An Experimental Analysis of Leadership, Cooperation, Competition and Decision-Making

- Five Essays in Behavioral and Experimental Economics -

Dissertation

zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.)
der Fakultät I -Bildungs- und Gesellschaftswissenschaften der Universität Vechta

vorgelegt von
Torben Kölpin

Erstprüfer: Prof. Dr. Gerald Eisenkopf
Zweiprüferin: Prof. Dr. Vanessa Mertins

Frankfurt am Main
2022

Examination Board

Prof. Dr. Gerald Eisenkopf (University of Vechta)

Prof. Dr. Vanessa Mertins (University of Vechta)

Prof. Dr. Nick Lin-Hi (University of Vechta)

PD Dr. Stefan Klößner (University of Vechta)

Prof. Dr. Robert Gillenkirch (Osnabrück University)

Promotionskommission

Prof. Dr. Gerald Eisenkopf (Universität Vechta)

Prof. Dr. Vanessa Mertins (Universität Vechta)

Prof. Dr. Nick Lin-Hi (Universität Vechta)

PD Dr. Stefan Klößner (Universität Vechta)

Prof. Dr. Robert Gillenkirch (Universität Osnabrück)

Acknowledgements

This work would not have been accomplished without the cooperation and contribution of many important people. My first and most important thanks goes to my PhD supervisor Prof. Dr. Gerald Eisenkopf. Thank you for creating an intellectually supportive and challenging environment that encouraged me in my research projects. In particular, his expertise and constructive criticism made an important contribution to this work. In addition to his personal qualities, I appreciate his sincere interest in science as well as his curiosity and willingness to always question findings critically. Furthermore, I am very grateful for the collegial cooperation as well as for the good supervision and guidance during the entire PhD period. I could not have asked for a better advisor. I would also like to thank Prof. Dr. Vanessa Mertins for her insightful suggestions and comments throughout the various stages of development of this research. I greatly appreciate her helpful feedback as well as the several conversations and suggestions that helped me especially at the beginning of the dissertation.

My sincere thanks also apply to the doctoral committee. I would like to thank Gerald Eisenkopf and Vanessa Mertins as well as Robert Gillenkirch, Nick Lin-Hi, and Stefan Klößner for the flawless procedure, always friendly communication, and stimulating discussion during the disputation.

During this time, the support of my current and former colleagues with whom I had the fortune to share the path to a PhD was invaluable to me. They all became very good friends. With you, even the most stressful times were very enjoyable, and I am sure that the friendly atmosphere contributed significantly to the success of this work. I am very grateful to Bernd Josef Leisen, whose feedback and ideas were always very helpful. To Christian Walter for his passion questioning things critically and discussing them constructively. To Maximilian Hiller for the many conversations, scientific and private, which I appreciated not only on the professional but also on the private level. To Vivien Breitrück for her valuable contributions and help with various questions. In addition, I would also like to thank Devin Kwasniok, Nicole Middendorf and Maximilian Kuntze, even though our time together was quite short. Last, but not least, I would like to thank Petra Tecker. Her open and cheerful manner always created a good atmosphere at the chair. I am very grateful for her open door and her open ear.

A big thank you also goes to all my friends, who helped me to keep life interesting and varied aside from my studies and doctorate. Without this balance, completing this thesis would have been much more difficult for me. Special thanks go to my parents, Frank and Regine Kölpin, to my siblings (Karen, Sven, Lars, and Inga), to my grandparents (Eckard Kölpin, Erika and Reinhard Fiedler), and to my family-in-law (Erwin, Herta, and Irina Mattrisch, Mehran Boroumand, Lea Freese, Sabrina Ganse as well as Ingo Mindemann), who were with me through all the ups and downs, and who always supported me unconditionally. A special thanks goes to my brother Lars Kölpin-Freese, who always supported me and my colleagues with all kinds of technical questions and the subsequent implementation. For his expertise and patience, I am very grateful. I would also like to express my special thanks to my uncle Wolfram Fiedler, who helped me with questions in word and deed and for whose feedback I am very grateful.

The biggest thanks go to my lovely partner Lea Mattrisch, who supported me on my life's journey with all her strength and love and carried me through all my decisions. For everything she has done for me and made possible, I am infinitely grateful. Without her, writing this thesis would not have been possible.

Thank you very much.

Danksagung

Diese Arbeit wäre ohne die Mitarbeit und den Beitrag vieler wichtiger Personen nicht zustande gekommen. Mein erster und wichtigster Dank geht an meinen Doktorvater Prof. Dr. Gerald Eisenkopf. Vielen Dank für die Schaffung eines intellektuell unterstützenden und herausfordernden Umfelds, das mich in meinen Forschungsprojekten ermutigt hat. Insbesondere seine Expertise und konstruktive Kritik haben einen wichtigen Beitrag zu dieser Arbeit geleistet. Neben seinen persönlichen Qualitäten schätze ich sein aufrichtiges Interesse an der Wissenschaft sowie seine Neugierde und Bereitschaft, Erkenntnisse stets kritisch zu hinterfragen. Darüber hinaus bin ich sehr dankbar für die kollegiale Zusammenarbeit sowie für die gute Betreuung und Begleitung während der gesamten Promotionszeit. Einen besseren Betreuer hätte ich mir nicht wünschen können. Ich möchte auch Prof. Dr. Vanessa Mertins für ihre aufschlussreichen Vorschläge und Kommentare während der verschiedenen Entwicklungsphasen dieser Arbeit danken. Ich schätze ihr hilfreiches Feedback sowie die zahlreichen Gespräche und Anregungen, die mir vor allem zu Beginn der Dissertation geholfen haben, sehr.

Mein aufrichtiger Dank gilt auch dem Promotionsausschuss. Ich danke Gerald Eisenkopf und Vanessa Mertins sowie Robert Gillenkirch, Nick Lin-Hi und Stefan Klößner für den reibungslosen Ablauf, die stets freundliche Kommunikation und die anregende Diskussion während der Disputation.

Während dieser Zeit war die Unterstützung meiner jetzigen und ehemaligen Kolleg*innen, mit denen ich das Glück hatte, den Weg zur Promotion zu teilen, von unschätzbarem Wert für mich. Sie alle sind sehr gute Freunde geworden. Mit Ihnen waren selbst die stressigsten Zeiten sehr angenehm, und ich bin sicher, dass die freundschaftliche Atmosphäre wesentlich zum Erfolg dieser Arbeit beigetragen hat. Ich bin Bernd Josef Leisen sehr dankbar, dessen Feedback und Ideen immer sehr hilfreich waren. Christian Walter für seine Leidenschaft, Dinge kritisch zu hinterfragen und konstruktiv zu diskutieren. Maximilian Hiller für die vielen Gespräche, wissenschaftlich und privat, die ich nicht nur auf beruflicher, sondern auch auf privater Ebene sehr geschätzt habe. Vivien Breitrück für ihre wertvollen Beiträge und ihre Hilfe bei verschiedenen Fragen. Darüber hinaus möchte ich mich auch bei Devin Kwasniok, Nicole Middendorf und Maximilian Kuntze bedanken, auch wenn unsere gemeinsame Zeit recht kurz war. Nicht zuletzt möchte ich mich bei Petra Tecker bedanken. Ihre offene und fröhliche Art hat immer für eine gute Atmosphäre am Lehrstuhl gesorgt. Ich bin sehr dankbar für ihre offene Tür und ihr offenes Ohr.

Ein großes Dankeschön geht auch an alle meine Freunde, die mir geholfen haben, das Leben neben dem Studium und der Promotion interessant und abwechslungsreich zu gestalten. Ohne diesen Ausgleich wäre die Fertigstellung dieser Arbeit für mich sehr viel schwieriger gewesen. Ein besonderer Dank gilt meinen Eltern, Frank und Regine Kölpin, meinen Geschwistern (Karen, Sven, Lars und Inga), meinen Großeltern (Eckard Kölpin, Erika und Reinhard Fiedler) und meiner Schwiegerfamilie (Erwin, Herta und Irina Mattrisch, Mehran Boroumand, Lea Freese, Sabrina Ganse sowie Ingo Mindemann), die mit mir durch alle Höhen und Tiefen gegangen sind und die mich immer bedingungslos unterstützt haben. Ein besonderer Dank geht an meinen Bruder Lars Kölpin-Freese, der mir und meinen Kollegen bei allen technischen Fragen und der anschließenden Umsetzung stets zur Seite stand. Für sein Fachwissen und seine Geduld bin ich ihm sehr dankbar. Einen besonderen Dank möchte ich auch meinem Onkel Wolfram Fiedler aussprechen, der mir bei Fragen mit Rat und Tat zur Seite stand und für dessen Feedback ich sehr dankbar bin.

Der größte Dank gilt meiner lieben Lebensgefährtin Lea Mattrisch, die mich auf meinem Lebensweg mit all ihrer Kraft und Liebe unterstützt und mich durch alle Entscheidungen getragen hat. Für alles, was sie für mich getan und möglich gemacht hat, bin ich ihr unendlich dankbar. Ohne sie wäre das Schreiben dieser Arbeit nicht möglich gewesen.

Ich danke euch allen von ganzem Herzen.

Abstract

This dissertation is a collection of five stand-alone research papers. Each study addresses a scientifically relevant research question in a management context that is answered using a controlled lab experiment. The introduction puts the papers in a more general context.

The first three studies examine the effects of Leading-by-Example on group cooperation. The *first study* uses a meta-analysis and examines the impact of Leading-by-Example in comparison with simultaneous contribution settings. The results show that the establishment of a leader leads to persistently higher contributions, while the aggregate effect remains stable over time and increases in group size. The *second study* investigates a causal relationship between leadership, (endogenous) team size and cooperation. The results show that high contributions of leaders encourage higher contributions of their followers which foster migration into their teams. However, the leader-effect diminishes with group size. Moreover, the results show that incumbents sacrifice economic benefits from potential entrants in order to maintain intra-team cooperation. The *third study* investigates the relationship between leadership, intragroup cooperation, and attrition rates in an online experiment. We observe that successful cooperation delays attrition. Moreover, groups with low initial contribution rates also suffer from rather premature attrition.

The *fourth study* investigates how group formation changes competitive behavior. The study systematically modifies the rules for prize allocation to explain behavioral differences between contests between individuals and between groups. The results show that group formation itself does not lead to a change in overall competitive behavior. The results indicate no evidence that group formation increases both outgroup hostility and unconditional ingroup favoritism. The results imply that group formation increases contest expenditure only in case of perceived fairness within the group.

Last, not least, the *fifth study* focuses on taking responsibility for decisions. The results show that a large proportion of participants disguise their responsibility for the decision behind an outcome of a diced lottery. Moreover, the results show that more selfish givers more often disguise their responsibility. We interpret these results as indicating that givers disguise their responsibility to avoid being perceived as selfish or greedy. However, we find no evidence that the disguise itself leads to a more selfish change in allocation.

Zusammenfassung

Diese Dissertation ist eine Sammlung von fünf eigenständigen Studien. Die Studien befassen sich mit wissenschaftlich relevanten Fragestellungen, die mithilfe von Laborexperimenten beantwortet werden. Die Einleitung stellt die Studien in einen allgemeinen Kontext.

Die ersten drei Studien untersuchen die Auswirkungen von *Leading-by-Example* auf das Kooperationsverhalten. Die erste Studie verwendet eine Meta-Analyse und untersucht *Leading-by-Example* im Vergleich zu simultanen Beitragssettings. Die Ergebnisse zeigen, dass die Etablierung eines *Leaders* zu anhaltend höheren Beiträgen führt, während der Gesamteffekt über die Zeit und mit zunehmender Gruppengröße stabil bleibt. In der zweiten Studie wird ein kausaler Zusammenhang zwischen Führung, (endogener) Teamgröße und Kooperation untersucht. Die Ergebnisse zeigen, dass hohe Beiträge der *Leader* zu höheren Beiträgen ihrer Gefolgsleute führen, was die Migration in ihre Teams fördert. Der Leader-Effekt hingegen nimmt mit der Gruppengröße ab. Darüber hinaus zeigen die Ergebnisse, dass wirtschaftliche Vorteile von potenziellen Neueinsteigern geopfert werden, um die teaminterne Zusammenarbeit aufrechtzuerhalten. Die dritte Studie untersucht in einem Online-Experiment die Beziehung zwischen Führung, Kooperation und Abbruchraten. Es zeigt sich, dass erfolgreiche Kooperation prinzipiell Abbrüche verzögert, Gruppen mit einer niedrigen anfänglichen Kooperationsrate jedoch besonders unter verfrühten Abbrüchen leiden.

In der vierten Studie wird untersucht, wie Gruppen das Wettbewerbsverhalten verändern. Die Studie modifiziert systematisch die Regeln für die Preisvergabe, um Verhaltensunterschiede in Wettbewerben zwischen Individuen und Gruppen zu erklären. Die Ergebnisse zeigen, dass die Gruppenbildung selbst nicht zu einer Veränderung des Wettbewerbsverhaltens führt. Stattdessen deuten die Ergebnisse darauf hin, dass die Wettbewerbsneigung erhöht wird, wenn innerhalb der Gruppe Fairness wahrgenommen wird.

Die fünfte Studie konzentriert sich auf die Übernahme von Verantwortung. Die Ergebnisse zeigen, dass Teilnehmer*innen die Verantwortung für eine Entscheidung eher hinter einem zufälligen Ereignis verstecken. Außerdem zeigen die Ergebnisse, dass egoistischere Geber ihre Verantwortung häufiger versuchen zu verschleiern. Wir interpretieren diese Ergebnisse dahingehend, dass die Geber ihre Verantwortung verstecken, um nicht als egoistisch oder gierig wahrgenommen zu werden. Wir finden jedoch keinen Hinweis darauf, dass die Verschleierung selbst zu einem egoistischeren Verhalten führt.

