** in your group according to this condition and the selected numbers Z (if

relevant).

- Finally, the computer will determine who gets the high payoff and who gets the lower payoff, given the final probabilities of winning.

Decision-Making Stage 1: Information about Z

- <INFO SELF$_{100}$ OPP$_0$ and INFO SELF$_{70}$ OPP$_0$> No player will be informed about the number Z **selected by the other player and whether or not this has an effect** on the probability of winning.

  <INFO SELF$_{100}$ OPP$_{70}$> Every player will always be informed whether or not the number Z he has selected has an effect on the probability of winning.

- <INFO SELF$_{100}$ OPP$_0$> **In any case**, you will be informed about the effect of the number Z **you** have selected.

  <INFO SELF$_{70}$ OPP$_0$> However, there is the possibility that you will be informed about the effect of the number Z **you** have selected.

  <INFO SELF$_{100}$ OPP$_{70}$> Additionally, there is the possibility that you will be informed about the number Z the **other player** has selected and whether this has an effect.

- <INFO SELF$_{100}$ OPP$_0$> Also, **the other player** will be informed about the effect of the number Z selected **by him** in any case.

  <INFO SELF$_{70}$ OPP$_0$> For the **other player**, there is also the possibility that he will be informed about the effect of the number Z **he** has selected.

  <INFO SELF$_{100}$ OPP$_{70}$> For the **other player**, there is also the possibility that he will be informed about the effect of the number Z **you** have selected.

- <INFO SELF$_{70}$ OPP$_0$ and INFO SELF$_{100}$ OPP$_{70}$> There are four scenarios in total which occur with different probabilities:

  – <INFO SELF$_{70}$ OPP$_0$> Scenario 1: **In 50% of the cases, both players receive information**. That means, **you** will be informed about the effect of the number Z **you** have selected, and the **other player** will be informed about the effect of the number Z **he** has selected.

    <INFO SELF$_{100}$ OPP$_{70}$> Scenario 1: **In 50% of the cases, both players receive information**. That means, **you** will be informed about number Z the **other player** has selected and its effect, and the **other player** will be informed about the number Z **you** have chosen and its effect.

  – <INFO SELF$_{70}$ OPP$_0$> Scenario 2: **In 20% of the cases, only you receive information**. That means, **you** will be informed about the effect of the number Z **you** have selected,

but the **other player**will **not** be informed about the effect of the number Z he has selected.

<INFO SELF$_{100}$ OPP$_{70}$> Scenario 2: **In 20% of the cases, only you receive information**. That means, **you** will be informed about the number Z the **other player** has selected and its effect, but the **other player** will **not** be informed about the number Z you have selected and its effect.

– <INFO SELF$_{70}$ OPP$_{0}$> Scenario 3: **In 20% of the cases, only the other player receives information**. That means, **you** will **not** be informed about the effect of the number Z you have selected, but the **other player** will be informed about the effect of the number Z **he** has selected.

<INFO SELF$_{100}$ OPP$_{70}$> Scenario 3: **In 20% of the cases, only the other player receives information**. That means, **you** will **not** be informed about the number Z the other player has selected and its effect, but the **other player** will be informed about the number Z **you** have selected and its effect.

– <INFO SELF$_{70}$ OPP$_{0}$> Scenario 4: **In 10% of the cases, no player receives information**. That means, **you** will **not** be informed about the effect of the number Z you have selected, and the **other player** will **not** be informed about the effect of the number Z he has selected.

<INFO SELF$_{100}$ OPP$_{70}$> Scenario 4: **In 10% of the cases, no player receives information**. That means, **you** will **not** be informed about the number Z the other player has selected and its effect, and the **other player** will **not** be informed about the Z you have selected and its effect.

<INFO SELF$_{70}$ OPP$_{0}$>

Scenario	Probability (in %) for this scenario to occur	You receive information about the effect of **your number** Z	The other player receives information about the effect of **his number** Z
1	50	Yes	Yes
2	20	Yes	No
3	20	No	Yes
4	10	No	No

<INFO SELF$_{100}$ OPP$_{70}$>

Scenario	Probability (in %) for this scenario to occur	You receive information about the **number Z of the other player**	The other player receives information about **your number Z**
1	50	Yes	Yes
2	20	Yes	No
3	20	No	Yes
4	10	No	No

Decision-Making Stage 1: Information about the Result

- At the end of the experiment, you will be informed about your payoff. The payoff is made up as follows:
  - winner of the competition: 75 tokens − costs for the choice of Z
  - loser of the competition: 25 tokens − costs for the choice of Z
- At the end of the experiment, we will inform you about the number Z you have selected and the associated costs from decision-making stage 1.
- <INFO SELF$_{100}$ OPP$_0$> You will be informed about the effect of the number Z **you** have selected.

  <INFO SELF$_{70}$ OPP$_0$> Additionally, you may be informed about the effect of the number Z **you** have selected, see description above.

  <INFO SELF$_{100}$ OPP$_{70}$> You will be informed about the effect of the number Z you have selected. Additionally, you may be informed about the effect of the number Z the **other player** has selected and its effect, see description above.

Before the working stage starts, you have the opportunity to become acquainted with the process decision-making stage 1 in 3 trial periods. For this purpose, you take the role of both players in a group and can determine the numbers Z. After the end of the trial periods, we will ask you to answer a couple of comprehensive questions. If you still have questions, please give us a signal.

Decision-Making Stage 2: Course of Action and Choice of Number Z

- At the beginning of this round, two players form a group. Note that in **no case** you will be assigned to a group with the same player as in decision-making stage 1.
- The process essentially corresponds to the previous decision-making stage. Again, there will be a **competition**. As in decision-making stage 1, your result and the result of the other player from the working stage prior to decision-making stage 1 will be used to determine the performance difference in your **new group**. The performance difference determines,

as in decision-making stage 1, the probability with which a player wins the competition and receives the high payoff of 75 tokens. Here you can see the table that is the same as the round before.

Absolute performance difference	Probability (in %) for the player with the **higher score** to receive the high payment (Probability for the player with the **lower score** to receive the high payment)
greater or equal to 100.5	90 (10)
70.5–100	80 (20)
45.5–70	70 (30)
20.5–45	60 (40)
0–20	50 (50)

- You and the other player select a number Z again, which is associated with the same consequences and costs as in decision-making stage 1. The cost associated with each number Z is given in the following table.

Number Z	0	1	2	3	4	5	6	7	8	9
Cost in token	0	1	2	4	6	9	12	16	20	25

- The only difference to round of decision 1 is that an **additional decision is made regarding the information about the number** Z. This additional decision is explained below.

Decision-Making Stage 2: Influencing Information about Z

- <INFO SELF_{100} OPP_0 and INFO SELF_{70} OPP_0> In contrast to decision-making stage 1, you now have the opportunity to **prevent** us from informing **you** about the effects of the number Z **you** have selected.

  <INFO SELF_{100} OPP_{70}> In contrast to decision-making stage 1, you now have the opportunity to **prevent** us from informing the **other player** about the number Z **you** have selected and its effect.

- <INFO SELF_{70} OPP_0 and INFO SELF_{100} OPP_{70}> Reminder: There are four scenarios which occur with different probabilities.

  - <INFO SELF_{70} OPP_0> Scenario 1: **In 50% of the cases, both players receive information**. That means, **you** will be informed about the effect of the number Z **you** have selected, and the **other player** will be informed about the effect of the number Z **he** has selected.

    <INFO SELF_{100} OPP_{70}> Scenario 1: **In 50% of the cases, both players receive information**. That means, **you** will be informed about number Z the **other player** has

selected and its effect, and the **other player** will be informed about the number Z **you** have chosen and its effect.

– <INFO SELF_{70} OPP_{0}> Scenario 2: **In 20% of the cases, only you receive information**. That means, **you** will be informed about the effect of the number Z **you** have selected, but the **other player** will **not** be informed about the effect of the number Z he has selected.

<INFO SELF_{100} OPP_{70}> Scenario 2: **In 20% of the cases, only you receive information**. That means, **you** will be informed about the number Z the **other player** has selected and its effect, but the **other player** will **not** be informed about the number Z you have selected and its effect.

– <INFO SELF_{70} OPP_{0}> Scenario 3: **In 20% of the cases, only the other player receives information**. That means, **you** will **not** be informed about the effect of the number Z you have selected, but the **other player** will be informed about the effect of the number Z **he** has selected.

<INFO SELF_{100} OPP_{70}> Scenario 3: **In 20% of the cases, only the other player receives information**. That means, **you** will **not** be informed about the number Z the other player has selected and its effect, but the **other player** will be informed about the number Z **you** have selected and its effect.

– <INFO SELF_{70} OPP_{0}> Scenario 4: **In 10% of the cases, no player receives information**. That means, **you** will **not** be informed about the effect of the number Z you have selected, and the **other player** will **not** be informed about the effect of the number Z he has selected.

<INFO SELF_{100} OPP_{70}> Scenario 4: **In 10% of the cases, no player receives information**. That means, **you** will **not** be informed about the number Z the other player has selected and its effect, and the **other player** will **not** be informed about the Z you have selected and its effect.

<INFO SELF_{100} OPP_{0}> By paying a certain price, you can ensure that you do **not** receive information about the number Z you have chosen in **any case**.

<INFO SELF_{70} OPP_{0}> By paying a certain price, you can ensure that you do **not** receive information about the number Z you have chosen **in any scenario** described above.

<INFO SELF_{100} OPP_{70}> By paying a certain price, you can ensure that the **other player** does not receive information about the number Z you have selected and its effect **in any scenario** described above.

- This price is between 1 and 10 tokens (in integer increments) and will be determined randomly by the computer at the end. Each price occurs with the same probability.
- Please indicate the maximum price you are willing to pay. This means, you are willing to

pay this price and any price below. For this purpose, we will show you a table in which you can enter your decision (see screenshot at the end of the instructions).

- <INFO SELF$_{70}$ OPP$_0$> If a price is drawn for which you have ticked the option "pay", you will **not** be informed about the effect of the number Z you have selected on the probability of winning **in any scenario**. At the end, the corresponding price will be deducted from your payoff.

  <INFO SELF$_{100}$ OPP$_{70}$> If a price is drawn for which you have ticked the option "pay", **the other player** will **not** be informed about the effect of the number Z **you** have selected on the probability of winning **in any scenario**. At the end, the corresponding price will be deducted from your payoff.

- <INFO SELF$_{100}$ OPP$_0$> If a price is drawn for which you have ticked the option "do **not** pay" option, you will be informed about the effect of the number Z you have selected **in any case**—just as in decision-making stage 1. In this case, nothing will be deducted from your payoff.

  <INFO SELF$_{70}$ OPP$_0$> If a price is drawn for which you have ticked the option "do **not** pay" option, you will be informed about the effect of the number Z you have selected **depending on the scenario**—just as in decision-making stage 1. In this case, nothing will be deducted from your payoff.

  <INFO SELF$_{100}$ OPP$_{70}$> If a price is drawn for which you have ticked the option "do **not** pay" option, the **other player** will be informed about the effect of the number Z **you** have selected **depending on the scenario**—just as in decision-making stage 1. In this case, nothing will be deducted from your payoff.

- <INFO SELF$_{100}$ OPP$_0$ and INFO SELF$_{70}$ OPP$_0$> For the other player, there is also the opportunity to **prevent him** from being informed about the effect of the number Z **he** has selected.

  <INFO SELF$_{100}$ OPP$_{70}$> For the other player, there is also the opportunity to **prevent you** from being informed about the number Z **he** has selected and its effect.

- The computer draws the price which applies to both players of a group.

Decision-Making Stage 2: Information about the Result

- At the end of the experiment, you will be informed about your payoff. The payoff is made up as follows:
  - winner of the competition: 75 tokens − costs for the choice of Z − cost of preventing information disclosure/price (if applicable)
  - loser of the competition: 25 tokens − costs for the choice of Z − cost of preventing information disclosure/price (if applicable)

- At the end of the experiment, we will inform you about the number Z you have selected and the associated costs from decision-making stage 2.
- <INFO SELF$_{100}$ OPP$_0$ and INFO SELF$_{70}$ OPP$_0$> Additionally, you may be informed about the effect of the number Z **you** have selected, see description above.

  <INFO SELF$_{100}$ OPP$_{70}$> You will be informed about the effect of the number Z you have selected. Additionally, you may be informed about the effect of the number Z the **other player** has selected and its effect, see description above.

Price		Your Decision
	I am not willing to pay a price	
1 Token	I am willing to pay at most 1 token	
2 Tokens	I am willing to pay at most 2 tokens	
3 Tokens	I am willing to pay at most 3 tokens	
4 Tokens	I am willing to pay at most 4 tokens	
5 Tokens	I am willing to pay at most 5 tokens	
6 Tokens	I am willing to pay at most 6 tokens	
7 Tokens	I am willing to pay at most 7 tokens	
8 Tokens	I am willing to pay at most 8 tokens	
9 Tokens	I am willing to pay at most 9 tokens	
10 Tokens	I am willing to pay at most 10 tokens	

Part 2

In this part, groups of two players are formed again. Note that you are in a group with a player with whom you have not played together in part 1.

You will make a decision in **10 situations** that can affect your payoff and the payoff of the other player. In each of the 10 situations, you can choose between **two alternatives**: “Left” and “Right”.

For each alternative, you will see two values: the first value is your potential payoff and the

second value indicates the possible payoff of the other player.

In each of the 10 situations, please select **one of the two alternatives** and tick the corresponding circle in the middle column of the table that will appear on the screen.

The other player also chooses one of the two alternatives in each of the 10 situations.

Example of a situation (there is a total of 10 situations):

Left				**Right**	
You receive	The other player receives			You receive	The other player receives
a tokens	*b* tokens	Left ◯	◯ Right	*c* tokens	*d* tokens

The letters *a*, *b*, *c*, and *d* are only for illustrative purposes and will be replaced by numbers.

Payoff

After you and the other player have made a decision for all 10 situations, the computer randomly draws **one situation** that is relevant to the payoff. Afterwards, the computer also randomly determines whether your decision or that of the other player will be implemented for this situation. So there are **two scenarios**.

If the computer has randomly drawn the situation described in the table above:

- Scenario 1: **Your** decision is relevant for payoff.
  - You have selected alternative **Left** for the selected situation: You will receive *a* tokens and the other player will receive *b* tokens.
  - You have selected alternative **Right** for the selected situation: You will receive *c* tokens and the other player will receive *d* tokens.
- Scenario 2: The decision of the **other player** is relevant for the payoff.
  - The other player has chosen the alternative **Left** for the selected situation: You will receive *b* tokens and the other player will receive *a* tokens.
  - The other player has chosen the alternative **Right** for the selected situation: You will receive *d* tokens and the other player will receive *c* tokens.

At the end of this part, you will be informed which line has been drawn for the payoff. You will also be informed whether your decision or the decision of the other player is relevant for the payoff. We will also inform you about the payoff from this part.

Appendix D Statistical Methods

D.1 Pearson's χ^2-Test

The Pearson's χ^2-test (named after Pearson, 1900) is used to test whether two unpaired samples of categorical data which are stored in a $k \times m$[1] contingency table are independent.

	$j=1$	$j=2$	$\cdots$	$j=m$	Σ_i
$i=1$	h_{11}	h_{12}	$\cdots$	h_{1m}	$N_{1\bullet}$
$i=2$	h_{21}	h_{22}	$\cdots$	h_{2m}	$N_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$i=k$	h_{k1}	h_{k2}	$\cdots$	h_{km}	$N_{k\bullet}$
Σ_j	$N_{\bullet 1}$	$N_{\bullet 2}$	$\cdots$	$N_{\bullet m}$	N

h_{ij} is the number of observations of the cell in the i-th row and the j-th column. $N_{i\bullet}$ and $N_{\bullet j}$ denote the sum of observations in the i-th row and the j-th column, respectively, and N is the total number of observations.

The test statistic is given by

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{m} \frac{(h_{ij} - e_{ij})^2}{e_{ij}},$$

where

$$e_{ij} = \frac{N_{i\bullet} \cdot N_{\bullet j}}{N}$$

is the expected number of observations of the cell in the i-th row and j-th column.

For a sufficiently high number of observations in each cell (expected number of observations in each cell must be at least 1; at most 20% of the cells may exhibit an expected number of observations of less than 5), the test statistic asymptotically approaches a χ^2-distribution with $(k-1)\cdot(m-1)$ degrees of freedom (Bortz and Lienert, 2008; Bortz *et al.*, 2008; Rasch *et al.*, 2014b).

To account for family-wise error rates due to multiple hypothesis testing in pairwise comparisons, the Holm-Bonferroni correction (Holm, 1979) is used. Let $P_{(1)}, \ldots, P_{(v)}$ be the sequence of p-values (in ascending order) of a family of v null hypotheses $H_{(1)}, \ldots, H_{(v)}$. Given a significance level α, the lowest rank $u \in \{1, \ldots, v\}$ is determined which satisfies the inequality

$$P_{(u)} > \frac{\alpha}{v+1-u}.$$

Finally, the null hypotheses $H_{(1)}, \ldots, H_{(u-1)}$ are rejected and the null hypotheses $H_{(u)}, \ldots, H_{(v)}$ are not.[2]

D.2 Bowker Test of Symmetry

The Bowker test of symmetry (Bowker, 1948) compares changes in a variable with k levels for a paired sample during two distinct measurements. h_{ij} denotes the observed frequency in the

i-th row and j-th column of the corresponding $k \times k$ contingency table. The null hypothesis is that changes in either direction between the two measurements are equally likely, i.e., the contingency table is symmetric with respect to the main diagonal. The test statistic is given by

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\left(h_{ij} - h_{ji}\right)^2}{h_{ij} + h_{ji}}, \quad i > j$$

and approximates the χ^2 distribution with $\binom{k}{2} = \frac{k \cdot (k-1)}{2}$ degrees of freedom. It is required that the expected frequencies $e_{ij} = \frac{(h_{ij} + h_{ji})}{2}$ are equal to or larger than 5 for at least 80% of the pairs (Bortz and Lienert, 2008; Bortz *et al.*, 2008).

D.3 Sign Test

The one-sample sign test is used to test whether the median of a sample is statistically different from a specific value η.[3] Under the null hypothesis, an equal number of positive and negative deviations from the median, N^+ and N^-, is expected and $N^+, N^- \sim \text{Binom}(n, 0.5)$, where n is the sample size.

If the actual median is larger than the hypothesized value η, the number of observed negative deviations, n^-, is lower than expected. The corresponding p-value is calculated by

$$p = \Pr(N^- \leq n^-) = \sum_{i=0}^{n^-} \binom{n}{i} \cdot \left(\frac{1}{2}\right)^n$$

for small samples ($n < 30$) (Bradley, 1968; Schaich and Hamerle, 1984).

D.4 Mann-Whitney U-Test

The Mann-Whitney U-test (Mann and Whitney, 1947; see also Wilcoxon, 1945; Kruskal, 1957) is used to test whether two unpaired samples of ordinal data are drawn from the same population. To calculate the test statistic, the observations are ranked in ascending order. Ties are assigned the average rank. The rank sum T_i of sample $i = 1, 2$ is denoted by

$$T_i = \sum_{m_i=1}^{n_i} R_{m_i},$$

where R_{m_i} is the rank of the m-th observation in sample i. For each sample, the test statistics

$$U_1 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - T_1$$

and

$$U_2 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - T_2$$

are calculated with n_1, n_2 as number of observations in the first and second sample, respectively. The smaller value of U_1 and U_2, $U = \min\{U_1, U_2\}$, is compared to the critical value for a given

significance level which is usually retrieved from tables for small samples (e.g., Table 8.2 in the appendix of Sani and Todman, 2006). For larger samples ($n_1, n_2 > 20$), the test statistic U approaches a normal distribution and the critical value is given by

$$z_U = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}}$$

To account for ties in larger samples ($n_1 > 20$ or $n_2 > 20$), the following correction is used:

$$z_{U(\text{corr})} = \frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2}{N \cdot (N-1)} \cdot \left(\frac{N^3 - N}{12} - \sum_{s=1}^{r} \frac{\tau_s^3 - \tau_s}{12} \right)}}$$

with $N = n_1 + n_2$ as total number of observations, r as total number of ties, and τ_s as number of ties for the s-th tied value (Bortz and Lienert, 2008; Bortz *et al.*, 2008; Rasch *et al.*, 2014b).

D.5 Kruskal-Wallis H-Test

Similar to the Mann-Whitney U-test, the Kruskal-Wallis H-test (Kruskal, 1952; Kruskal and Wallis, 1952) tests whether $k > 2$ samples of ordinal data belong to the same population. The test statistic in case of no ties is denoted by

$$H = \frac{12}{N \cdot (N+1)} \cdot \sum_{i=1}^{k} \frac{T_i^2}{n_i} - 3 \cdot (N+1),$$

where N is the total number of observations. Under the null hypothesis, the test statistic approaches a χ^2-distribution with $k-1$ degrees of freedom.

To account for ties in the ranks, the following correction can be applied:

$$H_{\text{corr}} = \frac{H}{1 - \frac{\sum_{s=1}^{r} (\tau_s^3 - \tau_s)}{N^3 - N}}$$

with r as total number of ties and τ_s as number of ties for the s-th tied value (Bortz and Lienert, 2008; Bortz *et al.*, 2008; Rasch *et al.*, 2014b).

To test which samples are statistically different from each other, the Dunn test (Dunn, 1964) is used as *post-hoc* test for the Kruskal-Wallis H-test. In particular, $h = \frac{k \cdot (k-1)}{2}$ pairwise tests are performed. Let A and B be a pair of samples which are to be compared. The corresponding test statistic is denoted by

$$z = \frac{|\bar{R}_A - \bar{R}_B|}{\sqrt{\left(\frac{N \cdot (N+1)}{12} - \frac{\sum_{s=1}^{r} (\tau_s^3 - \tau_s)}{12 \cdot (N-1)} \right) \cdot \left(\frac{1}{n_A} + \frac{1}{n_B} \right)}},$$

where $\bar{R}_A = \frac{R_A}{n_A}$ is the mean rank of sample A with R_A as rank sum and n_A as number of observations of sample A (and $\bar{R}_B$ similarly for sample B). To account for family-wise error rates, the resulting p-values have to be corrected, e.g., by using the Holm-Bonferroni method (Dinno, 2015).