Table of Contents

List of Tables

Chapter 5

Chapter 6

List of Figures

Chapter 1: Introduction

1.1 Motivation and aim

Decision-making, the core of management, is becoming increasingly complex. Globalization, technological progress, economic and political instability and global environmental problems have created new challenges for organizations, management and leaders (Buckley and Carter, 2004; Drucker, 2012; Ewoh, 2013). Even supposedly simple decisions and actions may lead to unforeseen and uncertain consequences that affect both decision-makers and those who are affected. Bad decisions not only lead to financial loss, but can also cause long lasting damage to the reputation of companies (Black et al., 2000; Schwaiger, 2004). Not at least for these reasons, management needs reliable and researched-based scientific knowledge to make its decisions in the best possible way (Rousseau, 2006).

To answer virtually any question of economic relevance, one must understand how economic actors behave. For example, the effect of good leadership can only be studied with a theory of how employees respond to the leader's behavior. Similarly, cooperation, e.g., in a team, can only be increased if one has an idea of the agent's motivation to behave cooperatively. Behavior in competitions can only be explained if one knows the propensities of the agents. However, nothing is as contradictory as human behavior. People fly to the moon, build computers, smartphones, or skyscrapers, but they fail the simplest tasks, overestimate their own abilities, and sometimes ignore the simplest rules of logic. The traditional economic approach to decision-making assumes rational, risk-neutral agents who have clearly defined preferences and choose their preferred, best possible alternative on this basis. While this approach provides a comprehensible and parsimonious model for studying economic behavior, observable behavior rarely matches these rational, utility-maximizing assumptions (see e.g., Zelmer, 2003 for an analysis in cooperation, Sheremeta, 2013 for an analysis in competition or Engel, 2011 for analysis of giving in dictator games).

During the last decades, the young field of behavioral economics has revolutionized economic research by placing human behavior at the center of research. More specifically, economists have begun to incorporate insights from related disciplines such as psychology and sociology to increase the explanatory power of economic analysis by acknowledging that people make mistakes, care about others, and generally do not behave as rationally as standard theoretical assumptions suggest. Research in this area has generated a vast literature that extends the standard theoretical assumptions to include social preferences such as altruism (Andreoni and Miller, 2002; Levine, 1998), inequality aversion (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999), or reciprocity (Cox et al., 2007; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006). Insights from this research have been applied to many different domains such as finance (e. g. Barberis and Thaler, 2003), labor market (e.g. Gächter and Fehr, 2002), employee performance (e. g. Chadi et al., 2017; Heinz et al., 2020), fundraising (e.g. Mertins and Waldner, 2020), voluntary work (e.g. Jeworrek and Mertins, 2021; Mertins and Walter, 2021), leadership (e.g. Eisenkopf, 2020; Eisenkopf and Walter, 2021; Güth et al., 2007) or management (e. g. Croson and Donohue, 2002).

The aim of this thesis is to provide a deeper understanding of decision-making in economically relevant contexts and to study the underlying behavioral mechanisms. Specifically, this dissertation provides the reader with clear and understandable data on five research topics that to date are new and, to the best of my knowledge, have never been studied in the given context before. The projects of this dissertation primarily address scientifically relevant research topics in a management context, such as leadership and cooperation (Chapters 2 - 4), competitiveness (Chapter 5), and responsibility (Chapter 6). In the following subchapters, the methodology used is explained first. Subsequently, the research questions and the structure of this thesis is addressed in greater detail.

1.2 Methodology

Research is often about how things are related. Thus, research usually tries to identify a cause of an effect. However, especially in the economic context, the dynamic interactions in natural settings complicates the evaluation of any causality. Such empirical data, while rich and numerous, reflect a variety of environmental factors that are difficult or impossible to disentangle. Even with large sample sizes, statistical relationships are often just correlations that do not allow for causal inference because of endogeneity concerns. For example, the work effort of employees in companies cannot be readily observed or measured, and employees face a variety of incentives. This makes it difficult to interpret differences in performance levels. Observed wage differentials may not reflect earnings opportunities, but may be due to company size, worker self-selection, or simply productivity differences. Even when a relationship between wages and work effort is observed, it does not necessarily reflect a relationship between fair pay and work effort but could be due to strategic considerations based on reputation and repeated interactions. Thus, the controlled variation of variables is the basis for empirical scientific knowledge.

One way that such tight control of the decision environment can be achieved is using laboratory experiments. Lab experiments are characterized by the fact that a decision-making situation is created in an artificial environment that makes it possible to test hypotheses established in advance (Friedman and Sunder, 1994). The artificial environment allows researchers to maintain an adequate control of all variables (Charness and Kuhn, 2011; Falk and Heckman, 2009; List, 2007). The experimenter has control over the information provided to the subjects and thus can systematically study the impact of the informativeness of the decision-makers on their decisions. This ability to vary variables allows researchers to ask exactly the question they are interested in and collect data for just that question (Weimann and Brosig-Koch, 2019). Thus, a major advantage is that researchers do not have to rely on economic reality to provide them with exactly the data they need to study a particular question. Rather, they can generate the data themselves. Together with the exogenous and random assignment of participants to treatment and control conditions, this allows for precise testing of theoretical predictions and a clean way to draw inferences about the causal relationship of interest. This leads to another major advantage of lab experiments: The use of specific theoretical models (e.g., cooperative behavior in social dilemmas). The artificial environment allows to replicate the requirements and assumptions of the theoretical models and to compare the predictions of the models (e.g., Nash equilibria) with the collected data. Moreover, competing explanations (e.g., inequality aversion, reciprocity, loss aversion) can be tested so that existing theories can be discarded or extended (e.g. Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Fehr and Schmidt, 1999 or Chapter 2 and Chapter 5).

Critics argue that in lab experiments, internal validity, i.e., the ability to draw confident causal conclusions, is achieved at the expense of external validity, i.e., the generalization of results to the world outside the laboratory (Levitt and List, 2007).[1] More precisely, the artificial and conceptual decision-making situations whose assumptions are highly abstracted from the realities of everyday economic life, hardly allow a generalization of the theories or findings to the real world. Some researchers argue that external validity does not play a major role, at least in classical experiments, because experiments are designed to test theories and therefore the question of whether the experimental results also apply in reality is irrelevant (Schram, 2005). However, even if one shares this view, the question of the meaning of theories tested in the lab

[1] Note that only the most common criticisms are addressed here. For a detailed scientific discussion of the advantages and disadvantages of lab experiments, please refer to Charness (2010); Falk and Heckman (2009); Levitt and List (2007).

remains. Without a reference to the real world, research loses its claim to be empirical. Successful research requires that the experiment actually tests the theory it is designed to test and that it leads to observations that contribute to a better understanding of real-world phenomena. Experimental research thus depends on being able to derive general relationships from its results. However, this does not apply to the single observation in a single experiment (Weimann and Brosig-Koch, 2019). The existence of general relationships can only be assumed if the observations are reproducible and if they have been shown to be robust to changes in experimental settings. Thus, generalizable conclusions are possible with a large number of independent observations showing the same or at least similar relationships. At this point, another feature of lab experiments comes into play: replicability (Charness, 2010). The artificial environment and adherence to a specific procedure, e. g. the same underlying game structures, allows researchers to replicate experiments as often as they like. In well-designed experiments, scientists will inevitably replicate important findings that remain stable under various environmental conditions. The comparability of studies also allows for e. g. meta-analyses to be conducted to test proposed effects in different scenarios and environments. Although economists have rarely relied on meta-analyses, the increasing popularity of experimental methods with standardized games has led to some meta-analyses across different experiments (e.g. Abeler et al., 2019; Croson and Marks, 2000; Engel, 2011; Johnson and Mislin, 2011; Larney et al., 2019; Zelmer, 2003 as well as Chapter 2).

Another common objection to lab experiments is that the subject pools primarily consisting of students, who are not representative sample of the population, and that the samples sizes are too small (Levitt and List, 2007). Of course, it is better to have a broader range of participants than students, who are the most convenient source of participants for conducting lab experiments. However, many of the economic models derive predictions that are independent of assumptions concerning participant pools (Falk and Heckman, 2009). Furthermore, it should be added that a large number of lab experiments have already been conducted with real-world participants (e.g., with soldiers (Fehr et al., 1998; Ruffle and Tobol, 2017), with business managers (Fehr and List, 2004), or with production workers (Barr and Serneels, 2009)). In many cases, no differences could be found between the results of the student and 'real-world' participants. Although it cannot be ruled out that there are differences between students and non-students (Anderson et al., 2013; Bortolotti et al., 2015; Falk et al., 2013), in many settings the advantages of a student sample, such as lower opportunity costs or better availability, outweigh the disadvantages (Friedman and Sunder, 1994; Weimann and Brosig-Koch, 2019).

Last but not least, another feature, but also a common objection, are the individual, incentive-based payoffs (Abbink 2006). More precisely, participants receive a payoff that depends on their decision. In economic experiments, it is therefore common to reward subjects in proportion to their gains in the experiment. This ensures that subjects have an appropriate incentive to maximize their payoff and make careful decisions (Clot et al., 2018; Gächter and Renner, 2010). Critics counter that the payoff incentives may be too small and insignificant to actually provide the incentives (Levitt and List, 2007). However, studies show that the amount of payoffs does not significantly change behavior (Carpenter et al., 2005; Fehr et al., 2014). A meta-analysis by Camerer and Hogarth (1999) confirms this finding. More precisely, they find that monetary incentives have an effect if it is possible for subjects to earn more money through higher effort. However, the marginal monetary return does not matter. All in all, this shows that incentives are very important. However, the level of payoffs in the experiment does not seem to be a fundamental problem (Weimann and Brosig-Koch, 2019).

With respect to the research questions of this dissertation (see Chapter 1.3), lab experiments represent a well suited workhorse because they circumvent the problem that heterogeneity in everyday life usually arises endogenously and is often difficult to measure. For this reason, all projects in this dissertation are based on lab experiments that were either self-conducted or whose data are based on lab experiments. In the next subsection, the research questions and the structure of this thesis are described in more detail.

1.3 Structure

In this thesis, five experiments were conducted in controlled laboratory settings. The first three research projects examine leadership in different contexts. More precisely, the first project studies the impact of leadership on cooperation compared to leaderless contexts using a meta-analysis (Chapter 2). The second study examines the effects of group size and endogenous growth on leadership (Chapter 3), while the third experiment investigates the effects of leadership on attrition in an online experiment (Chapter 4). The fourth research project examines the effects of group affiliation on competitive behavior (Chapter 5). Finally, the fifth experiment deals with taking responsibility for unpleasant decisions (Chapter 6). This subsection provides an overview of the research topics and briefly summarizes each research project.

Chapters 2 through 4 consider leadership in the context of a social dilemma. The voluntary contribution mechanism (Isaac and Walker, 1988a, 1988b) represents one of the most popular workhorses to experimentally analyze cooperative behavior in a social dilemma. It models the dilemma of the opposition between selfish preferences and social efficiency. A social dilemma can be characterized by three properties: First, each member of a given group can choose between a socially defecting or cooperative choice. Second, the socially defecting choice yields to a higher individual payoff, regardless of the decisions of the other members. Third, individuals, however, are better off if everybody choose the socially cooperative choice (Dawes, 1980). Social dilemmas are typical for many issues in everyday life. For example, they are prevalent in political contexts (Rothstein, 2000), in insurance contracts (Janssens and Kramer, 2016), in environmentally friendly behavior (Bucciol et al., 2019; Delmas and Keller, 2005), or in teamwork at the workplace (Blasco et al., 2019). Leadership in this context means that a first mover commits herself unilaterally to a decision. The other group members learn about this decision and decide on their own (Leading-by-Example). While first movers can inspire others to choose the cooperative choice, people may be reluctant to take the leading role because they fear others will not follow their example (Cappelen et al., 2016).

Cooperation is a key to success of human societies (Falk et al., 2003). Cooperation involves choices that require bearing individual costs in pursuit of collective benefits. Standard economic theory assumes that successful cooperation between individuals is only achieved when binding agreements (usually explicit contracts) are used. Everyone involved in a decision must be part of these agreements. In other words, coordination leads to successful cooperation only when people will commit to specific actions. However, even without such agreements with others, people tend to cooperate and voluntarily contribute to public goods, but it is often far from a socially efficient outcome (see e.g. Chaudhuri, 2011 for an overview). There is substantial evidence that (the threat of) punishment (Bochet et al., 2006; Eisenkopf and Walter, 2021; Fehr and Gächter, 2000), rewards (Gürerk et al., 2018; Sutter and Rivas, 2014) or communication (Bochet et al., 2006; Brosig and Weimann, 2003) can foster cooperation. These instruments are seen as central for some leadership styles (e. g. transactional leadership) (Podsakoff et al., 2006). However, many great leaders believe that the greatest instrument of leadership is to lead by example. For example, Albert Schweitzer, German physician and philosopher, stated that ‘the three most important ways to lead people are:… by example… by example… by example’.[2] John Wooden, head coach of the UCLA basketball team, explains that ‘a leader’s most powerful ally is his or her own example. Leaders don’t just talk about doing something; they do it’

[2] Note: This quote is attributed to Albert Schweitzer. However, no primary source could be found.

(Wooden and Carty, 2009). Barack Obama, 44th president of the United States stated 'We must lead the world, by deed and by example' (Obama, 2007). Economic experiments have provided evidence for the popular notion that Leading-by-Example is a successful instrument to coordinate groups (Dannenberg, 2015; Eichenseer, 2021; Eisenkopf, 2020; Güth et al., 2007; McCannon, 2018; Moxnes and Van der Heijden, 2003). However, this positive effect cannot always be confirmed (Gächter and Renner, 2018; Gürerk et al., 2018; Haigner and Wakolbinger, 2010; Sahin et al., 2015). To better understand the impact of such leadership style, the second chapter provides a parsimonious model of leadership in social dilemma situations and tests it with a meta-analysis of experimental studies. Meta-analyses are popular tools in leadership studies outside from economics (Bedi et al., 2016; Jong et al., 2016; Simons et al., 2015; Zhang et al., 2019). Economists rarely use meta-analyses but the increasing popularity of experimental methods with standardized games has provided some studies on public goods games (Croson and Marks, 2000; Zelmer, 2003), trust games (Johnson and Mislin, 2011), ultimatum games (Larney et al., 2019), dictator games (Engel, 2011), and the experimental paradigm of Fischbacher and Föllmi-Heusi (2013) on truth-telling (Abeler et al., 2019). The meta-analysis from Chapter 2 extends the existing literature, through a focus on Leading-by-Example. The meta-analysis relied on studies with treatments that allow for sequential contributions to a public good (as in Güth et al., 2007). The group members observe the contribution of a leader before contributing themselves. The results are compared with simultaneous contribution treatments from the same studies. The main finding is that the establishment of a leader results in persistently higher and more coordinated contributions. The aggregate effect remains stable over time and increases in group size even though leaders and followers have more divergent contribution patterns in larger groups.