D.6 Wilcoxon Signed-Rank Test

To consider statistical differences in two paired samples of ordinal data, the Wilcoxon signed-rank test (Wilcoxon, 1945) is used. The absolute difference $d_i = x_{Ai} - x_{Bi}$ for each pair of observations $i \in \{1, \ldots, N\}$ in both samples A and B is calculated and ordered in an ascending sequence. T_- and T_+ are the rank sums of differences in absolute terms with negative and positive sign, respectively. Tied values are assigned the average rank. The test statistic T is the smaller value of T_- and T_+, $T = \min\{T_-, T_+\}$. The critical values for small samples can be retrieved from tables (e.g., Table 8.1 in the appendix of Sani and Todman, 2006). For larger samples ($N > 50$), the test statistic T approaches a normal distribution and the critical value is given by

$$z_T = \frac{T - \frac{N \cdot (N+1)}{4}}{\sqrt{\frac{N \cdot (2N+1) \cdot (N+1)}{24}}}.$$

To correct for tied ranks and zero differences ($d_i = 0$), the critical value can be corrected in the following way:

$$z_{T(\text{corr})} = \frac{T - \frac{N \cdot (N+1)}{4}}{\sqrt{\frac{N \cdot (2N+1) \cdot (N+1) - \sum_{s=1}^{r} \frac{\left(\tau_s^3 - \tau_s\right)}{2}}{24}}}$$

with r as total number of ties and τ_s as number of ties for the s-th tied value (Bortz and Lienert, 2008; Bortz *et al.*, 2008; Rasch *et al.*, 2014b).

D.7 Z-Test

The Z-test is used to test whether the relative frequency of an event in a sample is statistically different from particular proportion (one-sample test) or whether the relative frequencies of two independent samples are statistically different from each other (two-sample test). Unlike the statistical tests discussed above, this test relies on the assumption that the data are approximately normally distributed.

The test statistic for the one-sample test is given by

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot (1-p)}{n}}},$$

where $\hat{p} = \frac{x}{n}$ is the observed relative frequency with x as the number of successful outcomes and n as the number of observations, and p the theoretical proportion. To meet the approximate normality assumption, it is required that $n \cdot p \geq 5$ and $n \cdot (1-p) \geq 5$.

The test statistic for the two-sample test is given by

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_c \cdot (1 - \hat{p}_c) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}},$$

where $\hat{p}_1 = \frac{x_1}{n_1}$ and $\hat{p}_2 = \frac{x_2}{n_2}$ are the observed relative frequencies with x_1, x_2 as the number of successful outcomes and n_1, n_2 as the number of observations in the two samples, and $\hat{p}_c = \frac{x_1 + x_2}{n_1 + n_2}$ is the is the observed relative frequency of the pooled samples. To meet the approximate normality assumption, it is required that $n_1, n_2 \geq 30$ (Zou *et al.*, 2003).

D.8 Spearman's ρ

Spearman's ρ (named after Spearman, 1904, 1906) measures the association between two ordinal variables. The rank correlation coefficient is given by

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^{N} d_i^2}{N \cdot (N-1)},$$

where $d_i = R_{x_i} - R_{y_i}$ is the difference of ranks between two variables x and y of individual i and N denotes the number of observations.

The rank correlation coefficient is asymptotically normally distributed with expected value $\mu = 0$ and variance $\sigma^2 = \frac{1}{N-1}$ for $N > 30$. Hence, the critical value is

$$z = \rho \cdot \sqrt{N-1}$$

(Bortz and Lienert, 2008; Bortz *et al.*, 2008; Rasch *et al.*, 2014a).

D.9 Ordinary Least Squares

The ordinary least squares approach estimates the coefficients of a linear model

$$y_i = \mathrm{x}_i' \boldsymbol{\beta} + u_i, \quad i = 1, \ldots, N$$

such that the sum of squared residuals, $\sum_{i=1}^{N} (y_i - \mathrm{x}_i' \boldsymbol{\beta})^2$, i.e., the difference between the observed and estimated values, is minimized. $\boldsymbol{\beta}$ is a $K \times 1$ column vector of coefficients which are to be estimated, $\mathbf{x}_i'$ is a $1 \times K$ row vector of observed regressors for individual i, and u_i is the residual.

The corresponding estimator $\hat{\boldsymbol{\beta}}_{\mathrm{OLS}}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\mathrm{OLS}} &= \left(\mathbf{X}'\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{y} \\ &= \left(\sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i'\right)^{-1} \sum_{i=1}^{N} \mathbf{x}_i y_i \end{aligned}$$

$$= \begin{bmatrix} \sum_{i=1}^{N} x_{1i}^2 & \sum_{i=1}^{N} x_{1i}x_{2i} & \cdots & \sum_{i=1}^{N} x_{1i}x_{Ki} \\ \sum_{i=1}^{N} x_{2i}x_{1i} & \sum_{i=1}^{N} x_{2i}^2 & \cdots & \sum_{i=1}^{N} x_{2i}x_{Ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} x_{Ki}x_{1i} & \sum_{i=1}^{N} x_{Ki}x_{2i} & \cdots & \sum_{i=1}^{N} x_{Ki}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} x_{1i}y_i \\ \sum_{i=1}^{N} x_{2i}y_i \\ \vdots \\ \sum_{i=1}^{N} x_{Ki}y_i \end{bmatrix},$$

where $\mathbf{X}$ is a $N \times K$ matrix of observed regressors for all individuals, i.e., $\mathbf{X}$ consists of N row vectors $\mathbf{x}_i'$ and $\mathbf{y}$ is a $N \times 1$ column vector of observed outcomes y_i for all individuals.[4]

To obtain consistent estimates it is required that

(i) the residuals have an expected value of zero and must not be correlated with the regressors (exogneity),

(ii) the variance σ^2 of the residuals is constant (homoskedasticity), and

(iii) $(y_i, \mathbf{x}_i), i = 1, \ldots, N$ are independent and identically distributed (independence).

If these assumptions are met, the variance-covariance matrix of the estimator (VCE) is asymptotically normally distributed and given by

$$\hat{V}_{\text{default}}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \left(\mathbf{X}'\mathbf{X}\right)^{-1}$$

with $\hat{\sigma}^2 = \frac{1}{N-K}\sum_{i=1}^{N} \hat{u}_i^2$ as consistent estimator of the variance σ and $\hat{u}_i = y_i - \mathbf{x}_i'\boldsymbol{\beta}$.

When residuals are correlated within clusters (groups), the third assumption is violated. To account for this issue, a cluster-robust VCE is obtained by

$$\hat{V}_{\text{cluster}}(\hat{\boldsymbol{\beta}}) = \left(\mathbf{X}'\mathbf{X}\right)^{-1} \left(\frac{G}{G-1} \frac{N-1}{N-K} \sum_{g=1}^{G} \mathbf{X}_g \hat{\mathbf{u}}_g \hat{\mathbf{u}}_g' \mathbf{X}_g' \right) \left(\mathbf{X}'\mathbf{X}\right)^{-1},$$

where $g = 1, \ldots, G$ denotes the clusters, $\hat{\mathbf{u}}_g$ is a vector of residuals for cluster g, and $\mathbf{X}_g$ is the corresponding matrix of regressors (Cameron and Trivedi, 2009; Wooldridge, 2010; Greene, 2012; Stock and Watson, 2015).

D.10 Generalized Least Squares with Random Effects

Unlike cross sectional data, panel data contain repeated measures of the same individuals over time and, hence, are valuable to analyze changes over time on individual level. Repeated measures of the same individual over time, however, are serially correlated, i.e., not independent.

The linear model is given by

$$y_{it} = \mathrm{x}_i'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it},$$

where α_i is the individual-specific and time-invariant component of the residual and ε_{it} is the component of the residual which varies over time and over individuals. Depending on the

treatment of α_i, the linear model is estimated with a fixed-effects or a random-effects estimator. If α_i is assumed be correlated with the regressors, the fixed-effects estimator is appropriate, whereas the random-effects estimator is to be used when no such correlation is assumed. If the random-effects estimator is erroneously used instead of the fixed-effects estimator, the results are not consistent due to endogeneity. On the other hand, the fixed-effects estimator does not allow to include time-invariant regressors in the model.

The estimator $\hat{\boldsymbol{\beta}}_{\text{RE}}$ is given by

$$\hat{\boldsymbol{\beta}}_{\text{RE}} = \left(\sum_{i=1}^{N} \mathbf{X}_i' \hat{\Omega}^{-1} \mathbf{X}_i \right)^{-1} \left(\sum_{i=1}^{N} \mathbf{X}_i' \hat{\Omega}^{-1} y_i \right)$$

with VCE

$$\hat{\Omega} = \hat{\sigma}_\varepsilon^2 \mathbf{I}_T + \hat{\sigma}_\alpha^2 \mathbf{j}_T' \mathbf{j}_T,$$

where $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\alpha^2$ are consistent estimators of the variance of ε_{it} and α_i, respectively. $\mathbf{I}_T$ is a $T \times T$ identity matrix and $\mathbf{j}_T$ is a $T \times 1$ column vector of ones (Cameron and Trivedi, 2009; Wooldridge, 2010; Greene, 2012; Andreß *et al.*, 2013).

D.11 Tobit

The Tobit (I) model (named after Tobin, 1958) is used to account for a limited dependent variable, i.e., the dependent variable assumes only values within a certain range. Coefficients estimated with ordinary least squares are not consistent in this case.

The regression model is given by

$$y_i^* = \mathbf{x}_i' \boldsymbol{\beta} + u_i,$$

where y_i^* is a latent variable, $\mathbf{x}_i'$ is a $K \times 1$ vector of regressors, and $u_i \sim \mathcal{N}(0, \sigma^2)$ is the error term. The relation between the observed variable y_i and the latent variable y_i^* is established by

$$y_i = \begin{cases} y_L & \text{if } y_i^* \leq y_L \\ y_i^* & \text{if } y_L < y_i^* < y_U \\ y_U & \text{if } y_i^* \geq y_U \end{cases},$$

where y_L and y_U denote the lower and upper boundary, respectively.

The estimator $\hat{\boldsymbol{\beta}}_{\text{Tobit}}$ is obtained using the following log-likelihood function:

$$\hat{\boldsymbol{\beta}}_{\text{Tobit}} = \arg\max_{\beta} \left\{ \sum_{y_i^* \leq y_L} \log\left(\Phi\left(\frac{y_L - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma} \right) \right) + \sum_{y_L < y_i^* < y_U} \log\left(\frac{1}{\sigma} \cdot \phi\left(\frac{y_i^* - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma} \right) \right) + \sum_{y_i^* \geq y_U} \log\left(1 - \Phi\left(\frac{y_U - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma} \right) \right) \right\},$$

where $\Phi(\bullet)$ and $\phi(\bullet)$ denote the cumulative distribution function and the probability density function of the standard normal distribution, respectively (Cameron and Trivedi, 2009; Wooldridge, 2010; Greene, 2012).