Chapter 3 and Chapter 4 investigate the effectiveness of Leading-by-Example as well as the impact of a leader in different contexts. More precisely, Chapter 3 studies a causal relationship between leadership, team size and cooperation. The chapter studies how the growth of teams affects leadership effectiveness and intra-group cooperation. The study approaches the question in two ways: On the one hand, the group size effect, and its impact on Leading-by-Example is considered. While studies with simultaneous contributions structures suggest a positive effect of team size on contributions (Isaac et al., 1994; Isaac and Walker, 1988b; Zelmer, 2003), insights about the impact on Leading-by-Example are rather limited. Thus, the focus on exogenous manipulation represents an innovation in the Leading-by-Example literature. On the other hand, this study extends the literature of endogenous group formation. Previous studies show that the rules of endogenous group formation are crucial for the maintenance of cooperation (Ahn et al., 2008; Chakravarty and Fonseca, 2017; Cinyabuguma et al., 2005; Maier-Rigaud et al., 2010). The focus on effective leadership as a determinant of endogenous growth and the examination of whether growth undermines leadership effectiveness represents an innovation in the existing literature. The results show a virtuous circle between leadership, team size and contributions. High contributions of leaders encourage higher per capita contributions of their followers which foster migration into their teams. In turn, larger teams experience even more courageous leadership and higher contributions, but the coordination effect diminishes.

Chapter 4 examines the relationship between leadership, intra-group cooperation, and attrition rates in an online experiment. More precisely, Chapter 4 investigates whether (successful) leadership in a social dilemma has an impact on task motivation and thus on (premature) attrition in our experiment. While endogenous attrition rates typically jeopardize the internal validity of the experiment (Zhou and Fishbach, 2016), we turn this bug into feature and use the attrition rate as a (noisy) signal of task motivation. The experimental design as well as the focus on effective leadership on task motivation represents a novel innovation in the existing literature. Our results show that, in general, successful cooperation decreases attrition while leaving group members foster both uncooperative behavior and attrition. Moreover, we do not find evidence that successful cooperation decreases premature attrition for leaderless groups. Specifically, we find that there is little difference in attrition rates between cooperative and uncooperative groups in the first part. In both cooperative and uncooperative groups, participants forgo the economic benefit of an additional task and prefer to leave the experiment, even if positive payoffs are still possible. However, for groups with a leader, such behavior occurs only for uncooperative groups. For cooperative groups, on the other hand, we find a positive impact of cooperation on premature attrition, but we cannot attribute this effect from the leader.

Chapter 5 tests a behavioral explanation on how group formation changes competitive behavior. The experimental treatments bridge the gap between contests of individuals and contests between groups by modifying prize sharing rules step by step. In a typical experimental contest between individuals, only the winner gets the prize, while all other contestants leave empty handed. In standard intergroup contests, the prize money is distributed among the members of the winning group. In many cases - and in most experimental intergroup contests (Abbink et al., 2010; Eisenkopf, 2014, 2018, 2020; Leibbrandt and Sääksvuori, 2012; Sheremeta, 2013) - all members get equal shares irrespective of their contributions, while meritocratic considerations and incentive concerns would apportion higher shares for people who provided more resources to victory. An equal distribution of prizes among group members implies that any contest expenditure generates a positive externality for the other group members. Hence, the formation of groups with egalitarian prize sharing should lead to less expenditure among risk neutral contestants without social preferences. However, some studies suggest that expenditures in intergroup contests exceed expenditures in contests between individuals (Abbink et al., 2010; Bornstein et al., 2008; Charness et al., 2007; Chen and Li, 2009). It is often argued that the formation of groups itself may generate a parochial altruism (Abbink et al., 2012; Bernhard et al., 2006b; Choi and Bowles, 2007) that induces contestants to take the benefits of their fellow group members into account. However, there is little evidence on how these behavioral mechanisms are elicited. The results show that group formation itself does not lead to a change in overall competitive behavior. We find no evidence that group formation increases both outgroup hostility and unconditional ingroup favoritism. The results imply that group formation increases contest expenditure only in case of perceived fairness within the group.

Last but not least, Chapter 6 focus on taking responsibility for unfair decisions. Giving in experimental studies, such as the dictator game, seems vastly reduced when the responsibility of the giver is diminished, e.g. if receivers' payoff is not only determined by the giver but also by chance (Andreoni and Bernheim, 2009; Dana et al., 2007) or the anonymity of the decision-maker is increased (Alpizar et al., 2008; Franzen and Pointner, 2012). The results of these studies suggest the underlying motivation for fair behavior could be self-interest, combined with the desire to maintain the illusion of not being selfish. Thus, fair behavior disappears when individuals have an excuse for not having to give. This is the focus of the study presented in Chapter 6. More specifically, Chapter 6 studies whether participants in a lab experiment are willing to deceive another participant in order to disguise their own responsibility for an unfair decision. The results show that a large proportion of participants disguise their responsibility for the decision behind a die roll. Moreover, the results show that more selfish givers more often disguise their responsibility. We interpret these results as indicating that givers disguise their responsibility to avoid being perceived as selfish or greedy. However, we find no evidence that the disguise itself leads to a more selfish change in allocation.

Each of the following chapters follow the same structure: First, each chapter is placed in a general context and the research question is inferred. Then, the underlying experimental design is explained and theoretical predictions are derived. In the main part of each chapter, the theoretical predictions are empirically tested. Finally, the results are discussed and conclusions are drawn. Chapter 2 – 5 were developed in cooperation with Gerald Eisenkopf. Chapter 6 was developed in cooperation with Vanessa Mertins and Robert Gillenkirch.

Chapter 2: Leading-by-Example: A Meta-Analysis

2.1 Introduction

The dynamic interaction between leaders and followers in groups complicates the evaluation of any leadership process. It is difficult to identify causal relationships between the characteristics of leaders and followers, their decisions and group performance. Even with large sample sizes, statistical relationships are often just correlations that do not allow for causal inference because of endogeneity concerns. Hence, leadership research has focused in recent years on experiments as a complimentary research method. A well-designed experimental study eliminates many statistical concerns in field studies (Antonakis et al., 2010) and circumvents related problems of selectivity. Ideally, such experiments take place in natural environments, but such circumstances often reflect idiosyncratic characteristics of the respective context. A clean identification strategy requires carefully controlled and replicable experiments with a design that tests the hypotheses of a specific leadership theory (Falk and Heckman, 2009).

We exploit the benefits of this scientific approach in this paper. We focus on experimental studies that use exemplary leadership in a social dilemma (as described in Güth et al., 2007). More specifically, a randomly selected leader visibly commits herself to a specific contribution to a public good before the fellow group members do so. Our paper studies four questions about the behavioral impact of such leadership. For each question we propose a hypothesis that derives from a parsimonious theory which combines standard economic reasoning with the simple reciprocity model proposed by Fehr and Schmidt (1999). We test this theory with a meta–analysis of experimental studies that allow for clean and replicable identification.[3]

In most studies, the decision sequence of the leadership game increases aggregate contributions and welfare relative to groups without a leader (Dannenberg, 2015; Eichenseer, 2021; Güth et al., 2007; McCannon, 2018; Moxnes and Van der Heijden, 2003) but this positive effect cannot always be confirmed (Gächter and Renner, 2018; Gürerk et al., 2018; Haigner and Wakolbinger, 2010; Sahin et al., 2015). Therefore, our first research question focuses on the aggregate impact of such leadership on contributions across several studies.

Question 1: *Does Leading-by-Example increase contributions?*

[3] Note that more recent models of reciprocity such as Dufwenberg and Kirchsteiger (2004), Cox et al. (2007); Falk and Fischbacher (2006) seem to provide a better explanation for reciprocal behavior in a variety of games. Our own calculations suggest that these more complex models do not provide qualitatively different results for the leadership game we consider. Hence we apply Occam's razor and stick to the simpler model of Fehr and Schmidt (1999).

Aggregate contributions are the most popular indicator of leadership effectiveness, but they do not constitute the only one. Our second research question focuses on the impact of such leadership on coordination. Fischbacher et al. (2001) show that many people are conditional cooperators, who prefer to match others' contributions. Barron and Nurminen (2020) demonstrate that even simple nudges lead to a remarkable increase in cooperation rates, because nudges provide a focal point, potentially helping conditional cooperators to coordinate. Thus, when followers are conditional cooperators, leaders' high efforts may provide a focal point and induce the followers to exert high efforts. Gächter et al. (2012), Frackenpohl et al. (2016) and Cartwright and Patel (2010) elicited followers contributions using the strategy method (Selten, 1967). Their results show a significant correlation between the contributions of followers and leaders. Hence, a cautious leader also has an impact. Their low contributions can decrease the expenditure of followers. Note that a limited degree of reciprocity among followers implies costs and risks for the leader. Some studies report that leaders end up worse than followers do (Cappelen et al., 2016; Eisenkopf, 2020; Gächter and Renner, 2018) and they may receive even lower payoffs than in a group without leadership. Here, the random allocation into experimental roles is particularly helpful for evaluation because voluntary leadership is a highly selective process that attracts only persons with specific characteristics (Alan et al., 2019; Arbak and Villeval, 2013).

Question 2: *Does Leading-by-Example induces a stronger alignment of group members' contributions?*

Third, we focus on the stability of any leadership effect over time. If leaders are also conditional cooperators they might be unwilling to uphold high levels of contribution in case of lower. Gächter and Renner (2018) as well as Teyssier (2012) show that leaders behave almost perfectly like conditional cooperators and match their contribution with the amount they believe the followers will contribute. Consequently, the positive effect of the Leading-by-Example would have to disappear if followers permanently undercut the contribution of the leader.

Question 3: *Do higher contributions with Leading-by-Example persist over time?*

Last, not least, we have an interest in the role of group size on the impact of leadership. Several public good experiments show a positive correlation between group size and contributions in related experiments without a leader (Goeree et al., 2002; Isaac et al., 1994; Isaac and Walker, 1988b). Even if this effect is not robust for all comparisons (Carpenter, 2007; Nosenzo et al., 2015), a meta study with 27 experiments confirm the positive correlation (Zelmer, 2003). However, only few studies deliberately investigate the role of group size in the context of leadership. Some studies suggest that the effect of Leading-by-Example is also present in larger groups (Figuieres et al., 2012), but the coordination effect of leaders seems to diminish with group size (Komai and Grossman, 2009). Therefore, our last research question focuses on the impact of group size on Leading-by-Example.

Question 4*: Does the impact of Leading-by-Example change with the size of the group?*

We obtained data from 14 studies with 369 groups as independent observations. Our results show that Leading-by-Example significantly increases contributions in comparison with leaderless settings. Followers reply in kind to the leader, but only to a certain extent. Consequently, leaders contribute significantly more than followers. Therefore, we can confirm the 'leader's curse' (Gächter and Renner, 2018): They earn less than their fellow group members. Nevertheless, leaders do not reduce their contributions more than those of their followers. The positive effect of Leading-by-Example is maintained over a longer period of time. Moreover, we find that contributions increase in group sizes. However, leaders in larger groups elicit less coherent responses from their followers.

Meta-analyses are popular tools in leadership studies from outside economics. Jong et al. (2016) observe in their study (which combines results from 112 studies) that trust between leaders and followers correlates significantly with team performance. Simons et al. (2015) find that the behavioral integrity of a leader has rather high correlations with trust, in-role task performance and organizational citizenship behavior. Their analysis relied on 35 samples. Similarly, Bedi et al. (2016) as well as Zhang et al. (2019) provide meta-analyses on the relationship between ethical leadership and follower outcomes. However, such aggregation of evidence still provides just correlational evidence and does eliminate the methodological limitations of field studies that do not exploit random variations in their identification strategies (Antonakis et al., 2010). Economists rarely use meta-analyses but the increasing popularity of experimental methods with standardized games has provided some studies on public goods games (Croson and Marks, 2000; Zelmer, 2003), trust games (Johnson and Mislin, 2011), ultimatum games (Larney et al., 2019), dictator games (Engel, 2011) and the experimental paradigm of Fischbacher and Föllmi-Heusi (2013) on truth-telling (Abeler et al., 2019). More often, economists review the relevant journals in dedicated outlets like the Journal of Economic Literature or the Journal of Economic Surveys. Regarding the topic of our paper, Eichenseer (2021) provides a thorough and thoughtful discussion of the relevant literature and interesting variations of the leadership game. That paper also includes an aggregate quantitative synthesis of the reported results, while our paper allows for an in-depth analysis of the interaction between individual leaders and followers within the groups.

The remainder of the paper is organized as follows. Section 2.2 explains the underlying game structure, our leadership model as well as the resulting theoretical predictions. Section 2.3 explains the methodology. Section 2.4 presents our main results, while section 2.5 concludes.

2.2 The underlying game and theoretical predictions

Our study focuses on behavior in the following voluntary contribution game (VCM) that has been by introduced by Isaac and Walker (1988b) and adapted to leadership by Güth et al. (2007).

Let $I = (1, \ldots, N\}$ denote a group of $N \geq 3$ individuals who interact for $t = 1, \ldots, T$ rounds. In each round t, individual $i \in I$ gets an endowment $e > 0$, which can be either privately consumed or contributed to a group activity. For our theoretical analysis, we set $T = 1$, standardize the endowment $e = 1$ and consider a binary decision regarding the contribution: $c_i \in \{0,1\}$. This simplification allows us to focus on the conditions that induce group members to make nonnegative contributions and how Leading-by-Example alters these conditions. The monetary payoff of individual i takes the following form:

$$\pi_i = 1 - c_i + q \sum_{j=1}^{N} c_j \tag{E2.1}$$

We have an interest in all studies in which $1 > q > \frac{1}{N}$ holds. Because of $1 > q$ the dominant strategy for each rational and selfish player is to contribute nothing. However, because of $q > \frac{1}{N}$, full contributions would generate the highest aggregate payoff of all group members.

We consider two variants of this game: the standard simultaneous VCM and the VCM with leadership. In the simultaneous VCM, all N group members make their contribution decisions privately and simultaneously. The VCM with leadership has two decision stages. First, the leader, L, chooses his contribution c_L, which is observed by the followers. Then, the followers F $(\neq L)$ decide simultaneously about their own c_F. Applying backward induction and assuming commonly known monetary payoff maximization, the theoretical prediction for the VCM with leadership do not differ from those for the standard-VCM: Because of $q < 1$, the followers' dominant strategy in stage 2 is to contribute zero. A rational leader will anticipate this and free-ride as well in stage 1.

However, we want to inquire whether reciprocal preferences induce more cooperation in the group and how Leading-by-Example fosters this cooperation. We assume that players suffer psychological losses from both advantageous and disadvantageous inequality (Fehr and Schmidt, 1999).[4] The parameter β_i (with $1 > \beta_i \geq 0$) measures the utility loss from advantageous inequality, while α_i with $\alpha_i \geq \beta_i$ indicates the loss from disadvantageous inequality. For simplicity, we assume that people have common knowledge about a homogeneous $\alpha_i = \alpha > (1-q)$ in the population. Moreover, regarding advantageous inequality aversion we consider only two types of persons ($\beta_i \in \left\{\underline{\beta}, \overline{\beta}\right\}$, with $\underline{\beta} = 0$ and $1 - q < \overline{\beta} \leq \alpha$). While the share of $\underline{\beta}$ in the entire population is common knowledge, individual realizations of β_i constitute private knowledge. Groups consist of randomly drawn samples from that population. Let x denote the expected share of people with $\beta_i = \underline{\beta}$ within a group (and $1 - x$ for $\beta_i = \overline{\beta}$). In case of simultaneous contributions, the utility function of a group member can be denoted as follows:

$$\begin{aligned} U_i = 1 - c_i + c_i q_i + q_j \sum_{j=1}^{N-1} c_j - \frac{\alpha}{N-1}\left(\sum_{j=1 \neq i}^{N-1} max\{c_i - c_j, 0\}\right) \\ - \frac{\beta_i}{N-1}\left(\sum_{j=1 \neq i}^{N-1} max\{c_j - c_i, 0\}\right) \end{aligned} \quad \text{(E2.2)}$$

FS have identified in proposition 4 (with some differences in notation) two conditions for potential asymmetric equilibria in which some group members increase their utility with a contribution. First, for any such member $\beta_i > (1-q)$ must hold. Hence, only people with $\beta_i = \overline{\beta}$ will contribute. Second, the benefits of a contribution must outweigh the costs of free-riding:

$$q + q(1-x)(N-1) - \alpha x > 1 + q(1-x)(N-1) - \overline{\beta}\,(1-x)$$

[4] Note also that FS provide a more detailed analysis regarding more differentiated contribution possibilities and social preferences.

This inequality implies the following condition:

$$\hat{x}_{Sim} = \frac{q - 1 + \overline{\beta}}{\alpha + \overline{\beta}} > x \tag{E2.3}$$

The term $\hat{x}_{Sim}$ denotes the maximum share of expected free riders at which contributions become profitable for conditional cooperators. We now study the case of the sequential VCM. Applying backward induction, we first look at the utility function of a follower F_i who has observed the decision of leader L:

$$\begin{aligned} U_{F_i} = 1 - c_{F_i} + qc_{F_i} + qc_L + q \sum_{j=1\neq i}^{N-2} c_j \\ - \frac{\alpha}{N-1}\left(\max\{c_{F_i} - c_L, 0\} + \sum_{j=1\neq i}^{N-2} max\{(c_{Fi} - c_{Fj}, 0\} \right) \\ - \frac{\beta_{Fi}}{N-1}\left(\max\{c_L - c_{F_i}, 0\} + \sum_{j=1\neq i}^{N-2} max\{c_{Fj} - c_{F_i}, 0\} \right) \end{aligned} \tag{E2.4}$$

In case of $c_L = 0$, the leader has already revealed her free riding. Again, the benefits of a follower's contribution must outweigh the costs of free-riding for any follower with $\beta_i = \overline{\beta}$:

$$q + q(1-x)(N-2) - \alpha\left(\frac{1 + x(N-2)}{N-1}\right) > 1 + q(1-x)(N-2) - \overline{\beta}\frac{(1-x)(N-2)}{N-1}$$

which leads to the following maximum share of free riders in the group for nonnegative contributions:

$$(\hat{x}_{Seq}|c_L = 0) = \frac{\left(q - 1 - \frac{\alpha}{N-1} + \frac{\overline{\beta}(N-2)}{N-1}\right)}{(\alpha + \overline{\beta})}\left(\frac{N-1}{N-2}\right) < \frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}} = \hat{x}_{Sim} \tag{E2.5}$$

If the leader has contributed her endowment instead ($c_L = 1$), the number of potential free riders within the group has decreased by one person. Hence, a conditionally contributive follower tolerates a higher share of free riders among the other followers than in the case of simultaneous contributions:

$$\left(\hat{x}_{Seq} \middle| c_L = 1\right) = \left(\frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}}\right)\left(\frac{N-1}{N-2}\right) > \frac{q + \overline{\beta} - 1}{\alpha + \overline{\beta}} = \hat{x}_{Sim} \tag{E2.6}$$

Now we study the benefit of a contribution for the leader. She will contribute if

$$(U_L | c_L = 1) > (U_L | c_L = 0)$$

holds which depends on the response of the followers towards the choice of the leader. As shown above, this response depends on the realization of x. We have three cases to distinguish:

- $x < \left(\hat{x}_{Seq} \middle| c_L = 0\right)$. In this case, there are enough conditional cooperators in the population. These people will always contribute irrespective of the choice of the leader. Any leader with $\beta_L = \overline{\beta}$ will contribute.
- $\left(\hat{x}_{Seq} \middle| c_L = 1\right) < x$. The followers will never cooperate irrespective of the choice of the leader.
- $\left(\hat{x}_{Seq} \middle| c_L = 0\right) \leq x < \left(\hat{x}_{Seq} \middle| c_L = 1\right)$. In this case, conditional cooperators follow the choice of the leader. This case implies $(U_L | c_L = 1) = q + q(N-1)(1-x) - x\alpha$ while $(U_L | c_L = 0) = 1$ holds. Hence, we obtain

$$x < \frac{qN - 1}{\alpha + q(N-1)} = \hat{x}_{Seq_L}$$

This inequality has an important implication because it provides an explanation why leaders accept lower average payoffs than followers (the leader's curse). The contribution decision in this critical case does not depend on β_L, the 'compassionate' part of a leader's inequality aversion. It rests on the expectation about how many followers the leader can induce to reciprocate, such that the expected net returns from the aggregated investments compensate the expected loss from disadvantageous inequality. This insight implies that leaders should not stop contributions if it pays off even if some group members are free-riders. Comparing $\hat{x}_{Seq_L}$ with $\hat{x}_{Sim} = \left(\frac{q+\overline{\beta}-1}{\alpha+\overline{\beta}}\right)$, the threshold in case of simultaneous contributions, it becomes obvious that a person is more likely to contribute as a leader if $\bar{\beta}$ does not exceed q or otherwise, if the group size is sufficiently large $\left(N > \frac{\overline{\beta}}{q} + 1\right)$.

Hence, our analysis has identified two sources of increased contributions via leadership. Conditional cooperators are more likely to contribute if the leader has contributed herself while assuming the leadership role itself tends to induce contributions in particular among otherwise uncooperative group members. Overall, our theoretical analysis suggests the following expectations regarding the research questions:

Prediction 2. 1*: Leading-by-Example increase contributions.*

Prediction 2.2: *Leading-by-Example induces a stronger alignment of group members' contributions.*

Prediction 2.3: *Higher contributions with Leading-by-Example persist over time.*

Group size has an ambiguous effect in this context. If x is in the appropriate parameter range, an increase in group size makes a leader's contribution more rewarding. At the same time, an increasing group size decreases that parameter range, as the leader's decision becomes less relevant for the fellow group members. However, this result relied on the assumption about common knowledge regarding the share of people with $\beta_i = \underline{\beta}$. This assumption is rather bold in the context of anonymous interaction between participants who do not know each other. Hence, followers will have a prior about the share. They can use a leader's decision to update this prior. Therefore, a contribution by the leader can increase the expectations of the followers regarding the share of participants with $\bar{\beta}$. This revision of expectations does not depend on the group size.

Prediction 2.4: *For a given level of marginal per capita return (the variable q), larger groups will see higher investments of leaders than smaller groups.*

2.3 Method

2.3.1 Search of Studies and Study criteria

We searched for relevant studies in early 2020. More specifically, we first looked for experimental studies that investigate leadership with a voluntary contribution mechanism. Leadership in this context means that one of the group members acts first and the others observe the behavior before they act themselves. We applied a three-step search procedure. First, we searched relevant databases for published studies, including e.g., Google Scholar, EconLit or IDEAS, using terms like *Leadership*, *Leading-by-Example, Voluntary Contribution Mechanism, Sequential Contribution, Public Goods*. This first search yielded 251 potential results. We removed all theoretical papers and those that did not use a voluntary contribution mechanism. Second, we checked the references in the remaining papers and looked at their citations in Google Scholar. These steps resulted in potential 33 studies. Last, not least, we sent a request to the e-mail list of the *Economic Science Association (ESA),* asking for additional publications as well as working papers and other unpublished research. The request identified 15 additional studies (including our own relevant working papers). Thus, our search procedure yielded 48 studies.

Subsequently, we further narrowed down our criteria. More specifically we looked for studies that met the following criteria: First, to ensure comparable procedures, we restricted the studies to experimental studies that used the voluntary contribution mechanism as in Güth et al. (2007). Second, each relevant study had to have at least one treatment with a randomly allocated leader in order to avoid self-selection among leaders. Third, all participants remained in their groups throughout the entire experiment. Applying these criteria resulted in a total of 20 potential studies for which we requested the data. We obtained the results in 15 cases. However, for one study we only got data on the aggregated level, so we were unable to recover the individual decisions. In total, we received full data from 14 studies, which we include in our meta-study. Table 2-1 informs about the included studies. We hope that future meta-analyses will be able to replicate our analyses with more samples from more diverse populations.

Table 2-1: Included Studies

Study	# of Rounds	Group Size	MPCR	Endow.	Simultaneous Treatment	Location
Centorrino & Concina (2013)	10	4	0.5	30	---	University of Venice, Italia
Dannenberg (2015)	10	4	0.4	25	✓	University of Magdeburg, Germany
Drouvelis & Nosenzo (2013)	10	3	0.5	20	---	University of Nottingham, England
Eisenkopf (2020)	20	3/6	0.5	100	✓	University of Konstanz, Germany
Eisenkopf & Kölpin (2022)	20	3	0.5	100	✓	University of Hamburg, Germany
Eisenkopf & Walter (2021)	20	3	0.5	100	✓	University of Hamburg, Germany
Frackenpohl et al. (2016)	10	4	0.4	20	---	University of Bonn, Germany
Gächter & Renner (2018)	10	4	0.4	20	✓	University of Erfurt, Germany
Gürerk et al. (2018)	20	4	0.4	20	✓	Aachen University, Germany
Güth et al. (2007)	16	4	0.4	25	✓	Max-Planck-Institute, Germany
Moxness & van der Heijden (2003)	10	3	0.4	20	✓	Norwegian School of Economics, Norway
Rivas & Sutter (2011)	16	4	0.4	25	✓	Max-Planck-Institute, Germany
Sahin et al. (2015)	20	6	0.2	9	✓	Virginia Tech & University of Texas, US
Yu & Kocher (2020)	10	4	0.8/0.4	20	✓	University of Munich, Germany

Note: Eisenkopf and Kölpin (2022) report four Leading-by-Example treatments with different group sizes (3 and 6). Moxnes and Van der Heijden (2003) used a public bad environment. We consider only the contribution into the non-damaging good. Yu and Kocher (2020) implement a public goods experiment with heterogeneous marginal per capita returns.

2.3.2 Coding and Data preparation

For each study, we first transcribed the individual observations into a general form. This data form contained all variables necessary to answer our research questions. Our prime dependent variable was the individual contribution towards to public good. However, since all studies has different maximum stakes, we calculated the individual contribution as a percentage of the respective endowment. This procedure facilitated the comparison of the results of the different studies. Moreover, we included the respective round, the role in the experiment (leader or follower), the subject number as well as the respective group. Note, that we later assigned a unique identification number to each study, group, and subject, to ensure identification. Besides these subject dependent variables, we added study specific characteristics that have been hypothesized in the literature to affect contributions towards a public goods. Such characteristics include the marginal per capita return (mpcr), group size, number of rounds, endowment and the exchange rate of the tokens/points into monetary amounts (Ledyard, 1995; Zelmer, 2003). To ensure comparability, we have converted the exchange rate into euros and adjusted it for inflation (as of summer 2021). Last, not least, we added treatment variables that indicate whether a group played in a sequential or simultaneous contribution structure. After applying the described procedure to all studies, all data sets were transferred into one complete data set.