D.12 Truncated Normal Hurdle Model

The truncated normal hurdle model (Cragg, 1971) is a two-part model which extends the Tobit model by allowing for different mechanisms to describe an individual's decision *(i)* whether or not to take an action (participation) and *(ii)* which quantity of this action to choose if the previous decision was positive (intensity), i.e., it separates the extensive and intensive margin. More formally, the first part of the model is estimated using the Probit estimator

$$\Pr(y_i^* > 0) = \Phi\left(\mathbf{x}_i'\boldsymbol{\gamma}\right),$$

where y_i^* is a latent variable and

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0 \\ - & \text{if } y_i^* \leq 0 \end{cases}.$$

The linear regression model is given by

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + u_i.$$

The estimator $\hat{\boldsymbol{\beta}}_{\text{TNH}}$ is obtained using the following log-likelihood function:

$$\hat{\boldsymbol{\beta}}_{\text{TNH}} = \arg\max_{\beta} \Big\{ \mathbb{1}_{[y_i=0]} \cdot \log\left(1 - \Phi(\mathbf{x}_i\boldsymbol{\gamma})\right) + \mathbb{1}_{[y_i>0]} \cdot \log\left(\Phi(\mathbf{x}_i\boldsymbol{\gamma})\right) + \mathbb{1}_{[y_i>0]} \cdot \left[-\log\left(\Phi\left(\frac{\mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right)\right) + \log\left(\phi\left(\frac{y_i - \mathbf{x}_i\boldsymbol{\beta}}{\sigma}\right)\right) - \log(\sigma) \right] \Big\}$$

(Wooldridge, 2010; Greene, 2012).

D.13 Ordered Logit

The Ordered Logit model is used to estimate coefficients in a model with a latent variable y_i^*:

$$y_i^* = \mathbf{x}_i'\boldsymbol{\beta} + u_i.$$

The observed outcome variable y_i can assume one of $k = 1, \ldots, m$ discrete values s_k (in ascending order). The relation between the latent variable y_i^* and the observed variable y_i is denoted by

$$y_i = \begin{cases} s_1 & \text{if } y_i^* \leq \delta_1 \\ s_2 & \text{if } \delta_1 < y_i^* < \delta_2 \\ \vdots & \\ s_{m-1} & \text{if } \delta_{m-2} < y_i^* < \delta_{m-1} \\ s_m & \text{if } y_i^* \geq \delta_{m-1} \end{cases},$$

where $\delta_\bullet$ are $m-1$ thresholds which separate the observed values of the outcome variable.

The estimator $\hat{\boldsymbol{\beta}}_{\text{Ordered Logit}}$ is obtained using the following log-likelihood function:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{\text{Ordered Logit}} = \arg\max_{\beta} \Bigg\{ & \sum_{y_i^* \leq \delta_1} \log\left(\frac{\exp\{\delta_1 - \mathbf{x}_i'\boldsymbol{\beta}\}}{1+\exp\{\delta_1 - \mathbf{x}_i'\boldsymbol{\beta}\}} \right) \\ & + \sum_{k=2}^{m-1} \left(\sum_{\delta_k < y_i^* < \delta_{k+1}} \log\left(\frac{\exp\{\delta_k - \mathbf{x}_i'\boldsymbol{\beta}\}}{1+\exp\{\delta_k - \mathbf{x}_i'\boldsymbol{\beta}\}} - \frac{\exp\{\delta_{k-1} - \mathbf{x}_i'\boldsymbol{\beta}\}}{1+\exp\{\delta_{k-1} - \mathbf{x}_i'\boldsymbol{\beta}\}} \right) \right) \\ & + \sum_{y_i^* \geq \delta_{m-1}} \log\left(1 - \frac{\exp\{\delta_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}\}}{1+\exp\{\delta_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}\}} \right) \Bigg\} \end{aligned}$$

(Greene and Hensher, 2010).

Notes

1. $k, m \geq 2$. The case of $k, m = 2$ refers to binary data.
2. To facilitate the interpretation, p-values adjusted with the Holm-Bonferroni method in this thesis are multiplied with $(v+1-u)$ for $P_{(1)}, \ldots, P_{(u-1)}$ and can, thus, be directly compared to conventional significance levels.
3. Note that the sign test is a special case of the Binomial test (Schaich and Hamerle, 1984).
4. Note that the first column of $\mathbf{X}$ typically consists of ones to allow for the intercept.

Bibliography

Abs, Hermann Josef, Martina Diedrich, Helge Sickmann, and Eckhard Klieme (2007). *Evaluation im BLK-Modellprogramm Demokratie Lernen und Leben: Skalen zur Befragung von Schüler/-Innen, Lehrer/-Innen und Schulleitungen. Dokumentation der Erhebungsinstrumente 2006*. Vol. 20. Materialien zur Bildungsforschung. Frankfurt/Main: German Institute for International Educational Research. URN: urn:nbn:de:0111-opus-19006.

Acosta, Pablo and Noël Muller (2018). "The Role of Cognitive and Socio-Emotional Skills in Labor Markets." In: *IZA World of Labor* 453. DOI: 10.15185/izawol.453.

Alchian, Armen A. and Harold Demsetz (1972). "Production, Information Costs, and Economic Organization." In: *American Economic Review* 62(5), pp. 777–795. JSTOR: 1815199.

Alem, Yonas, Håkan Eggert, Martin G. Kocher, and Remidius D. Ruhinduka (2018). "Why (Field) Experiments on Unethical Behavior Are Important: Comparing Stated and Revealed Behavior." In: *Journal of Economic Behavior & Organization* 156, pp. 71–85. DOI: 10.1016/j.jebo.2018.08.026.

Almås, Ingvild, Alexander W. Cappelen, Kjell G. Salvanes, Erik Ø. Sørensen, and Bertil Tungodden (2016). "Willingness to Compete: Family Matters." In: *Management Science* 62(8), pp. 2149–2162. DOI: 10.1287/mnsc.2015.2244.

Almond, Douglas and Janet Currie (2011). "Human Capital Development Before Age Five." In: *Handbook of Labor Economics*. Ed. by David Card and Orley Ashenfelter. Vol. 4. Amsterdam: Elsevier. Chap. 15, pp. 1315–1486. DOI: 10.1016/S0169-7218(11)02413-0.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). "Childhood Circumstances and Adult Outcomes: Act II." In: *Journal of Economic Literature* 56(4), pp. 1360–1446. DOI: 10.1257/jel.20171164.

Andreoni, James (1988). "Why Free Ride?: Strategies and Learning in Public Goods Experiments." In: *Journal of Public Economics* 37(3), pp. 291–304. DOI: 10.1016/0047-2727(88)90043-6.

Andreoni, James (1989). "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." In: *Journal of Political Economy* 97(6), pp. 1447–1458. DOI: 10.1086/261662.

Andreoni, James (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." In: *Economic Journal* 100(401), pp. 464–477. DOI: 10.2307/2234133.

Andreoni, James (1995). "Cooperation in Public-Goods Experiments: Kindness or Confusion?" In: *American Economic Review* 85(4), pp. 891–904. JSTOR: 2118238.

Andreoni, James (1998). "Toward a Theory of Charitable Fund-Raising." In: *Journal of Political Economy* 106(6), pp. 1186–1213. DOI: 10.1086/250044.

Andreoni, James and B. Douglas Bernheim (2009). "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects." In: *Econometrica* 77(5), pp. 1607–1636. DOI: 10.3982/ECTA7384.

Andreß, Hans-Jürgen, Katrin Golsch, and Alexander Schmidt (2013). *Applied Panel Data Analysis for Economic and Social Surveys*. Heidelberg: Springer. DOI: 10.1007/978-3-642-32914-2.

Ariely, Dan, Anat Bracha, and Stephan Meier (2009). "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." In: *American Economic Review* 99(1), pp. 544–555. DOI: 10.1257/aer.99.1.544.

Balafoutas, Loukas, Florian Lindner, and Matthias Sutter (2012). "Sabotage in Tournaments: Evidence from a Natural Experiment." In: *Kyklos* 65(4), pp. 425–441. DOI: 10.1111/kykl.12000.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul (2013). "Team Incentives: Evidence from a Firm Level Experiment." In: *Journal of the European Economic Association* 11(5), pp. 1079–1114. DOI: 10.1111/jeea.12028.

Battigalli, Pierpaolo and Martin Dufwenberg (2009). "Dynamic Psychological Games." In: *Journal of Economic Theory* 144(1), pp. 1–35. DOI: 10.1016/j.jet.2008.01.004.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse (2012). "The Relationship Between Economic Preferences and Psychological Personality Measures." In: *Annual Review of Economics* 4(1), pp. 453–478. DOI: 10.1146/annurev-economics-080511-110922.

Becker, Gordon M., Morris H. Degroot, and Jacob Marschak (1964). "Measuring Utility by a Single-Response Sequential Method." In: *Behavioral Science* 9(3), pp. 226–232. DOI: 10.1002/bs.3830090304.

Bénabou, Roland and Jean Tirole (2006). "Incentives and Prosocial Behavior." In: *American Economic Review* 96(5), pp. 1652–1678. DOI: 10.1257/aer.96.5.1652.

Blau, David and Janet Currie (2006). "Pre-School, Day Care, and After-School Care: Who's Minding the Kids?" In: *Handbook of the Economics of Education*. Ed. by Eric Hanushek and Finis Welch. Vol. 2. Amsterdam: Elsevier. Chap. 20, pp. 1163–1278. DOI: 10.1016/S1574-0692(06)02020-4.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch (2014). "hroot: Hamburg Registration and Organization Online Tool." In: *European Economic Review* 71, pp. 117–120. DOI: 10.1016/j.euroecorev.2014.07.003.

Booth, Alison L. (2009). "Gender and Competition." In: *Labour Economics* 16(6), pp. 599–606. DOI: 10.1016/j.labeco.2009.08.002.

Booth, Alison and Patrick Nolen (2012). "Choosing to Compete: How Different are Girls and Boys?" In: *Journal of Economic Behavior & Organization* 81(2), pp. 542–555. DOI: 10.1016/j.jebo.2011.07.018.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel (2008). "The Economics and Psychology of Personality Traits." In: *Journal of Human Resources* 43(4), pp. 972–1059. DOI: 10.3368/jhr.43.4.972.

Bortz, Jürgen and Gustav A. Lienert (2008). *Kurzgefasste Statistik für die klinische Forschung. Leitfaden für die verteilungsfreie Analyse kleiner Stichproben.* Heidelberg: Springer Medizin Verlag. DOI: 10.1007/978-3-540-75738-2.

Bortz, Jürgen, Gustav A. Lienert, and Klaus Boehnke (2008). *Verteilungsfreie Methoden in der Biostatistik.* Heidelberg: Springer Medizin Verlag. DOI: 10.1007/978-3-540-74707-9.

Bös, Klaus, Annette Worth, Elke Opper, Jennifer Oberger, and Alexander Woll (2009). *Motorik-Modul: Eine Studie zur motorischen Leistungsfähigkeit und körperlich-sportlichen Aktivität von Kindern und Jugendlichen in Deutschland.* Vol. 5. Forschungsreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend. Baden-Baden: Nomos. URL: https://www.bmfsfj.de/blob/94390/dc4ceb29b7415827c48a6a313b224602/motorik-modul-data.pdf (Retrieved: Oct. 15, 2018).

Bowker, Albert H. (1948). "A Test for Symmetry in Contingency Tables." In: *Journal of the American Statistical Association* 43(244), pp. 572–574. DOI: 10.1080/01621459.1948.10483284.

Bradley, James V. (1968). *Distribution-Free Statistical Tests.* Englewood Cliffs, NJ: Prentice-Hall.

Brandts, Jordi and David J. Cooper (2006a). "A Change Would Do You Good.... An Experimental Study on How to Overcome Coordination Failure in Organizations." In: *American Economic Review* 96(3), pp. 669–693. DOI: 10.1257/aer.96.3.669.

Brandts, Jordi and David J. Cooper (2006b). "Observability and Overcoming Coordination Failure in Organizations: An Experimental Study." In: *Experimental Economics* 9(4), pp. 407–423. DOI: 10.1007/s10683-006-7056-5.

Brandts, Jordi, David J. Cooper, and Enrique Fatas (2007). "Leadership and Overcoming Coordination Failure with Asymmetric Costs." In: *Experimental Economics* 10(3), pp. 269–284. DOI: 10.1007/s10683-007-9182-0.

Breuer, Christoph, Pamela Wicker, Sören Dallmeyer, and Jiři Dvořák (2016). *The Economic "Return on Investment" in Physical Education, Physical Activity and Sport.* Working Paper. German Sport University Cologne. DOI: 10.13140/RG.2.2.32330.57283.

Cabane, Charlotte and Andrew E. Clark (2015). "Childhood Sporting Activities and Adult Labour-Market Outcomes." In: *Annals of Economics and Statistics* 119/120, pp. 123–148. DOI: 10.15609/annaeconstat2009.119-120.123.

Cabane, Charlotte and Michael Lechner (2016). "Physical Activity of Adults: A Survey of Correlates, Determinants, and Effects." In: *Journal of Economics and Statistics* 235(4-5), pp. 376–402. DOI: 10.1515/jbnst-2015-4-504.

Cadsby, Charles Bram and Elizabeth Maynes (1999). "Voluntary Provision of Threshold Public Goods with Continuous Contributions: Experimental Evidence." In: *Journal of Public Economics* 71(1), pp. 53–73. DOI: 10.1016/S0047-2727(98)00049-8.

Camerer, Colin F. (2003). *Behavioral Game Theory. Experiments in Strategic Interaction.* Princeton, NJ: Princeton University Press.

Cameron, A. Colin and Pravin K. Trivedi (2009). *Microeconometrics Using Stata.* College Station, TX: Stata Press.

Cappelen, Alexander W., Trond Halvorsen, Erik Ø. Sørensen, and Bertil Tungodden (2017). "Face-Saving or Fair-Minded: What Motivates Moral Behavior?" In: *Journal of the European Economic Association* 15(3), pp. 540–557. DOI: 10.1093/jeea/jvw014.

Cappelen, Alexander W., Bjørn-Atle Reme, Erik Ø. Sørensen, and Bertil Tungodden (2016). "Leadership and Incentives." In: *Management Science* 62(7), pp. 1944–1953. DOI: 10.1287/mnsc.2015.2225.

Carpenter, Jeffrey, Peter Hans Matthews, and John Schirm (2010). "Tournaments and Office Politics: Evidence from a Real Effort Experiment." In: *American Economic Review* 100(1), pp. 504–517. DOI: 10.1257/aer.100.1.504.

Cartwright, Edward, Joris Gillet, and Mark Van Vugt (2013). "Leadership by Example in the Weak-Link Game." In: *Economic Inquiry* 51(4), pp. 2028–2043. DOI: 10.1111/ecin.12003.

Cartwright, Edward and Matheus L.C. Menezes (2014). "Cheating to Win: Dishonesty and the Intensity of Competition." In: *Economics Letters* 122(1), pp. 55–58. DOI: 10.1016/j.econlet.2013.10.016.

Caucutt, Elizabeth M., Lance Lochner, and Youngmin Park (2017). "Correlation, Consumption, Confusion, or Constraints: Why Do Poor Children Perform so Poorly?" In: *Scandinavian Journal of Economics* 119(1), pp. 102–147. DOI: 10.1111/sjoe.12195.

Celse, Jérémy, Michel Nicolas, and Pierre Schilling (2017). "Are Athletes More Cooperative Than Nonathletes? A Laboratory Experiment." In: *Managerial and Decision Economics* 38(8), pp. 1248–1261. DOI: 10.1002/mde.2862.

Charness, Gary and Martin Dufwenberg (2006). "Promises and Partnership." In: *Econometrica* 74(6), pp. 1579–1601. DOI: 10.1111/j.1468-0262.2006.00719.x.

Chaudhuri, Ananish (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." In: *Experimental Economics* 14(1), pp. 47–83. DOI: 10.1007/s10683-010-9257-1.

Chowdhury, Subhasish M. and Oliver Gürtler (2015). "Sabotage in Contests: A Survey." In: *Public Choice* 164(1), pp. 135–155. DOI: 10.1007/s11127-015-0264-9.

Cipriani, Marco, Paola Giuliano, and Olivier Jeanne (2013). "Like Mother Like Son? Experimental Evidence on the Transmission of Values from Parents to Children." In: *Journal of Economic Behavior & Organization* 90, pp. 100–111. DOI: 10.1016/j.jebo.2013.03.002.

Coase, Ronald H. (1937). "The Nature of the Firm." In: *Economica* 4(16), pp. 386–405. DOI: 10.1111/j.1468-0335.1937.tb00002.x.

Coats, Jennifer C., Timothy J. Gronberg, and Brit Grosskopf (2009). "Simultaneous versus Sequential Public Good Provision and the Role of Refunds—An Experimental Study." In: *Journal of Public Economics* 93(1), pp. 326–335. DOI: 10.1016/j.jpubeco.2008.06.002.

Cobb-Clark, Deborah A. (2015). "Locus of Control and the Labor Market." In: *IZA Journal of Labor Economics* 4(3). DOI: 10.1186/s40172-014-0017-x.

Corazzini, Luca, Christopher Cotton, and Paola Valbonesi (2015). "Donor Coordination in Project Funding: Evidence from a Threshold Public Goods Experiment." In: *Journal of Public Economics* 128, pp. 16–29. DOI: 10.1016/j.jpubeco.2015.05.005.

Cornelißen, Thomas and Christian Pfeifer (2009). "Sport und Arbeitseinkommen—Individuelle Ertragsraten von Sportaktivitäten in Deutschland." In: *Review of Economics* 59(3), pp. 344–255. DOI: 10.1515/roe-2008-0305.

Cragg, John G. (1971). "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." In: *Econometrica* 39(5), pp. 829–844. JSTOR: 1909582.

Croson, Rachel T.A. and Melanie Beth Marks (2000). "Step Returns in Threshold Public Goods: A Meta- and Experimental Analysis." In: *Experimental Economics* 2(3), pp. 239–259. DOI: 10.1023/A:1009918829192.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov (2006). "Interpreting the Evidence on Life Cycle Skill Formation." In: *Handbook of the Economics of Education*. Ed. by Eric Hanushek and Finis Welch. Vol. 1. Amsterdam: Elsevier. Chap. 12, pp. 697–812. DOI: 10.1016/S1574-0692(06)01012-9.

Currie, Janet (2001). "Early Childhood Education Programs." In: *Journal of Economic Perspectives* 15(2), pp. 213–238. DOI: 10.1257/jep.15.2.213.

Dana, Jason, Daylian M. Cain, and Robyn M. Dawes (2006). "What You Don't Know Won't Hurt Me: Costly (But Quiet) Exit in Dictator Games." In: *Organizational Behavior and Human Decision Processes* 100(2), pp. 193–201. DOI: 10.1016/j.obhdp.2005.10.001.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang (2007). "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." In: *Economic Theory* 33(1), pp. 67–80. DOI: 10.1007/s00199-006-0153-z.

Dannenberg, Astrid (2015). "Leading by Example versus Leading by Words in Voluntary Contribution Experiments." In: *Social Choice and Welfare* 44(1), pp. 71–85. DOI: 10.1007/s00355-014-0817-8.

Dato, Simon and Petra Nieken (2014). "Gender Differences in Competition and Sabotage." In: *Journal of Economic Behavior & Organization* 100, pp. 64–80. DOI: 10.1016/j.jebo.2014.01.012.

De Hooge, Ilona E., Marcel Zeelenberg, and Seger M. Breugelmans (2007). "Moral Sentiments and Cooperation: Differential Influences of Shame and Guilt." In: *Cognition and Emotion* 21(5), pp. 1025–1042. DOI: 10.1080/02699930600980874.

DellaVigna, Stefano, John A. List, and Ulrike Malmendier (2012). "Testing for Altruism and Social Pressure in Charitable Giving." In: *Quarterly Journal of Economics* 127(1), pp. 1–56. DOI: 10.1093/qje/qjr050.

Deming, David J. (2017). "The Growing Importance of Social Skills in the Labor Market." In: *Quarterly Journal of Economics* 132(4), pp. 1593–1640. DOI: 10.1093/qje/qjx022.

Destatis, ed. (2017). *Statistisches Jahrbuch 2017*. Wiesbaden: Federal Statistical Office. URL: https://www.destatis.de/DE/Publikationen/StatistischesJahrbuch/StatistischesJahrbuch2017.pdf?___blob=publicationFile (Retrieved: Oct. 15, 2018).

Dinno, Alexis (2015). "Nonparametric Pairwise Multiple Comparisons in Independent Groups Using Dunn's Test." In: *Stata Journal* 15(1), pp. 292–300. DOI: 10.1177/1536867X1501500117.

DIW/SOEP, ed. (2014). *Erhebungsinstrumente 2013 (Welle 30) des Sozio-oekonomischen Panels: Personenfragebogen, Altersstichproben.* SOEP Survey Papers 180. HDL: 10419/96127.

DIW/SOEP, ed. (2015a). *Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Haushaltsfragebogen, Altstichproben.* SOEP Survey Papers 275. HDL: 10419/123290.

DIW/SOEP, ed. (2015b). *Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Personenfragebogen, Altstichproben.* SOEP Survey Papers 274. HDL: 10419/123294.

DIW/SOEP, ed. (2015c). *Erhebungsinstrumente 2015 (Welle 32) des Sozio-oekonomischen Panels: Schülerinnen und Schüler (11–12 Jahre), Altstichproben.* SOEP Survey Papers 286. HDL: 10419/125862.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2011). "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." In: *Journal of the European Economic Association* 9(3), pp. 522–550. DOI: 10.1111/j.1542-4774.2011.01015.x.

Dong, Lu, Maria Montero, and Alex Possajennikov (2018). "Communication, Leadership and Coordination Failure." In: *Theory and Decision* 84(4), pp. 557–584. DOI: 10.1007/s11238-017-9617-9.

Dunn, Olive Jean (1964). "Multiple Comparisons Using Rank Sums." In: *Technometrics* 6(3), pp. 241–252. DOI: 10.1080/00401706.1964.10490181.

Eckel, Catherine C. and Philip J. Grossman (2005). "Managing Diversity by Creating Team Identity." In: *Journal of Economic Behavior & Organization* 58(3), pp. 371–392. DOI: 10.1016/j.jebo.2004.01.003.

Eder, Ferdinand (1998). *Linzer Fragebogen zum Schul- und Klassenklima für die 8. bis 13. Klasse (LFSK 8–13).* Göttingen: Hogrefe.

Englmaier, Florian and Georg Gebhardt (2016). "Social Dilemmas in the Laboratory and in the Field." In: *Journal of Economic Behavior & Organization* 128, pp. 85–96. DOI: 10.1016/j.jebo.2016.03.006.