2.4 Results

Before we focus on our research questions in detail, we have a brief look at the descriptive statistics. Table 2-2 presents the average relative contributions for the simultaneous and Leading-by-Example structures, as well as the number of included studies, groups and subjects. Appendix A.I report additional descriptive statistics for each included study separately. For the sake of clarity, we subdivide the results part in different subsection. Each subsection provides results for at least one of our predictions presented in section 2.2. Note that we report results from additional robustness checks in the Appendix. More specifically, we looked whether individual studies had an outsized effect on the aggregate outcomes in any of the regression models. We replicated the estimations and eliminated each individual study in a specific subsample for each robustness check. In a few cases, elimination leads to minor changes. The robustness checks are shown in the Appendix A.II. We refer to the checks in the corresponding sections.

Table 2-2: Descriptive Statistics

	Simultaneous	Leading-by-Example
Average contributions (in percentages of the endowment)		
All	0.379 (0.245)	0.507 (0.265)
Leader	---	0.596 (0.276)
Follower	---	0.478 (0.276)
Studies	11	14
Groups	179	248
Subjects	686	970

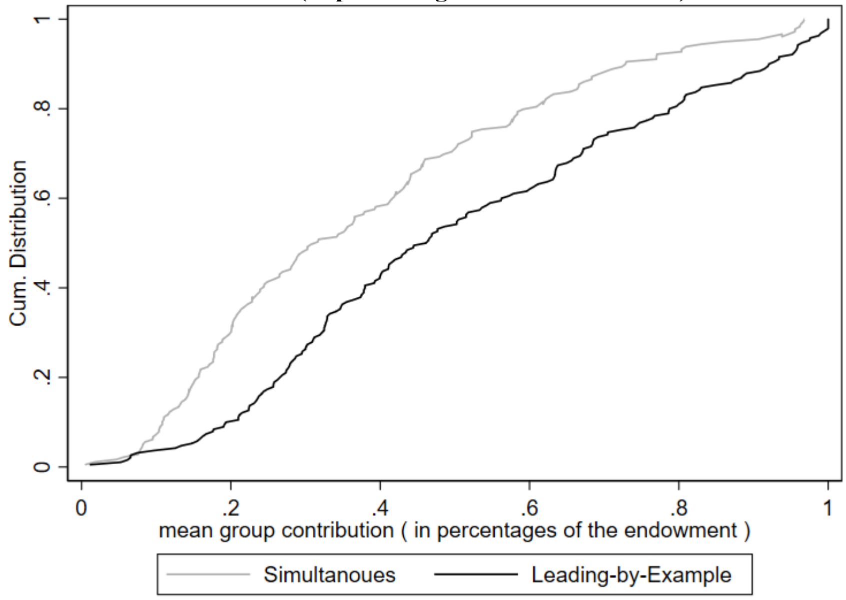
Note: Standard Deviations in parentheses

2.4.1 Does Leading-by-Example increase aggregate cooperation?

To address our first research question, we include only those studies on our subsequent analysis that allow a comparison between sequential and simultaneous contributions (i.e., with and without leader). This leaves us with 11 studies[5], including 369 groups (179 for simultaneous, 190 for Leading-by-Example) as independent observations. Figure 2-1 demonstrates the cumulative distribution of means from all included groups separated by the contribution (in percentages of the endowment). The distribution for groups with a simultaneous contribution structure is always above the distribution for groups with a leader. A two-sample Kolmogorov-Smirnov test confirms that the distributions are statistically different ($p < .01$).

[5] Included studies: Dannenberg (2015); Eisenkopf (2020); Eisenkopf and Kölpin (2022); Eisenkopf and Walter (2021); Gächter and Renner (2018); Gürerk et al. (2018); Güth et al. (2007); Moxnes and Van der Heijden (2003); Rivas and Sutter (2011); Sahin et al. (2015); Yu and Kocher (2020).

Figure 2-1: Cumulative distribution of the mean contribution per group (in percentages of the endowment)



To address our first research question in greater detail, Table 2-3 shows the overall impact of Leading-by-Example on contributions. The dependent variable is a subject's contribution to the public good as a percent of the endowment, with standard errors clustered at the group level. The simultaneous decision structure serves as the benchmark. The variable Leading-by-Example denotes the dummy variable for the groups with leaders. Model I studies the impact of Leading-by-Example. The coefficient of the variable is highly significant and indicates a positive impact of a leader. The effect of Leading-by-Example remains positive and highly significant even if we control for characteristics of the experimental public good environment (model II) or include fixed effects for the studies (model III).

Result 2.1: *Leading-by-Example enhances cooperation in comparison with simultaneous decisions.*

Table 2-3: Leading-by-Example in comparison with simultaneous decisions

Dep. Var.: Indiv. contribution in percentages of the endowment	**Both Treatments Benchmark: Simultaneous** **I**	**II**	**III**
Leading-by-Example	0.127***	0.110***	0.110***
	(0.030)	(0.025)	(0.025)
Group size		0.117***	0.092***
		(0.015)	(0.021)
Exchange rate (in €)		-1.491***	-39.413
		(0.360)	(92.601)
MPCR		0.127	0.685***
		(0.124)	(0.103)
Endowment		-0.001**	-0.050
		(0.001)	(0.112)
Total number of rounds		0.018***	0.366
		(0.004)	(0.773)
Fixed Effects for Studies	--	--	✓
Constant	0.414***	-0.303***	-2.298
	(0.021)	(0.114)	(3.621)
Observations	23,544	23,544	23,544
Groups	369	369	369
R-squared	0.026	0.137	0.156

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

2.4.2 To which extent do group members follow the leader?

Next, we focus on the impact of Leading-by-Example on followers. We use the data from groups with leaders of the 11 studies from the previous subsection, but we also include groups from three other studies (Centorrino and Concina, 2013; Drouvelis and Nosenzo, 2013; Frackenpohl et al., 2016) that did not allow comparison with simultaneous contribution structures but investigated Leading-by-Example in other contexts. In Appendix A.II.I we provide robustness checks without the three additional studies. Excluding these three studies does not alter the results significantly. Table 2-4 presents OLS-regressions with the individual contribution as a percent of the endowment as the dependent variable. We include the characteristics of the experimental public good environment as independent variables. In addition, the dummy variable *Leader not fixed* indicates whether the leader remains in her role during the experiment (=0) or whether the leader role is changed between rounds (=1). Models I - III address contribution differences between leaders and followers. These models include a dummy variable which indicates whether a subject is in the role of a leader (=1) or a follower (=0). All three models show a positive and highly significant coefficient for the *Leader* variable. Thus, leaders contribute significantly more to the public good than followers. Moreover, contributions in groups with a fixed leader are higher than in groups with different leaders.

Table 2-4: The impact of Leading-by-Example

Dep. Var.: Indiv. contribution in percentages of the endowment	**Leading-by-Example**					
	Leader & Follower			**Only Followers**		
	I	**II**	**III**	**IV**	**V**	**VI**
Leader	0.090*** (0.012)	0.112*** (0.010)	0.111*** (0.010)			
Rel. Leader contribution				0.676*** (0.023)	0.692*** (0.023)	0.705*** (0.022)
Group Size		0.076*** (0.019)	0.074** (0.029)		-0.002 (0.013)	-0.007 (0.016)
Exchange rate (in €)		0.471 (0.374)	-3.114 (8.550)		-0.245 (0.275)	5.380 (4.512)
MPCR		0.446** (0.174)	0.653*** (0.127)		0.495*** (0.152)	0.794*** (0.157)
Endowment		-0.000 (0.001)	-0.012* (0.007)		-0.001** (0.001)	-0.002 (0.005)
Total number of rounds		0.014*** (0.005)	0.042 (0.068)		0.016*** (0.004)	-0.027 (0.035)
Leader not fixed	-0.16*** (0.038)	-0.13*** (0.044)	-0.166** (0.073)	-0.056** (0.027)	-0.071** (0.028)	-0.1*** (0.033)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.531*** (0.021)	-0.231* (0.136)	-0.231 (0.425)	0.089*** (0.013)	-0.32*** (0.111)	-0.098 (0.233)
Observations	14,392	14,392	14,392	10,676	10,676	10,676
Groups	248	248	248	248	248	248
R-squared	0.028	0.080	0.108	0.396	0.419	0.461

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

To identify the impact of the example set by the leader, model IV - VI in Table 2-4 include only observations of the followers. Again, the dependent variable is the individual contribution as percentages of the endowment. The variable *Rel. Leader contribution* indicates the contribution of the leader in a given round. In all estimation models, this variable is positive and highly significant. However, both models indicate that followers, on average, employ an imperfect matching strategy.

Table 2-5: Standard deviations of individual contributions within groups

Dep. Var.: Standard deviations within groups	**Both Treatments Benchmark: Simultaneous**		
	I	**II**	**III**
Leading-by-Example	-0.020*	-0.029***	-0.029***
	(0.011)	(0.009)	(0.010)
Group Size		0.039***	0.040***
		(0.010)	(0.012)
Exchange rate (in €)		0.170	-30.799
		(0.206)	(32.870)
MPCR		0.122	0.037
		(0.076)	(0.091)
Endowment		-0.000	-0.040
		(0.000)	(0.040)
Total number of rounds		-0.004**	0.247
		(0.002)	(0.274)
Fixed Effects for Studies	---	---	✓
Constant	0.190***	0.074	-0.929
	(0.007)	(0.064)	(1.290)
Observations	6,126	6,126	6,126
Groups	369	369	369
R-squared	0.004	0.126	0.133

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

The previous findings confirm that Leading-by-Example influences follower behavior. Our prediction 2.2 predicts that this impact leads to a stronger alignment of the group members' contributions. Thus, the heterogeneity of individual contributions should be smaller in groups with a leader. Table 2-5 shows the alignment of group members contributions in detail. The estimations reported in Table 2-5 use the data from the 369 groups of the 11 studies that allow for a comparison between Leading-by-Example and simultaneous contributions. We use the standard deviation within a group in each round as measure of intra-group heterogeneity. The variable *Leading-by-Example* denotes a dummy variable which indicates the groups with leaders. All three models show that the standard deviation within groups is significantly lower in the Leading-by-Example treatments in comparison with simultaneous contributions. Thus, Leading-by-Example leads to a stronger alignment of contributions within the group. This observation confirms our prediction 2.2. Note, however, that the effect vanishes when we exclude Eisenkopf and Walter (2021) or Güth et al. (2007) (see Table A-V in Appendix A.II).

> ***Result 2.2:*** *Leading-by-Example leads to a stronger alignment of the contributions of the group members.*

2.4.3 What is the long-term impact of Leading-by-Example?

To answer our third research question, Table 2-6 present results from an OLS-regression, which focus on long-term effects of Leading-by-Example. The first three models estimate the long-term effect for Leading-by-Example in comparison with simultaneous contribution structures. To estimate the long-term effect, we only include the groups of the 11 studies used in section 2.4.1. The dependent variable is the individual contribution as percentages of the endowment. *Leading-by-Example* denote the dummy for groups with a leader. Model II controls for characteristics of the experimental public good environment, while model III includes fixed effects for studies. Our estimations show that contributions generally decrease over time. To test the long-term differences between groups with and without leaders, we implement an interaction term between Leading-by-Example and the round in our estimations. The interaction term enters positively and significantly in all three models. Thus, while during an experiment the contributions generally decrease, the effect is less pronounced in groups with leaders than in groups with simultaneous contributions. Note, however, that this finding is not robust against all specifications. More precisely, the interaction term for model I - III turns insignificant if we exclude Eisenkopf and Walter (2021) from the analysis. However, the Leading-by-Example as well as the round effect remain highly significant. In Appendix A.II we provide the results from regressions in which we exclude this study.

Model IV-VI now focus on the long-term effect in groups with leaders. For our estimations we rely on the groups of the 14 studies used in section 2.4.2. Again, in Appendix A.II.I we provide robustness checks that exclude the three additional studies. The variable *Leader* denotes whether a participant held the role of leader (=1) or a follower (=0). Model V controls for characteristics of the experimental public good environment, while model VI additionally control for the studies. Again, our models show that contributions decrease during an experiment. Leaders generally contribute more than followers. However, we do not find that leader and followers do react differently to the progress of the experiment which is in line with our prediction 2.3.

Result 2.3: *Leading-by-Example also has a positive effect on contributions in the long-term view. Leaders do not decrease their contributions more than followers.*

Table 2-6: Long-term impact of leadership

Dep. Var.: Indiv. contribution in percentages of the endowment	**Both treatments**			**Leading-by-Example**		
	Benchmark: Simultaneous					
	I	**II**	**III**	**IV**	**V**	**VI**
Leading-by-Example (LbE)	0.108***	0.082***	0.080***			
	(0.0320)	(0.0284)	(0.0294)			
LbE*Round	0.0050**	0.0049**	0.0049**			
	(0.0024)	(0.0023)	(0.0022)			
Leader				0.097***	0.11***	0.104***
				(0.0171)	(0.0162)	(0.0155)
Leader*Round				-0.0008	0.0003	0.0008
				(0.0013)	(0.0012)	(0.0011)
Round	-0.01***	-0.01***	-0.01***	-0.0023	-0.01***	-0.01***
	(0.0017)	(0.0014)	(0.0014)	(0.0018)	(0.0015)	(0.0015)
Group Size		0.116***	0.092***		0.076***	0.074**
		(0.0152)	(0.0216)		(0.0195)	(0.0289)
Exchange rate (in €)		-1.50***	-43.1579		0.4705	18.10***
		(0.3651)	(93.0367)		(0.3739)	(5.9670)
MPCR		0.2587**	0.685***		0.4455**	0.653***
		(0.1246)	(0.1028)		(0.1737)	(0.1268)
Endowment		-0.002**	-0.0510		-0.0004	0.0006
		(0.0006)	(0.0936)		(0.0009)	(0.0047)
Total number of Rounds		0.022***	0.3983		0.018***	-0.129**
		(0.0042)	(0.7974)		(0.0055)	(0.0504)
Leader not fixed	-0.19***	-0.13***	-0.116**	-0.17***	-0.13***	(omitted)
	(0.0391)	(0.0433)	(0.0588)	(0.0377)	(0.0443)	
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.494***	-0.30***	-2.3872	0.552***	-0.2264*	0.6710
	(0.0214)	(0.1114)	(4.2298)	(0.0220)	(0.1367)	(0.4268)
Observations	23,544	23,544	23,544	14,392	14,392	14,392
Groups	369	369	369	248	248	248
R-squared	0.0482	0.1606	0.1761	0.0292	0.0883	0.1154

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** $p<0.01$, ** $p<0.05$, * $p<0.1$*

2.4.4 Do contributions increase with group size?

Last, not least, we focus on group sizes effects on contributions. Table 2-7 presents results from an OLS-regression based on the groups of the 11 studies that compare Leading-by-Example and simultaneous contribution structures (see section 2.4.1). The dependent variable is the individual contribution to the public goods as percentages of the endowment, with standard errors clustered at the group level. Model I shows that contributions increase with larger groups, but that contributions are higher in the Leading-by-Example treatments. Model

II examines the differential effects of group size on the two treatments. It suggests a positive effect of group size, but indicates that this effect is smaller in groups with a leader. These findings remain highly significant even when we control for the public goods environment (model III) and include study fixed effects (model IV).