Enste, Dominik, Mara Grunewald, and Louisa Kürten (2018). "Vertrauenskultur als Wettbewerbsvorteil in digitalen Zeiten: Neue experimentelle und verhaltensökonomische Ergebnisse." In: *IW-Trends—Vierteljahresschrift zur empirischen Wirtschaftsforschung* 45(2). DOI: 10.2373/1864-810X.18-02-04.

Erev, Ido and Amnon Rapoport (1990). "Provision of Step-Level Public Goods: The Sequential Contribution Mechanism." In: *Journal of Conflict Resolution* 34(3), pp. 401–425. DOI: 10.1177/0022002790034003002.

Erkal, Nisvan, Lata Gangadharan, and Nikos Nikiforakis (2011). "Relative Earnings and Giving in a Real-Effort Experiment." In: *American Economic Review* 101(7), pp. 3330–3348. DOI: 10.1257/aer.101.7.3330.

European Commission, ed. (2016). *Expert Group on Human Resources Development in Sport: Recommendations on the Contribution of Sport to the Employability of Young People, Including Young Professional Sportsmen and Women, and the Creation of Jobs in the Sport and Sport-Related Labour Market*. URL: http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetailDoc&id=31296 (Retrieved: July 17, 2018).

European Commission, ed. (2017). *Study on the Contribution of Sport to the Employability of Young People in the Context of the Europe 2020 Strategy: Final Report*. URL: https://ec.europa.eu/sport/news/study-contribution-sport-employability-young-people-published_en (Retrieved: July 17, 2018).

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde (2016). *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences*. IZA Discussion Paper 9674. RePEc: iza:izadps:dp9674.

Fan, Chinn-Ping (2000). "Teaching Children Cooperation—An Application of Experimental Game Theory." In: *Journal of Economic Behavior & Organization* 41(3), pp. 191–209. DOI: 10.1016/S0167-2681(99)00072-4.

Farrington, Camille A., Melissa Roderick, Elaine Allensworth, Jenny Nagaoka, Tasha Seneca Keyes, David W. Johnson, and Nicole O. Beechum (2012). *Teaching Adolescents to Become Learners. The Role of Noncognitive Factors in Shaping School Performance—A Critical Literature Review*. Chicago, IL: University of Chicago Consortium on Chicago School Research. URL: https://consortium.uchicago.edu/sites/default/files/publications/Noncognitive%20Report.pdf (Retrieved: Nov. 6, 2018).

Felfe, Christina, Michael Lechner, and Andreas Steinmayr (2016). "Sports and Child Development." In: *PLOS ONE* 11(5), art.-no. e0151729. DOI: 10.1371/journal.pone.0151729.

Fenigstein, Allan, Michael F. Scheier, and Arnold H. Buss (1975). "Public and Private Self-Consciousness: Assessment and Theory." In: *Journal of Consulting and Clinical Psychology* 43(4), pp. 522–527. DOI: 10.1037/h0076760.

Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." In: *Experimental Economics* 10(2), pp. 171–178. DOI: 10.1007/s10683-006-9159-4.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." In: *Economics Letters* 71(3), pp. 397–404. DOI: 10.1016/S0165-1765(01)00394-9.

Fryer, Roland G. (2017). "The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments." In: *Handbook of Economic Field Experiments*. Ed. by Abhijit Vinayak Banerjee and Esther Duflo. Vol. 2. Amsterdam: Elsevier. Chap. 2, pp. 95–322. DOI: 10.1016/bs.hefe.2016.08.006.

Gächter, Simon, Daniele Nosenzo, Elke Renner, and Martin Sefton (2010). "Sequential vs. Simultaneous Contributions to Public Goods: Experimental Evidence." In: *Journal of Public Economics* 94(7), pp. 515–522. DOI: 10.1016/j.jpubeco.2010.03.002.

Gächter, Simon, Daniele Nosenzo, Elke Renner, and Martin Sefton (2012). "Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example." In: *Economic Inquiry* 50(4), pp. 953–967. DOI: 10.1111/j.1465-7295.2010.00295.x.

Gächter, Simon and Elke Renner (2018). "Leaders as Role Models and 'Belief Managers' in Social Dilemmas." In: *Journal of Economic Behavior & Organization* 154, pp. 321–334. DOI: 10.1016/j.jebo.2018.08.001.

Galizzi, Matteo M. and Daniel Navarro-Martinez (2018). "On the External Validity of Social Preference Games: A Systematic Lab-Field Study." In: *Management Science*. DOI: 10.1287/mnsc.2017.2908. Forthcoming.

Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989). "Psychological Games and Sequential Rationality." In: *Games and Economic Behavior* 1(1), pp. 60–79. DOI: 10.1016/0899-8256(89)90005-5.

German Bundestag, ed. (2014). *13. Sportbericht der Bundesregierung*. Bundestagsdrucksache 18/3523 (Parliament Paper). URL: http://dip21.bundestag.de/dip21/btd/18/035/1803523.pdf (Retrieved: Oct. 15, 2018).

German Bundestag, ed. (2017). *15. Kinder- und Jugendbericht: Bericht über die Lebenssituation junger Menschen und die Leistungen der Kinder- und Jugendhilfe in Deutschland*. Bundestagsdrucksache 18/11050 (Parliament Paper). URL: http://dipbt.bundestag.de/doc/btd/18/110/1811050.pdf (Retrieved: Oct. 15, 2018).

Glöckner, Andreas, Bernd Irlenbusch, Sebastian Kube, Andreas Nicklisch, and Hans-Theo Normann (2011). "Leading with(out) Sacrifice? A Public-Goods Experiment with a Privileged Player." In: *Economic Inquiry* 49(2), pp. 591–597. DOI: 10.1111/j.1465-7295.2010.00314.x.

Gneezy, Uri, Andreas Leibbrandt, and John A. List (2016). "Ode to the Sea: Workplace Organizations and Norms of Cooperation." In: *Economic Journal* 126(595), pp. 1856–1883. DOI: 10.1111/ecoj.12209.

Golman, Russell, David Hagmann, and George Loewenstein (2017). "Information Avoidance." In: *Journal of Economic Literature* 55(1), pp. 96–135. DOI: 10.1257/jel.20151245.

Greenberg, Adam Eric, Paul Smeets, and Lilia Zhurakhovska (2015). *Promoting Truthful Communication through Ex-Post Disclosure*. SSRN Working Paper 2544349. DOI: 10.2139/ssrn.2544349.

Greene, William H. (2012). *Econometric Analysis*. Boston, MA: Pearson.

Greene, William H. and David A. Hensher (2010). *Modeling Ordered Choices. A Primer*. Cambridge: Cambridge University Press.

Greiner, Ben (2015). "Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE." In: *Journal of the Economic Science Association* 1(1), pp. 114–125. DOI: 10.1007/s40881-015-0004-4.

Grossman, Zachary (2014). "Strategic Ignorance and the Robustness of Social Preferences." In: *Management Science* 60(11), pp. 2659–2665. DOI: 10.1287/mnsc.2014.1989.

Güth, Werner, M. Vittoria Levati, Matthias Sutter, and Eline van der Heijden (2007). "Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments." In: *Journal of Public Economics* 91(5), pp. 1023–1042. DOI: 10.1016/j.jpubeco.2006.10.007.

Gutman, Leslie Morrison and Ingrid Schoon (2013). *The Impact of Non-Cognitive Skills on Outcomes for Young People: Literature Review*. London: Education Endowment Foundation. VOECD: 10707/287500.

Haas, André (2019). *First Movers in a Threshold Public Good Game*. Mimeo. Karlsruhe Institute of Technology.

Haas, André, Sandra Ludwig, and Petra Nieken (2019a). *Guilt and Shame*. Mimeo. Karlsruhe Institute of Technology.

Haas, André, Petra Nieken, Hagen Wäsche, Rita Wittelsberger, and Alexander Woll (2019b). *Cooperation in the Classroom*. Mimeo. Karlsruhe Institute of Technology.

Haas, André, Petra Nieken, Hagen Wäsche, Rita Wittelsberger, and Alexander Woll (2019c). *Study Protocol for the Movigen Project*. Mimeo. Karlsruhe Institute of Technology.

Harbaugh, William T. and Kate Krause (2000). "Children's Altruism in Public Good and Dictator Experiments." In: *Economic Inquiry* 38(1), pp. 95–109. DOI: 10.1111/j.1465-7295.2000.tb00006.x.

Harbring, Christine and Bernd Irlenbusch (2004). "Anreize zu produktiven und destruktiven Anstrengungen durch relative Entlohnung." In: *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* 56(6), pp. 546–576. DOI: 10.1007/BF03372749.

Harbring, Christine and Bernd Irlenbusch (2005). "Incentives in Tournaments with Endogenous Prize Selection." In: *Journal of Institutional and Theoretical Economics* 161(4), pp. 636–663. DOI: 10.1628/093245605775075951.

Harbring, Christine and Bernd Irlenbusch (2008). "How Many Winners Are Good to Have?: On Tournaments with Sabotage." In: *Journal of Economic Behavior & Organization* 65(3), pp. 682–702. DOI: 10.1016/j.jebo.2006.03.004.

Harbring, Christine and Bernd Irlenbusch (2011). "Sabotage in Tournaments: Evidence from a Laboratory Experiment." In: *Management Science* 57(4), pp. 611–627. DOI: 10.1287/mnsc.1100.1296.

Harbring, Christine, Bernd Irlenbusch, Matthias Kräkel, and Reinhard Selten (2007). "Sabotage in Corporate Contests—An Experimental Analysis." In: *International Journal of the Economics of Business* 14(3), pp. 367–392. DOI: 10.1080/13571510701597445.

Harbring, Christine and Jan Wilhelm (2016). *Deliberate Harming or Failing to Act?—Active and Passive Sabotage in Contests*. Mimeo. RWTH Aachen University.

Harsanyi, John C. and Reinhard Selten (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

Heckman, James J. (2000). "Policies to Foster Human Capital." In: *Research in Economics* 54(1), pp. 3–56. DOI: 10.1006/reec.1999.0225.

Heckman, James J. and Tim Kautz (2012). "Hard Evidence on Soft Skills." In: *Labour Economics* 19(4), pp. 451–464. DOI: 10.1016/j.labeco.2012.05.014.

Heinemann, Wolfgang (1979). "The Assessment of Private and Public Self-Consciousness: A German Replication." In: *European Journal of Social Psychology* 9(3), pp. 331–337. DOI: 10.1002/ejsp.2420090311.

Heinz, Matthias and Heiner Schumacher (2017). "Signaling Cooperation." In: *European Economic Review* 98, pp. 199–216. DOI: 10.1016/j.euroecorev.2017.06.017.

Hermalin, Benjamin E. (1998). "Toward an Economic Theory of Leadership: Leading by Example." In: *American Economic Review* 88(5), pp. 1188–1206. JSTOR: 116866.

Hermes, Henning, Florian Hett, Mario Mechtel, Felix Schmidt, Daniel Schunk, and Valentin Wagner (2019). "Do Children Cooperate Conditionally? Adapting the Strategy Method for First-Graders." In: *Journal of Economic Behavior & Organization*. DOI: 10.1016/j.jebo.2018.12.032. Forthcoming.

Hirshleifer, Jack (1983). "From Weakest-Link to Best-Shot: The Voluntary Provision of Public Goods." In: *Public Choice* 41(3), pp. 371–386. DOI: 10.1007/BF00141070.

Holm, Sture (1979). "A Simple Sequentially Rejective Multiple Test Procedure." In: *Scandinavian Journal of Statistics* 6(2), pp. 65–70. JSTOR: 4615733.

Howard-Jones, Paul A., Elizabeth V. Washbrook, and Sara Meadows (2012). "The Timing of Educational Investment: A Neuroscientific Perspective." In: *Developmental Cognitive Neuroscience* 2, Supplement 1, S18–S29. DOI: 10.1016/j.dcn.2011.11.002.

Isaac, R. Mark, David Schmidtz, and James M. Walker (1989). "The Assurance Problem in a Laboratory Market." In: *Public Choice* 62(3), pp. 217–236. DOI: 10.1007/BF02337743.

Jerusalem, Matthias, Stephanie Drössler, Dietmar Kleine, Johannes Klein-Heßling, Waldemar Mittag, and Bettina Röder (2009). *Förderung von Selbstwirksamkeit und Selbstbestimmung im Unterricht. Skalen zur Erfassung von Lehrer- und Schülermerkmalen.* Berlin: Department of Education Studies, Humboldt University. URL: https://www.erziehungswissenschaften.hu-berlin.de/de/paedpsych/forschung/Skalenbuch_FoSS.pdf (Retrieved: Oct. 12, 2018).

John, Katrin and Stephan L. Thomsen (2015). "School-Track Environment or Endowment: What Determines Different Other-Regarding Behavior across Peer Groups?" In: *Games and Economic Behavior* 94, pp. 122–141. DOI: 10.1016/j.geb.2015.10.007.

Kajackaite, Agne (2015). "If I Close My Eyes, Nobody Will Get Hurt: The Effect of Ignorance on Performance in a Real-Effort Experiment." In: *Journal of Economic Behavior & Organization* 116, pp. 518–524. DOI: 10.1016/j.jebo.2015.05.020.

Kandel, Eugene and Edward P. Lazear (1992). "Peer Pressure and Partnerships." In: *Journal of Political Economy* 100(4), pp. 801–817. DOI: 10.1086/261840.

Kaplan, Todd R., Bradley J. Ruffle, and Ze'ev Shtudiner (2018). "Cooperation through Coordination in Two Stages." In: *Journal of Economic Behavior & Organization* 154, pp. 206–219. DOI: 10.1016/j.jebo.2018.08.008.

Kautz, Tim, James J. Heckman, Ron Diris, Bas ter Weel, and Lex Borghans (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success.* NBER Working Paper 20749. DOI: 10.3386/w20749.

Kerschbamer, Rudolf (2015). "The Geometry of Distributional Preferences and a Non-Parametric Identification Approach: The Equality Equivalence Test." In: *European Economic Review* 76, pp. 85–103. DOI: 10.1016/j.euroecorev.2015.01.008.

Keser, Claudia and Frans van Winden (2000). "Conditional Cooperation and Voluntary Contributions to Public Goods." In: *Scandinavian Journal of Economics* 102(1), pp. 23–39. DOI: 10.1111/1467-9442.00182.

Kim, Oliver and Mark Walker (1984). "The Free Rider Problem: Experimental Evidence." In: *Public Choice* 43(1), pp. 3–24. DOI: 10.1007/BF00137902.

Klinowski, David (2018). "Gender Differences in Giving in the Dictator Game: The Role of Reluctant Altruism." In: *Journal of the Economic Science Association* 4(2), pp. 110–122. DOI: 10.1007/s40881-018-0058-1.

Knaus, Michael C., Michael Lechner, and Anne K. Reimers (2018). *For Better or Worse? The Effects of Physical Education on Child Development*. IZA Discussion Paper 11268. RePEc: iza:izadps:dp11268.

Konow, James (2000). "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." In: *American Economic Review* 90(4), pp. 1072–1091. DOI: 10.1257/aer.90.4.1072.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk (2018). *The Formation of Prosociality: Causal Evidence on the Role of Social Environment*. CESifo Working Paper 7068. RePEc: ces:ceswps:_7068.

Kruskal, William H. (1952). "A Nonparametric Test for the Several Sample Problem." In: *Annals of Mathematical Statistics* 23(4), pp. 525–540. DOI: 10.1214/aoms/1177729332.

Kruskal, William H. (1957). "Historical Notes on the Wilcoxon Unpaired Two-Sample Test." In: *Journal of the American Statistical Association* 52(279), pp. 356–360. DOI: 10.1080/01621459.1957.10501395.

Kruskal, William H. and W. Allen Wallis (1952). "Use of Ranks in One-Criterion Variance Analysis." In: *Journal of the American Statistical Association* 47(260), pp. 583–621. DOI: 10.1080/01621459.1952.10483441.

Kunter, Mareike, Gundel Schümer, Cordula Artelt, Jürgen Baumert, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Petra Stanat, Klaus-Jürgen Tilmann, and Manfred Weiß (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Vol. 72. Materialien aus der Bildungsforschung. Berlin: Max-Planck-Institute for Educational Research. eDoc: 14414.

Kurth, Bärbel-Maria (2007). "Der Kinder- und Jugendgesundheitssurvey (KiGGS): Ein Überblick über Planung, Durchführung und Ergebnisse unter Berücksichtigung von Aspekten eines Qualitätsmanagements." In: *Bundesgesundheitsblatt—Gesundheitsforschung—Gesundheitsschutz* 50(5), pp. 533–546. DOI: 10.1007/s00103-007-0214-x.

Langmeyer, Alexandra and Ursula Winklhofer (2014). *Taschengeld und Gelderziehung—Eine Expertise zum Thema Kinder und ihr Umgang mit Geld mit aktualisierten Empfehlungen zum Taschengeld.* Munich: German Youth Institute. URL: https://www.dji.de/fileadmin/user_upload/bibs2014/DJI_Expertise_Taschengeld.pdf (Retrieved: Mar. 23, 2018).

Larson, Tara and C. Monica Capra (2009). "Exploiting Moral Wiggle Room: Illusory Preference for Fairness? A Comment." In: *Judgment and Decision Making* 4(6), pp. 467–474. RePEc: jdm:journl:v:4:y:2009:i:6:p:467-474.

Lazear, Edward P. (1989). "Pay Equality and Industrial Politics." In: *Journal of Political Economy* 97(3), pp. 561–580. DOI: 10.1086/261616.

Lazear, Edward P., Ulrike Malmendier, and Roberto A. Weber (2012). "Sorting in Experiments with Application to Social Preferences." In: *American Economic Journal: Applied Economics* 4(1), pp. 136–163. DOI: 10.1257/app.4.1.136.

Lazear, Edward P. and Paul Oyer (2012). "Personnel Economics." In: *The Handbook of Organizational Economics.* Ed. by Robert Gibbons and John Roberts. Princeton, NJ: Princeton University Press. Chap. 12, pp. 479–519.

Lazear, Edward P. and Sherwin Rosen (1981). "Rank-Order Tournaments as Optimum Labor Contracts." In: *Journal of Political Economy* 89(5), pp. 841–864. DOI: 10.1086/261010.

Lazear, Edward P. and Kathryn L. Shaw (2007). "Personnel Economics: The Economist's View of Human Resources." In: *Journal of Economic Perspectives* 21(4), pp. 91–114. DOI: 10.1257/jep.21.4.91.

Lechner, Michael (2009). "Long-Run Labour Market and Health Effects of Individual Sports Activities." In: *Journal of Health Economics* 28(4), pp. 839–854. DOI: 10.1016/j.jhealeco.2009.05.003.

Lechner, Michael (2015). "Sports, Exercise, and Labor Market Outcomes." In: *IZA World of Labor* 126. DOI: 10.15185/izawol.126.

Lechner, Michael and Nazmi Sari (2015). "Labor Market Effects of Sports and Exercise: Evidence from Canadian Panel Data." In: *Labour Economics* 35, pp. 1–15. DOI: 10.1016/j.labeco.2015.04.001.

Ledyard, John O. (1995). "Public Goods: A Survey of Experimental Research." In: *Handbook of Experimental Economics*. Ed. by John H. Kagel and Alvin E. Roth. Princeton, NJ: Princeton University Press. Chap. 2, pp. 111–194.

Levati, M. Vittoria, Matthias Sutter, and Eline van der Heijden (2007). "Leading by Example in a Public Goods Experiment with Heterogeneity and Incomplete Information." In: *Journal of Conflict Resolution* 51(5), pp. 793–818. DOI: 10.1177/0022002707302796.

Long, James E. and Steven B. Caudill (1991). "The Impact of Participation in Intercollegiate Athletics on Income and Graduation." In: *Review of Economics and Statistics* 73(3), pp. 525–531. JSTOR: 2109580.

Luce, Robert Duncan and Howard Raiffa (1957). *Games and Decisions. Introduction and Critical Survey*. New York, NY: Wiley.

Mann, Harold B. and Donald R. Whitney (1947). "On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other." In: *Annals of Mathematical Statistics* 18(1), pp. 50–60. DOI: 10.1214/aoms/1177730491.

Matthey, Astrid and Tobias Regner (2011). "Do I Really Want to Know? A Cognitive Dissonance-Based Explanation of Other-Regarding Behavior." In: *Games* 2(1), pp. 114–135. DOI: 10.3390/g2010114.

Mazar, Nina, On Amir, and Dan Ariely (2008). "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." In: *Journal of Marketing Research* 45(6), pp. 633–644. DOI: 10.1509/jmkr.45.6.633.

Meidinger, Claude and Marie-Claire Villeval (2002). *Leadership in Teams: Signaling or Reciprocating?* GATE Working Paper 2002-13. HAL: halshs-00178474.

Milinski, Manfred, Dirk Semmann, Hans-Jürgen Krambeck, and Jochem Marotzke (2006). "Stabilizing the Earth's Climate Is Not a Losing Game: Supporting Evidence from Public Goods Experiments." In: *Proceedings of the National Academy of Sciences* 103(11), pp. 3994–3998. DOI: 10.1073/pnas.0504902103.

Moxnes, Erling and Eline van der Heijden (2003). "The Effect of Leadership in a Public Bad Experiment." In: *Journal of Conflict Resolution* 47(6), pp. 773–795. DOI: 10.1177/0022002703258962.

Murnighan, J. Keith, Jae Wook Kim, and A. Richard Metzger (1993). "The Volunteer Dilemma." In: *Administrative Science Quarterly* 38(4), pp. 515–538. JSTOR: 2393335.

Nalbantian, Haig R. and Andrew Schotter (1997). "Productivity under Group Incentives: An Experimental Study." In: *American Economic Review* 87(3), pp. 314–341. JSTOR: 2951348.

Niederle, Muriel and Lise Vesterlund (2007). "Do Women Shy Away from Competition? Do Men Compete Too Much?" In: *Quarterly Journal of Economics* 122(3), pp. 1067–1101. DOI: 10.1162/qjec.122.3.1067.

Normann, Hans-Theo and Holger A. Rau (2015). "Simultaneous and Sequential Contributions to Step-Level Public Goods: One versus Two Provision Levels." In: *Journal of Conflict Resolution* 59(7), pp. 1273–1300. DOI: 10.1177/0022002714530429.