Table 2-7: The impact of group size

Dep. Var.: Indiv. contribution in percent of the endowment	**Both treatments** **Benchmark: Simultaneous** **I**	**II**	**III**	**IV**
Leading-by-Example (LbE)	0.109*** (0.026)	0.328*** (0.095)	0.346*** (0.093)	0.370*** (0.093)
Group Size	0.093*** (0.012)	0.122*** (0.014)	0.148*** (0.016)	0.126*** (0.023)
Group Size * LbE		-0.054** (0.022)	-0.058*** (0.021)	-0.063*** (0.021)
Exchange rate (in €)			-1.469*** (0.365)	-43.893 (92.826)
MPCR			0.115 (0.125)	0.685*** (0.103)
Endowment			-0.001* (0.001)	-0.055 (0.094)
Total number of Rounds			0.018*** (0.004)	0.406 (0.796)
Fixed Effects for Studies	---	---	---	✓
Constant	0.042 (0.048)	-0.075 (0.059)	-0.432*** (0.119)	-2.655 (4.223)
Observations	23,544	23,544	23,544	23,544
Groups	369	369	369	369
R-squared	0.097	0.103	0.144	0.164

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

To examine the effects of group size on Leading-by-Example in more detail, we focus below on the groups of the 14 studies that focused on Leading-by-Example (see section 2.4.2). Again, Appendix A.II.II provides robustness checks that exclude the three additional studies. Table 2-8 presents results from an OLS-regression. Model I – III focus on the leaders, while the remaining models consider the followers. The first two models confirm our prediction 2.4. Leaders increase their own contributions in larger groups. The effect remains significant even if we control for other characteristics of the experimental public good environment (model II) and when we include the fixed effects for studies (model III). Turning to the followers, model IV shows that followers also increase their relative contributions with increasing group sizes.

However, the group size variables turn insignificant once we control for the leader contribution, whereas the leader contribution variable enters positive and highly significant (model V). Model VI and VII identifies the impact of group size on the leader-follower relationship. Model VI shows a group size effect for the followers. However, the effect vanishes if we control for the leader's contribution (model V). Model VI and VII investigate the coordination impact of the leader. Both models show that the coordination impact of leaders on followers becomes weaker for increasing group sizes. However, model I indicates that leaders in smaller groups are more timid. Hence, even though the leader has less coordination power in the larger groups she generates more contributions. Note, however, that the group size effect for followers in model IV turns insignificant when we exclude Sahin et al. (2015). All other models do not change significantly (see Table A-VII in Appendix A.II).

Table 2-8: The impact of group size in Leading-by-Example treatments

Dep. Var.: Indiv. contribution in percent of the endowment	**Leading-by-Example**						
		Leader			**Follower**		
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Group Size	0.10***	0.11***	0.10***	0.06***	-0.011	0.044**	0.07***
	(0.015)	(0.018)	(0.027)	(0.017)	(0.011)	(0.019)	(0.025)
Rel. Leader contr. (RLC)					0.69***	0.96***	1.10***
					(0.024)	(0.096)	(0.092)
Group Size * RLC						-0.1***	-0.1***
						(0.025)	(0.025)
Exchange rate (in €)		1.12***	2.32***				-3.0***
		(0.380)	(0.456)				(0.486)
MPCR		0.559**	0.77***				0.79***
		(0.222)	(0.238)				(0.157)
Endowment		-0.000	-0.006				-0.01**
		(0.001)	(0.005)				(0.003)
Total number of Rounds		0.010*	0.008				0.04***
		(0.006)	(0.008)				(0.006)
Leader not fixed	-0.1***	-0.1***	-0.087	-0.2***	-0.06**	-0.07**	-0.2***
	(0.046)	(0.050)	(0.057)	(0.038)	(0.027)	(0.027)	(0.035)
Fixed Effects for Studies	---	---	✓	---	---	---	✓
Constant	0.23***	-0.243	-0.264	0.26***	0.16***	-0.079	-0.7***
	(0.069)	(0.160)	(0.301)	(0.073)	(0.039)	(0.066)	(0.191)
Observations	3,692	3,692	3,692	10,700	10,676	10,676	10,676
Groups	248	248	248	248	248	248	248
R-squared	0.080	0.097	0.121	0.050	0.397	0.400	0.468

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

The result confirms our theoretical prediction from prediction 2.4. Rather reluctant leadership in small groups explain this gap even though leaders in large teams elicit less coordinated responses from their fellow group members.

Result 2.4: *Contributions increase with in group size, but the coordination impact of leaders decreases.*

2.5 Conclusion

Experimental studies become more and more important in leadership studies, mainly for three reasons. First, they eliminate endogeneity concerns in the identification of causal relationships. Second, they can be tailored to test a specific theory. Third, they allow for replications by other researchers. We exploited these benefits and merged data from 14 studies in a meta-analysis to answer four questions about the impact of exemplary leadership in the light of a theory that combined standard economic reasoning with a simple model of reciprocity.

First, we hypothesized that Leading-by-Example increases contributions in a social dilemma. The results support this hypothesis. The establishment of a first-moving leader generates significantly higher contributions in comparison to groups without a leader. Our second focus was on the alignment of decisions between leader and followers. Our model predicted that conditional cooperators follow the leader's decision, while the rest refuse to make any contribution. As a result, leaders with high contributions will end up worse than followers (the ‘leader’s curse’). We observe that leadership generates a greater alignment of group members' contributions even though some followers contribute much less than their leader. Hence, on average, leaders contribute more than followers. We then inquired whether higher contributions with Leading-by-Example persist over time. Our model predicted that even selfish leaders should not stop contributing with sufficiently few likely free riders in their group because their own economic losses from a breakdown in cooperation are too large. The results support these insights. Despite the relatively small gains of leaders, they do not reduce their contributions more than followers over time which establishes a positive long-term effect of Leading-by-Example. These findings highlight the importance of cooperative leadership for successful groups. Last but not least, our fourth question deals with the impact of group size. While our simple model predicts an ambiguous effect, further considerations of Bayesian updating suggest that the effect of Leading-by-Example is stronger in larger groups. Our results show that contributions increase with group size, independent of the contributions structure. This result is consistent with previous literature examining the effects of group size (Goeree et al., 2002; Isaac et al., 1994; Isaac and Walker, 1988b; Zelmer, 2003). At the same time, however, our results show that the effect of a leader decreases as group size increases. Further analysis shows that this is particularly related to the fact that leaders in larger groups elicit fewer coherent responses from their followers. This result suggests that the benefits of Leading-by-Example do not extent beyond a certain group size.

We hope that future meta-analyses can rely on a larger and more diverse sample. Such studies could also test, and potentially falsify, specific extensions of our rather simple leadership model. Moreover, we did not investigate any leadership instruments such as communication, monitoring or punishment. Furthermore, most leaders emerge endogenously within a group or that they come as outsiders into the group. Nevertheless, we consider our results as encouraging because they derive from a systematic, replicable, and theory-guided research agenda that may complement and inspire future research in the lab and the ‘real life’.

Chapter 3: Leadership and Cooperation in Growing Teams

3.1 Introduction

Successful leaders foster trust and cooperation within their team and mitigate the adverse incentive effects of imperfect contracts (Boezeman and Ellemers, 2014; Güth et al., 2007; Palanski and Yammarino, 2011). This leader-induced voluntary cooperation raises the aggregate benefits of the team. Hence, people from less effective teams try to join those better performing ones. However, many successful leaders do not want large teams. There are obvious limits, such as increased coordination costs or a limited supply of potential members with the right set of skills. Hence, Steve Jobs, cofounder of Apple Inc., claimed that '*a small team of A+ players can run circles around a giant team of B and C players'* while Amazon's CEO Jeff Bezos advises the 'two pizza rule'[6] to determine team size.

Perhaps less obvious is the threat that newcomer may turn out to be mere opportunists who undermine the cooperative team culture. Google's project 'Aristotle' indicates that trust is a key factor for effective and cooperative teams (Google, 2014). In this context, leadership is critical to fostering trust (Dirks and Ferrin, 2002; Hassan and Ahmed, 2011) and ensuring cooperation (Dannenberg, 2015; Eichenseer, 2021; Güth et al., 2007). However, the effectiveness of leadership as well as trust among members apparently decreases with team sizes (Komai and Grossman, 2009; La Macchia et al., 2016; Soboroff et al., 2020). Hence, successful teams might prefer to stay small to maintain an effective cooperative culture. They establish institutions to manage access to the team and provide for an assessment and voting process before they let new members enroll. In a hiring process, they carefully look at a candidate's capacity for teamwork.

Our paper investigates a causal relationship between leadership, team size and cooperation. More specifically, we study how the endogenous growth of successful teams affects leadership effectiveness and cooperation. This focus implies an obvious interest in the impact of team size on the leader-follower relationship but we also study trust among followers. We test whether participants see a trade-off between a team's openness to strangers and its cooperative culture and therefore. We also test whether they keep their teams small in order to maintain voluntary cooperation even if larger teams could provide a pareto-dominant outcome. The impact of team size and leadership on the dynamics of internal trust are difficult to disentangle with observations from the field. Idiosyncratic factors like the leader's charisma or integrity

[6] According to a statement by Bezos at the Pathfinder Awards 2016, the ideal is the 'two pizza team': '[…] No teams should be so large that it cannot be fed with just two pizzas […]'.

play an important in this context (Avolio and Gardner, 2005; Conger et al., 2000). We therefore deploy a lab experiment with random assignments of participants into teams and roles that examines size effects in teams with leadership.

To address our research questions, we put the participants into matching groups of six people and subdivide these matching groups into one or two teams, depending on the treatment. All teams face the same social dilemma. The participants make voluntary contributions to a club good, which is, like a public good, non-rivalrous in consumption but accessible for the respective team members only (Buchanan, 1965). Any contribution imposes net private costs on the contributor but generates a benefit for the fellow team members. More specifically, an additional team member does not impose any direct costs and her contributions raise the payoffs of all team members.

In some treatments, one team has a randomly chosen leader who can foster cooperation in her team by visibly committing herself to a contribution before the others can do so. Such a leader can act as a ‘prominent’ agent, whose visible actions can establish social norms of high cooperation (Acemoglu and Jackson, 2015). This type of leadership is, in general, rather effective (Dannenberg, 2015; Figuieres et al., 2012; Güth et al., 2007; Moxnes and Van der Heijden, 2003; Pogrebna et al., 2011) even though it does not change the underlying economic incentives against contribution.

Our experimental design consists of seven treatments. In our three benchmark treatments the participants remain in their team throughout the experiment. The first benchmark treatment uses the design as explained above. The matching group is divided into two independent teams of three members and one of the teams has a randomly chosen leader. In the other two benchmark treatments, all six participants in a matching-group are in one team of six members. In one of these two treatments, a randomly chosen team member leads the group and contributes ahead of the others.

We compare the evidence from this exogenous variation in team size with results from three treatments with endogenous variations. These treatments allow for migration between these teams at the end of a round. In two of the three treatments, one team has a randomly chosen leader. Both treatments differ in the mechanism of this endogenous team formation. In one of these treatments, all participants (except for the leader) can change the teams without further restrictions (free entry/exit-mechanism). In the other treatment, the current team members can also reject potential entrants with a majority vote (restrictive mechanism). We compare these results with a third change-treatment in which all participants are allowed to change teams without restriction, but none of the teams has a leader. Last, but not least, we have a treatment that replicates the benchmark treatment with two separate teams of three members. However, each participant receives information about the contributions of all members of the matching group. This treatment bridges the gap between the benchmark treatments and our change treatments, because the disclosure of all contributions is an additional variation in these change treatments. Behavioral spillovers across teams are likely to influence cooperation within the teams (Böhm and Rockenbach, 2013; Burton-Chellew and West, 2012; Tan and Bolle, 2007).

It is a feature of our design that new team members do not impose any direct cost on the incumbents. Instead, the maximum team size of six establishes the opportunity for the most efficient outcome. By design, all other detrimental factors of increased group size (e.g., coordination and communication costs) do not apply in this experimental setup. Several experiments show a positive correlation between group size and contributions in public goods experiments for constant MPCRs (Goeree et al., 2002; Isaac et al., 1994; Isaac and Walker, 1988b; Zelmer, 2003). Hence, team members have an incentive to integrate more people into their team. The efficiency gains from new members rules out direct costs as an explanation for any decline in the incumbent's contribution with increasing team size or any outright rejection of potential new members.

In the context of our research questions, the leader's behavior and its impact on the other team members is particularly interesting. Because the leader's response to new team members provides a direct signal of (dis-)trust to the followers in the team. A decline in the coordination success with increasing team size then informs about the relative importance of this fear for the other team members. Note that the comparison with fixed groups controls for the team size effect in this context. This exogenous variation also represents a novel contribution to the literature.

Our focus on leadership provides a new perspective to the literature on endogenous team formation. Previous studies show that exclusion mechanisms reduce free-riding and that low contributors are more likely to be expelled from the initial group (Chakravarty and Fonseca, 2017; Cinyabuguma et al., 2005; Maier-Rigaud et al., 2010). In contrast, we introduce an admission process in which incumbents can reject the access of other participants to the group.