Olson, Mancur (1965). *The Logic of Collective Action. Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Organization for Economic Cooperation and Development, ed. (2015). *Skills for Social Progress. The Power of Social and Emotional Skills*. OECD Skills Studies. Paris: OECD Publishing. DOI: 10.1787/9789264226159-en.

Palfrey, Thomas R. and Jeffrey E. Prisbrey (1997). "Anomalous Behavior in Public Goods Experiments: How Much and Why?" In: *American Economic Review* 87(5), pp. 829–846. JSTOR: 2951327.

Pawlowski, Tim, Ute Schüttoff, Paul Downward, and Michael Lechner (2018). "Can Sport Really Help to Meet the Millennium Development Goals? Evidence from Children in Peru." In: *Journal of Sports Economics* 19(4), pp. 498–521. DOI: 10.1177/1527002516661601.

Pearson, Karl (1900). "χ. On the Criterion That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling." In: *Philosophical Magazine* 50(302). Series 5, pp. 157–175. DOI: 10.1080/14786440009463897.

Pedersen, Bente Klarlund and Bengt Saltin (2015). "Exercise as Medicine—Evidence for Prescribing Exercise as Therapy in 26 Different Chronic Diseases." In: *Scandinavian Journal of Medicine & Science in Sports* 25(S3), pp. 1–72. DOI: 10.1111/sms.12581.

Physical Activity Guidelines Advisory Committee, ed. (2018). *2018 Physical Activity Guidelines Advisory Committee Scientific Report*. Washington, DC: U.S. Department of Health and Human Services. URL: https://health.gov/paguidelines/second-edition/report.aspx (Retrieved: Oct. 12, 2018).

Piopiunik, Marc, Guido Schwerdt, Lisa Simon, and Ludger Woessmann (2018). *Skills, Signals, and Employability: An Experimental Investigation*. IZA Discussion Paper 11283. RePEc: iza:izadps:dp11283.

Pogrebna, Ganna, David H. Krantz, Christian Schade, and Claudia Keser (2011). "Words versus Actions as a Means to Influence Cooperation in Social Dilemma Situations." In: *Theory and Decision* 71(4), pp. 473–502. DOI: 10.1007/s11238-011-9248-5.

Potters, Jan, Martin Sefton, and Lise Vesterlund (2005). "After You—Endogenous Sequencing in Voluntary Contribution Games." In: *Journal of Public Economics* 89(8), pp. 1399–1419. DOI: 10.1016/j.jpubeco.2004.02.008.

Potters, Jan, Martin Sefton, and Lise Vesterlund (2007). "Leading-by-Example and Signaling in Voluntary Contribution Games: An Experimental Study." In: *Economic Theory* 33(1), pp. 169–182. DOI: 10.1007/s00199-006-0186-3.

Prendergast, Canice (1999). "The Provision of Incentives in Firms." In: *Journal of Economic Literature* 37(1), pp. 7–63. DOI: 10.1257/jel.37.1.7.

Rasch, Björn, Malte Friese, Wilhelm Hofmann, and Ewald Naumann (2014a). *Quantitative Methoden 1. Einführung in die Statistik für Psychologen und Sozialwissenschaftler.* Berlin: Springer. DOI: 10.1007/978-3-662-43524-3.

Rasch, Björn, Malte Friese, Wilhelm Hofmann, and Ewald Naumann (2014b). *Quantitative Methoden 2. Einführung in die Statistik für Psychologen und Sozialwissenschaftler.* Berlin: Springer. DOI: 10.1007/978-3-662-43548-9.

Reiner, Miriam, Christina Niermann, Darko Jekauc, and Alexander Woll (2013). "Long-Term Health Benefits of Physical Activity—A Systematic Review of Longitudinal Studies." In: *BMC Public Health* 13(1), pp. 813–821. DOI: 10.1186/1471-2458-13-813.

Rooth, Dan-Olof (2011). "Work Out or Out of Work—The Labor Market Return to Physical Fitness and Leisure Sports Activities." In: *Labour Economics* 18(3), pp. 399–409. DOI: 10.1016/j.labeco.2010.11.006.

Roth, Alvin E. (1995). "An Introduction to Experimental Economics." In: *Handbook of Experimental Economics.* Ed. by John H. Kagel and Alvin E. Roth. Princeton, NJ: Princeton University Press. Chap. 1, pp. 3–109.

Sahin, Selhan Garip, Catherine Eckel, and Mana Komai (2015). "An Experimental Study of Leadership Institutions in Collective Action Games." In: *Journal of the Economic Science Association* 1(1), pp. 100–113. DOI: 10.1007/s40881-015-0010-6.

Samak, Anya C. (2013). "Is There a Gender Gap in Preschoolers' Competitiveness? An Experiment in the U.S." In: *Journal of Economic Behavior & Organization* 92, pp. 22–31. DOI: 10.1016/j.jebo.2013.04.014.

Sani, Fabi and John Todman (2006). *Experimental Design and Statistics for Psychology. A First Course.* Oxford: Blackwell. DOI: 10.1002/9780470776124.

Savikhin Samek, Anya and Roman M. Sheremeta (2014). "Recognizing Contributors: An Experiment on Public Goods." In: *Experimental Economics* 17(4), pp. 673–690. DOI: 10.1007/s10683-013-9389-1.

Schaich, Eberhard and Alfred Hamerle (1984). *Verteilungsfreie statistische Prüfverfahren. Eine anwendungsorientierte Darstellung.* Berlin: Springer.

Schmidt, Steffen C.E., Annette Henn, Claudia Albrecht, and Alexander Woll (2017). "Physical Activity of German Children and Adolescents 2003–2012: The MoMo-Study." In: *International Journal of Environmental Research and Public Health* 14(11), art.-no. 1375. DOI: 10.3390/ijerph14111375.

Schmidt, Steffen, Nadine Will, Annette Henn, Anne Reimers, and Alexander Woll (2016). *Der Motorik-Modul Aktivitätsfragebogen Momo-AFB : Leitfaden zur Anwendung und Auswertung.* KIT Scientific Working Papers 53. DOI: 10.5445/IR/1000062199.

Schüttoff, Ute, Tim Pawlowski, Paul Downward, and Michael Lechner (2017). "Sports Participation and Social Capital Formation During Adolescence." In: *Social Science Quarterly* 99(2), pp. 683–698. DOI: 10.1111/ssqu.12453.

Schwieren, Christiane and Doris Weichselbaumer (2010). "Does Competition Enhance Performance or Cheating? A Laboratory Experiment." In: *Journal of Economic Psychology* 31(3), pp. 241–253. DOI: 10.1016/j.joep.2009.02.005.

Selten, Reinhard (1967). "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperimentes." In: *Beiträge zur Experimentellen Wirtschaftsforschung.* Ed. by Heinz Sauermann. Vol. 1. Tübingen: Mohr Siebeck. Chap. 5, pp. 136–168.

Sheremeta, Roman M. (2016). "The Pros and Cons of Workplace Tournaments." In: *IZA World of Labor* 302. DOI: 10.15185/izawol.302.

Smither, Robert D. and John M. Houston (1992). "The Nature of Competitiveness: The Development and Validation of the Competitiveness Index." In: *Educational and Psychological Measurement* 52(2), pp. 407–418. DOI: 10.1177/0013164492052002016.

Spearman, Charles (1904). "The Proof and Measurement of Association between Two Things." In: *American Journal of Psychology* 15(1), pp. 72–101. JSTOR: 1412159.

Spearman, Charles (1906). "'Footrule' for Measuring Correlation." In: *British Journal of Psychology* 2(1), pp. 89–108. DOI: 10.1111/j.2044-8295.1906.tb00174.x.

Stock, James H. and Mark W. Watson (2015). *Introduction to Econometrics.* Boston, MA: Pearson.

Sutter, Matthias, Claudia Zoller, and Daniela Glätzle-Rützler (2019). "Economic Behavior of Children and Adolescents—A First Survey of Experimental Economics Results." In: *European Economic Review* 111, pp. 98–121. DOI: 10.1016/j.euroecorev.2018.09.004.

Tobin, James (1958). "Estimation of Relationships for Limited Dependent Variables." In: *Econometrica* 26(1), pp. 24–36. JSTOR: 1907382.

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." In: *American Economic Review* 80(1), pp. 234–248. JSTOR: 2006745.

Vandegrift, Donald and Abdullah Yavas (2010). "An Experimental Test of Sabotage in Tournaments." In: *Journal of Institutional and Theoretical Economics* 166(2), pp. 259–285. DOI: 10.1628/093245610791343012.

Vesterlund, Lise (2003). "The Informational Value of Sequential Fundraising." In: *Journal of Public Economics* 87(3), pp. 627–657. DOI: 10.1016/S0047-2727(01)00187-6.

Warburton, Darren E.R. and Shannon S.D. Bredin (2016). "Reflections on Physical Activity and Health: What Should We Recommend?" In: *Canadian Journal of Cardiology* 32(4), pp. 495–504. DOI: 10.1016/j.cjca.2016.01.024.

Warburton, Darren E.R., Crystal Whitney Nicol, and Shannon S.D. Bredin (2006). "Health Benefits of Physical Activity: The Evidence." In: *Canadian Medical Association Journal* 174(6), pp. 801–809. DOI: 10.1503/cmaj.051351.

Wartenberg, Josephine, Thomas Borchert, and Ralf Brand (2014). "A Longitudinal Assessment of Adolescent Student-Athletes' School Performance: (Not) Worse in School and (Not) Putting Their Education at Risk?!" In: *Sportwissenschaft* 44(2), pp. 78–85. DOI: 10.1007/s12662-014-0331-2.

Weber, Roberto, Colin Camerer, Yuval Rottenstreich, and Marc Knez (2001). "The Illusion of Leadership: Misattribution of Cause in Coordination Games." In: *Organization Science* 12(5), pp. 582–598. DOI: 10.1287/orsc.12.5.582.10090.

Weinhardt, Michael and Jürgen Schupp (2011). *Multi-Itemskalen im SOEP Jugendfragebogen.* SOEP Data Documentation 60. RePEc: diw:diwddc:dd60.

Wilcoxon, Frank (1945). "Individual Comparisons by Ranking Methods." In: *Biometrics Bulletin* 1(6), pp. 80–83. JSTOR: 3001968.

Woll, Alexander, Claudia Albrecht, and Annette Worth (2017). "Motorik-Modul (MoMo)—das Modul zur Erfassung der motorischen Leistungsfähigkeit und der körperlich-sportlichen

Aktivität in KiGGS Welle 2." In: *Journal of Health Monitoring* 2(S3), pp. 66–73. DOI: 10.17886/RKI-GBE-2017-104.

Woll, Alexander, Dominik Böttger, Rita Wittelsberger, André Haas, Petra Nieken, and Hagen Wäsche (2018). *Movigen: Ein Interventionskonzept zur Förderung sozialer Kompetenzen im Schulsport.* KIT Scientific Working Papers 100. DOI: 10.5445/IR/1000086350.

Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

World Health Organization, ed. (2010). *Global Recommendations on Physical Activity for Health.* Geneva: World Health Organization. HDL: 10665/44399.

Zou, Kelly H., Julia R. Fielding, Stuart G. Silverman, and Clare M.C. Tempany (2003). "Hypothesis Testing I: Proportions." In: *Radiology* 226(3), pp. 609–613. DOI: 10.1148/radiol.2263011500.