This mechanism reveals skepticism towards newcomers and their effect on cooperation. Ahn et al. (2008) show that such an admission process affects the level of contributions although the direction of the effect depends on the entry and exit rules. Our experiment extends the findings in important ways. We introduce effective leadership as a determinant of endogenous growth and study, whether growth undermines the effectiveness of leadership. We also compare these results with exogenous variations in team size.

Our experimental results show that individual contributions indeed react to different team sizes. Individual contributions to the club good increase with larger teams, but the coordination effect of the leaders diminish. Moreover, in all three change-treatments, virtually all participants want to join the team with higher contributions. In treatments with a leader, they overwhelmingly attach themselves to the team with the leader. In the treatment with endogenous access restrictions, low contributors are initially from entry into that team. As a consequence, we find significantly higher per-capita contributions in comparison with the unrestricted entry and exit.

The remainder of the paper is organized as follows. Section 3.2 explains the experimental design in greater detail. In section 3.3, we present the behavioral predictions. Section 3.4 presents our main results, while section 3.5 concludes.

3.2 The experimental design and procedures

At the beginning of the experiment, we randomly allocated the subjects in matching groups of six participants each. Depending on the treatments, each matching group included either one or two teams. All participants in the experiment played a variation of the standard voluntary contribution mechanism game (Isaac and Walker, 1988a, 1988b). In each of the 20 rounds, a player received an endowment of 100 points.[7] She could keep them in a private account or contribute them towards a club good in order to enhance the welfare of all team members.

Only the actual team members benefited from the club good. The members of the other team in the matching group benefited from their own club good. The payoff function for individual i in a team of size n is

[7] We exchanged 200 points into 1€, about 1.12 US-$ in early 2019.

$$\pi_i = 100 - c_i + 0.5 * \sum_{j=1}^{n} c_j \qquad \text{(E3.1)}$$

in which c_i denotes $i's$ contribution to the club good with $0 \leq c_i \leq 100$. One point invested implied a marginal per-capita return for each club member of 0.5 points, which implies that there is no rivalry in consumption of the club good. The marginal per capita benefit of contribution does not change with team size, but the aggregate social benefit increases in the number of team members. More specifically, in teams with three or more members it is efficient to contribute the entire endowment.

Baseline treatments

We observed contributions in three baseline treatments where team size was systematically varied exogenously. In the *SMALL-TEAM Treatment,* we divided our matching group in two teams, with three participants per team. These assignments did not change across the 20 rounds. The two teams played the VCM game described above in different sequences. In team 1, all team members contributed simultaneously to the club good. In team 2, one randomly chosen team member was put into a leadership position. This person remained in that role throughout the experiment. The leader contributed first. The other team members – but not the members from the leaderless team – could observe the leader's contribution when they decided about their own. At the end of the round, the participants learned about the individual contributions of the others team members (but not the other team's) and their own payoff in that round.[8]

We compare the results with two Large-Team-Treatments. In both treatments all six members of the matching group formed one team. In the *LARGE-TEAM Treatment (w/o Leader)*, all six team members decided simultaneously about their contributions. In the *LARGE-TEAM Treatment (w Leader)* one randomly chosen team member was assigned into a permanent leadership position. The other five team members could observe the leader's contribution when they decided about their own. At the end of the round, all six participants learned about the individual contributions of the others and their own payoff.

The comparison of the two teams in the *SMALL-TEAM Treatment* as well as the comparison between the two *LARGE-TEAM Treatments (w Leader* and *w/o Leader)* assesses the impact of leadership on contributions in different team sizes. Moreover, the comparison

[8] Technically, the separation into two completely independent teams with or without leadership is equivalent to a separation into two different treatments. However, unlike in all other treatments, participants in one team in this Small-Team-Treatment also studied the instructions of the other team.

between the *SMALL-TEAM* with the *LARGE-TEAM Treatments* assesses the impact of team size on the individual contributions.

Change-Treatments

We compare our three benchmark treatments with four additional treatments. Three of these treatments allow for team changes. These three change-treatments started like the *SMALL-TEAM Treatment.* All matching group members were initially divided into two teams of three participants each. Both teams played the VCM game for 20 rounds. At the end of a round, all matching group members learned about the contributions of all other matching group members. After observing the individual contributions, all matching group members could decide about leaving one team for the other (free entry/exit mechanism). Hence, the size of the individual teams could vary across the rounds. In the *OPEN-CHANGE Treatment (w/o Leader)* both teams were leaderless. This treatment measures the impact of group changes and shows whether a coordination effect towards an efficient group is formed. We compare these results with the *OPEN-CHANGE Treatment (w Leader).* In this treatment, one of the two teams had a randomly chosen leader. As in the benchmark treatments, this leader visibly commits herself to a specific contribution. After observing individual contributions at the end of a round, all matching group members, except the leader, could decide about leaving one team for the other. Since we expect participants to join the team with the leader (see predictions section below) our interest in this treatment focuses on how an *endogenous* change in team size influences individual contributions in that particular team.

The *RESTRICTED-CHANGE Treatment* used the same procedure as the *OPEN-CHANGE-Treatment (w Leader)* with one addition: In each team, the current members could prevent members of the other team from joining their own team. This occurred via a voting mechanism at the end of a round. When participants observed the individual contributions of all matching group members of the previous round, they could decide whether they would like to reject or accept each member of the other team. If more than 50% of the current members of the receiving team accepted a participant, she could choose to join or not join the team. If less than 50%, she had to stay in her old team in any case. Overall, we get detailed information about who wants to change teams and which people accept or reject such potential new team members.

Since the change treatments revealed the contributions of all members of the matching group at the end of a round, the *PEER Treatment* measured the effects of spillover on contributions. The procedure of the *PEER Treatment* is identical to the *SMALL-TEAM Treatment*. The difference between the two treatments is that the *PEER Treatment* discloses the contributions of all matching group members, whereas the *SMALL-TEAM Treatment* reveals only the contributions within one's own team. Thus, the design of the *PEER Treatment* explicitly allows for spillover effects between the teams. A comparison between the *SMALL-TEAM Treatment* and the *PEER Treatment* provides us with information about peer effects and their impact on contributions. Table 3-1 summarizes the design and reports the number of subjects assigned to each treatment.

We conducted twelve sessions in late 2018 and early 2019 and additional nine sessions in the summer of 2021. The experiment was programmed with the help of z-tree and conducted at the WiSo-Forschungslabor (University of Hamburg). Recruitment was done via hroot (Bock et al., 2014). Each subject participated just once. Each session lasted approximately 90 minutes. Within each session, the participants were randomly assigned into the different treatments, matching groups and role. Every participant sat on a randomly assigned computer. The instructions were handed out, as well as read out loud. Control questions were provided to ensure a common level of understanding of the experiment. All subjects received their payment (on average 19.83 €, or about 22.92 US-$ in autumn 2021) privately at the end of their session.

Table 3-1: Experimental Design

	Small-Team		**Large-Team (w/o Leader)**	**Large-Team (w Leader)**	**Open-Change (w/o Leader)**		**Open-Change (w Leader)**		**Re-stricted-Change**		**Peer**	
	Team 1	Team 2	Team 1	Team 2	Team 1	Team 2	Team 1	Team 2	Team 1	Team 2	Team 1	Team 2
Stage 1 Random appointment of the leader (only in round 1)	---	✓	---	✓	---	---	---	✓	---	✓	---	✓
Stage 2 Leader invests and team members learn about the investment	---	✓	---	✓	---	---	---	✓	---	✓	---	✓
Stage 3 (Remaining) Team members invest simultaneously	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stage 4 a) Team members learn about the contributions of the own members and of the members of the other team	---	---	---	---	✓	✓	✓	✓	✓	✓	✓	✓
b) Team members learn about the contributions of the own members	✓	✓	✓	✓	---	---	---	---	---	---	---	---
Stage 5 Team members can prevent members of the other team from joining the team	---	---	---	---	---	---	---	---	✓	✓	---	---
Stage 6 Subjects decide about changing the team	---	---	---	---	✓	✓	✓	✓	✓	✓	---	---
Subjects	54		54	66	66		66		72		66	
Matching Groups	9		9	11	11		11		12		11	

3.3 Behavioral predictions

All teams in the experiment face the same social dilemma. The social benefit of club good contributions increases in the number of team members. In any team with at least three members, it is efficient to contribute the entire endowment to the club good. Meanwhile the private return from a one-point-contribution to the club good is half a point. Hence, zero contributions constitute the Nash equilibrium if we assume common knowledge that all participants just care about their own private payoff. However, it is well known that many people are conditionally cooperative. They contribute to the aggregate welfare of a team if they expect others to do so as well (Fischbacher et al., 2001; Fischbacher and Gächter, 2004; Keser and Van Winden, 2000). Evidence on Leading-by-Example suggests that sequential contribution structures generally increase cooperation. In particular, several studies show that group members who are informed about the contribution of a first-mover (or leader) make significantly higher contributions than members in groups with simultaneous contributions (Dannenberg, 2015; Eichenseer, 2021; Eisenkopf, 2020; Güth et al., 2007; McCannon, 2018; Moxnes and Van der Heijden, 2003; Pogrebna et al., 2011).

Prediction 3.1: *In the SMALL-TEAM Treatment, contributions to the club good are higher in the team with a leader than in the leaderless team.*

Studies investigating Leading-by-Example in larger groups are rather limited. From a theoretical point of view, the effect of group size on Leading-by-Example is ambiguous (Figuieres et al., 2012). While on the one hand, and especially in our design, larger groups may generate higher social benefits, a larger group size implies a higher probability of at least one purely selfish team member. Such a person may induce other team members to lower their contribution as well. However, the literature on group size effects in simultaneous contribution structures suggests that the positive effects should prevail and the contributions increase with group size (Eisenkopf and Kölpin, 2022; Isaac et al., 1994; Isaac and Walker, 1988b; Zelmer, 2003).

In context of our research question, the impact of team size on Leading-by-Example and especially on the leader's behavior is particularly interesting. Figuieres et al. (2012) show that in smaller groups contributions, on average, are significantly larger with sequential public good contributions than in a simultaneous setting, while they only find insignificant differences for larger groups. Sahin et al. (2015) also find no significant differences in contributions between groups with and without leaders in groups of six players. However, their experimental design does not allow for comparison with smaller groups. Komai and Grossman (2009) show that the effectiveness of Leading-by-Example decreases with group size, since followers do not mimic the leader's behavior in larger groups as much as they do in smaller groups. A meta-analysis by Eisenkopf and Kölpin (2022) confirms this finding. They show that the leader's coordination power – measured as the correlation between the contributions of the leader and the followers – decreases with group size.

Prediction 3.2:

1. *In both LARGE-TEAM Treatments (with and without leader), contributions are higher than in the respective teams in the SMALL-TEAM Treatment.*
2. *The correlation between the contribution of the leader and the contributions of the followers is smaller in the LARGE-TEAM Treatment with leader than in the team with leader in the SMALL-TEAM Treatment.*

Now the focus turns on the change in team size. Here the per capita contributions in the *LARGE-TEAM (with and without leadership)* and the two teams in the *SMALL-TEAM Treatment* provide the benchmarks. In our three Change-Treatments participants can migrate to the other team after round 1 (in the *RESTRICTED-CHANGE Treatment* only with permission). Since contributions in our game generate a public good without rivalry in consumption, each additional team member generates additional benefits (or at least does not impose any costs) for the fellow team members. Thus, all participants benefit from moving into larger groups with high levels of cooperation. Hence, we expect to replicate the results from previous studies on endogenous group formation showing that participants migrating into the more cooperative team (Ahn et al., 2008; Brekke et al., 2011; Charness and Yang, 2014). As discussed above in prediction 1, we expect higher contributions in teams with a leader which implies a net migration towards them. Thus, we predict that participants migrate into the efficient team (in treatments with leadership: into the team with a leader). Furthermore prediction 2 states that per capita contribution should increase with team size, while the additional positive leadership-effect decreases. As a consequence of the previous statements, the following prediction emerges:

Prediction 3.3: *In the treatments with endogenous team composition*

1. *Participants migrate into the team with higher contributions (for treatments with leadership: into teams with leadership).*
2. *Per capita contributions increase in team size but the additional benefit of leadership decreases.*

Finally, we focus on the access vote in the *RESTRICTED-CHANGE Treatment.* By design, team members do not get any immediate material benefit from rejecting other participants. However, conditionally cooperative team members have an interest in denying perceived free-riders access to their team, e.g. because of equality concerns (Fehr and Schmidt, 1999), perceived altruism (Levine, 1998) or if they expect that other incumbent team members reduce their contributions in consequence. Experimental studies on restrictive team formation show that low contributors are more likely to be expelled or not allowed to join the group (Ahn et al., 2008; Chakravarty and Fonseca, 2017; Cinyabuguma et al., 2005; Maier-Rigaud et al., 2010; Swope, 2002). Moreover, Ahn et al. (2008) as well as Weber (2006) show that a managed growth of teams significantly increases cooperation. Hence, such a restrictive entry mechanism should foster the per-capita contributions for a given team size.

Prediction 3.4:

1. *In the RESTRICTED-CHANGE Treatment, participants in the team with leaders are more likely to accept high contributors than low contributions from the other team.*
2. *For a given team size, contributions in teams with leaders are higher in the RESTRICTED-CHANGE Treatment than in the OPEN-CHANGE Treatment (w Leader).*

3.4 Results

Before we test the predictions in detail, we have a brief look at the descriptive statistics. Table 3-2 shows the average contributions of participants across treatments and teams. If applicable, the table also differentiates between contributions of leaders and followers. In the three treatments with changes between the teams we also report the contributions of incumbent followers (those who have started the experiment in the respective team) and newcomers who had changed from the other team. For the *OPEN-CHANGE Treatment (w/o leader)*, a distinction between team 1 or 2 is meaningless because no team has any leader. Instead, we report the average team size across all rounds and distinguish between the larger and the smaller team. If not noted otherwise, all subsequent nonparametric tests use average contributions of the relevant members in a matching group across all 20 rounds as an independent observation. All reported test results are two-sided.

Table 3-2: Descriptive Statistics

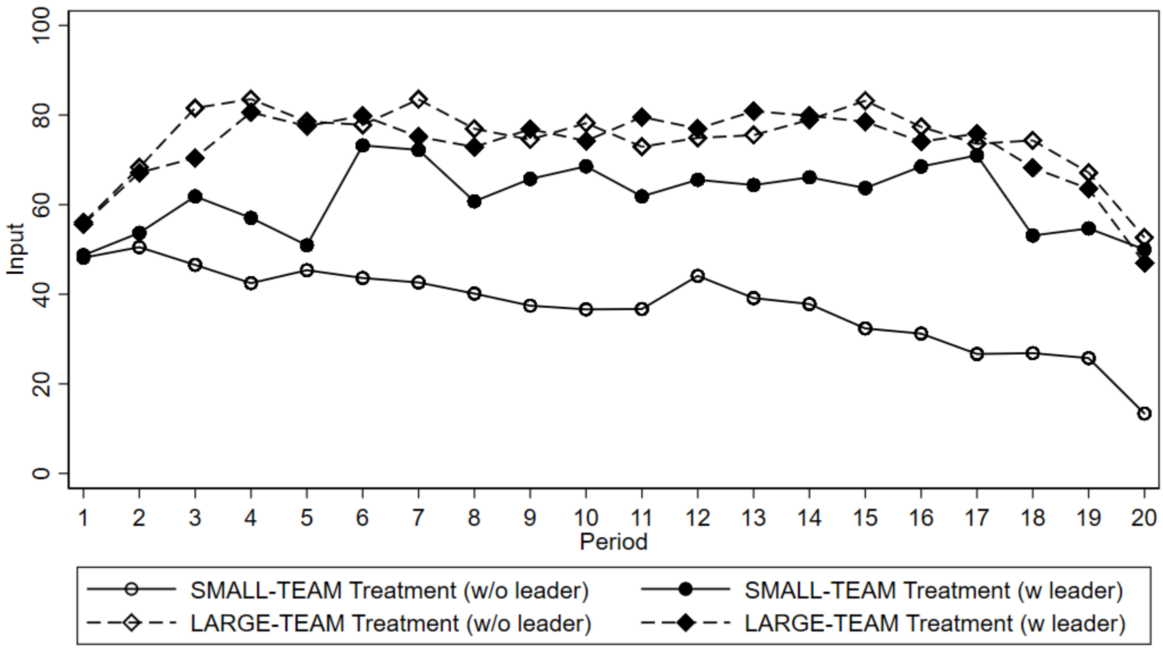
Treatment	**Team**	**Team Size**	**Contributions: All**	**Leader**	**Follower: All**	**Follower: Incumbents**	**Follower: Newcomer**
SMALL-TEAM (Matching Groups = 9)	Team 1 (w/o leader)	3	37.38 (31.28)	---	---	---	---
	Team 2 (w Leader)	3	61.58 (24.42)	62.16 (26.48)	61.29 (23.73)	---	---
LARGE-TEAM (w/o Leader) (Matching Groups = 9)	Team 1 (w/o leader)	6	74.48 (17.61)	---	---	---	---
LARGE-TEAM (w Leader) (Matching Groups = 11)	Team 2 (w Leader)	6	72.74 (20.05)	84.55 (18.09)	70.38 (21.62)	---	---
OPEN-CHANGE (w/o Leader) (Matching Groups = 11)	Smaller (w/o leader)	0.58	36.87 (12.92)	---	---	---	---
	Larger (w/o leader)	5.42	64.31 (17.02)	---	---	63.99 (15.65)	64.93 (23.33)
OPEN-CHANGE (w Leader) (Matching Groups = 11)	Team 1 (w/o leader)	0.53	45.33 (23.2)	---	---	---	---
	Team 2 (w Leader)	5.48	65.22 (31.34)	73.59 (25.98)	63.35 (33.13)	61.20 (35.41)	64.66 (33.12)
RESTRICTED-CHANGE (Matching Groups = 12)	Team 1 (w/o leader)	1.75	57.64 (30.11)	---	---	---	---
	Team 2 (w Leader)	4.25	67.25 (28.6)	70.77 (37.02)	75.06 (17.19)	73.64 (16.25)	82.03 (17.31)
PEER (Matching Groups = 11)	Team 1 (w/o leader)	3	39.21 (34.38)	---	---	---	---
	Team 2 (w Leader)	3	44.91 (31.39)	48.53 (30.36)	43.1 (32.23)	---	---

Note: Standard deviation in parentheses

3.4.1 Leading-by-Example in small and large teams

Figure 3-1 displays the average contributions of both teams in the *SMALL-TEAM* and both *LARGE-TEAM Treatments* (with and without leadership) over the rounds. Our results show that contributions differ significantly between the two teams in the *SMALL-TEAM Treatment* ($p < 0.1$, Wilcoxon signed-rank test), while we find fairly similar contributions over time for the two *LARGE-TEAM Treatments* (with and without leader) ($p = .91$, Mann-Whitney-U-Test). Moreover, we find that contributions in teams without a leader are significantly higher in larger than in smaller teams ($p < .05$, Mann-Whitney-U-Test). With leadership, however, we only observe insignificant differences ($p = .31$, Mann-Whitney-U-Test).

Figure 3-1: Average contributions in the SMALL-TEAM- and LARGE-TEAM-Treatments for both teams



To test the effect of team size on contribution in detail, Table 3-3 presents the results of a GLS random effects regression. The dependent variable is the per capita contribution to the club good, with standard-errors clustered at the matching-group level. The variables *Small-Team (w Leader), Large-Team (w/o Leader)* and *Large-Team (w Leader)* denote dummy variables for the respective teams. For model I, the *Small-Team (w/o Leader)* serves as the benchmark. The estimates in model I indicate two salient results. First, leadership enhances contributions in small teams. Second, larger teams (with and without leadership) experience on average higher contributions than small teams without leaders.

The remaining models II and III now investigate the team size impact on leadership. Model II measures the leadership impact in large teams. The *LARGE-TEAM Treatment (w/o Leader)* servers as the benchmark. The model shows that the contribution differences between large teams with or without leader are only marginal and insignificant. Last but not least, model III investigates the impact of team size on leadership. Therefore, model III only considers teams with a leader. The *SMALL-TEAM Treatment (w Leader)* denotes the benchmark. The model shows only insignificant differences in contributions between all teams with a leader.

Table 3-3: Individual contribution to the club good in the SMALL-TEAM and LARGE-TEAM Treatments

Dep. Var.: Per Capita Contribution to club good	**All Teams**	**Only Large Teams**	**Only Teams with Leader**
Benchmark:	SMALL-TEAM w/o Leader **I**	LARGE-TEAM w/o Leader **II**	SMALL-TEAM w Leader **III**
Small-Team (w Leader)	24.20**		
	(9.528)		
Large-Team (w/o Leader)	37.10***	1.743	
	(11.49)	(8.308)	
Large-Team (w Leader)	35.36***		11.16
	(11.68)		(9.940)
Round	-0.358	-0.233	-0.120
	(0.238)	(0.270)	(0.355)
Constant	41.14***	75.18***	62.84***
	(10.29)	(5.981)	(7.970)
Observations	3,480	2,400	1,860
Subjects	174	120	93
Matching Groups	29	20	20
R-Squared	0.113	0.002	0.02

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

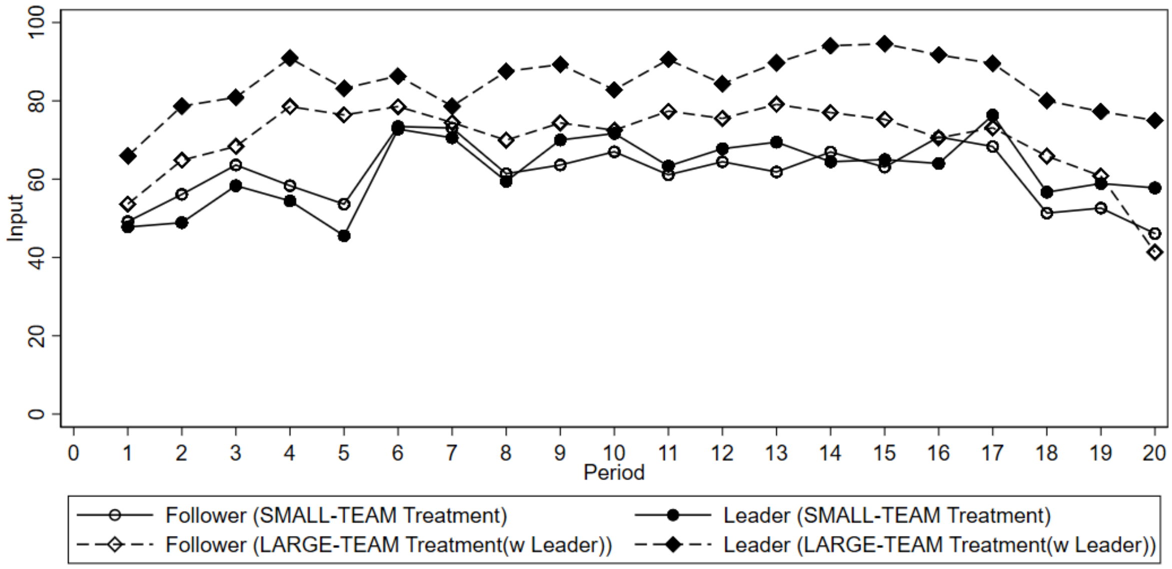
Result 3.1:

1. *In small teams, Leading-by-Example leads to higher contributions.*
2. *Larger teams without leader generally experience higher contributions than smaller teams.*
3. *With leadership, the contributions in small and large teams do not differ significantly.*

3.4.2 Leadership in small and large teams

We now study the impact of team size on per capita contributions in teams with leadership in greater detail. The preceding results (see result 1 and Table 3-2) show that in teams without leaders, contributions are generally higher in larger than in smaller teams. However, these differences are no longer present in teams with leaders. To address this observation in greater detail, we now disentangle the differences between the teams with leadership in the *SMALL-TEAM Treatment* and the *LARGE-TEAM Treatment*. Figure 3-2 shows the average leaders' and followers' contributions in both treatments over the twenty rounds. It displays higher contributions for larger teams in each round. However, the differences in contributions are significant for the leaders only ($p < .05$), but not for the followers ($p = .52$, Mann-Whitney-U-Test). In addition, we find that leaders contribute significantly more than followers in the *LARGE-TEAM Treatment* ($p < .01$), but not in the *SMALL-TEAM Treatment* ($p = .68$, Wilcoxon signed-rank test). However, for both small and large teams, we find a highly and significant correlation between the contributions of the leaders and the followers ($r = .77$ for the *SMALL-TEAM Treatment*, $r = .46$ for the *LARGE-TEAM Treatment*, $p < .01$, Pearson correlation). Figure 3-2 also shows interesting differences in leader behavior. While the followers' contributions are influenced by their leader's choices, the leaders' initial contributions provide an exogenous variation. We observe that the contributions of leaders in the *LARGE-TEAM Treatment* are already higher in the first round than those of leaders in the *SMALL-TEAM Treatment* ($p < .1$, Mann-Whitney-U-Test). This difference does not vanish until the end of the experiment. This observation indicates a positive impact of team size on the contributions of leaders.

Figure 3-2: Average leader and follower contributions over the rounds in the SMALL- and LARGE-Team Treatments



To test the previous observations in greater detail, Table 3-4 studies the interaction between leaders and followers in these two treatments. It provides results from random effect GLS regressions. The dependent variable is a subject's contribution to the club good, with standard errors clustered at the level of the matching group. The *SMALL-TEAM Treatment (w Leader)* serves as the benchmark, *Large-Team (w Leader)* denotes the dummy variable for the other treatment. Model I studies the contributions of the leaders only. It shows that leaders in larger teams contribute significantly more than leaders in smaller teams. This result confirms our previous observation in Figure 3-2. Models II-IV now focus on the contributions of the followers. Model II indicates only insignificant differences in contributions of the followers, while model III shows that leaders strongly influence the decisions of the follower with their own contribution. They achieve a rather high degree of coordination. If a leader increases her contribution by one point, the followers increase their contributions by about .57 points. Model IV identifies the impact of team size on the leader-follower relationship. It shows that this coordination impact of leaders on followers is about .26 points stronger in the *SMALL-TEAM Treatment* than in the *LARGE-TEAM Treatment*.

Table 3-4: Follower and Leader contributions to the club good in the SMALL- and LARGE-TEAM Treatments

Dep. Variable: Per capita Contribution to club good	**Benchmark: SMALL-TEAM (w Leader)**			
	Leader	**Follower**		
	I	**II**	**III**	**IV**
Large-Team (w Leader)	22.39**	9.081	-3.757	14.57*
	(10.09)	(9.968)	(5.615)	(8.400)
Leader contribution			0.573***	0.739***
			(0.0768)	(0.103)
LT* Leader contribution				-0.260**
				(0.131)
Round	0.417	-0.267	-0.476**	-0.477**
	(0.501)	(0.336)	(0.190)	(0.195)
Constant	57.78***	64.10***	30.66***	20.40***
	(9.565)	(7.661)	(5.176)	(4.858)
Observations	400	1,460	1,460	1,460
Subjects	20	73	73	73
Matching Groups	20	20	20	20
R-Squared	0.105	0.001	0.312	0.312

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

The preceding observations give us insights into how team size shapes the impact of leadership. First, leaders increase their contributions as team size increases. Second, for a given level of leader contributions, we find no differences in contributions of the followers. Third, we find that followers mimic the leader's contributions to a greater extent in smaller teams than in larger teams. Thus, in conjunction with Result 1, these observations suggest that Leading-by-Example has a rather limited additional effect in larger teams due to the uncoordinated responses of followers.

Result 3.2: *Leaders in large teams contribute significantly more than leaders in smaller teams. However, the statistical relationship between the contribution of the leader and the contributions of the followers is significantly smaller.*

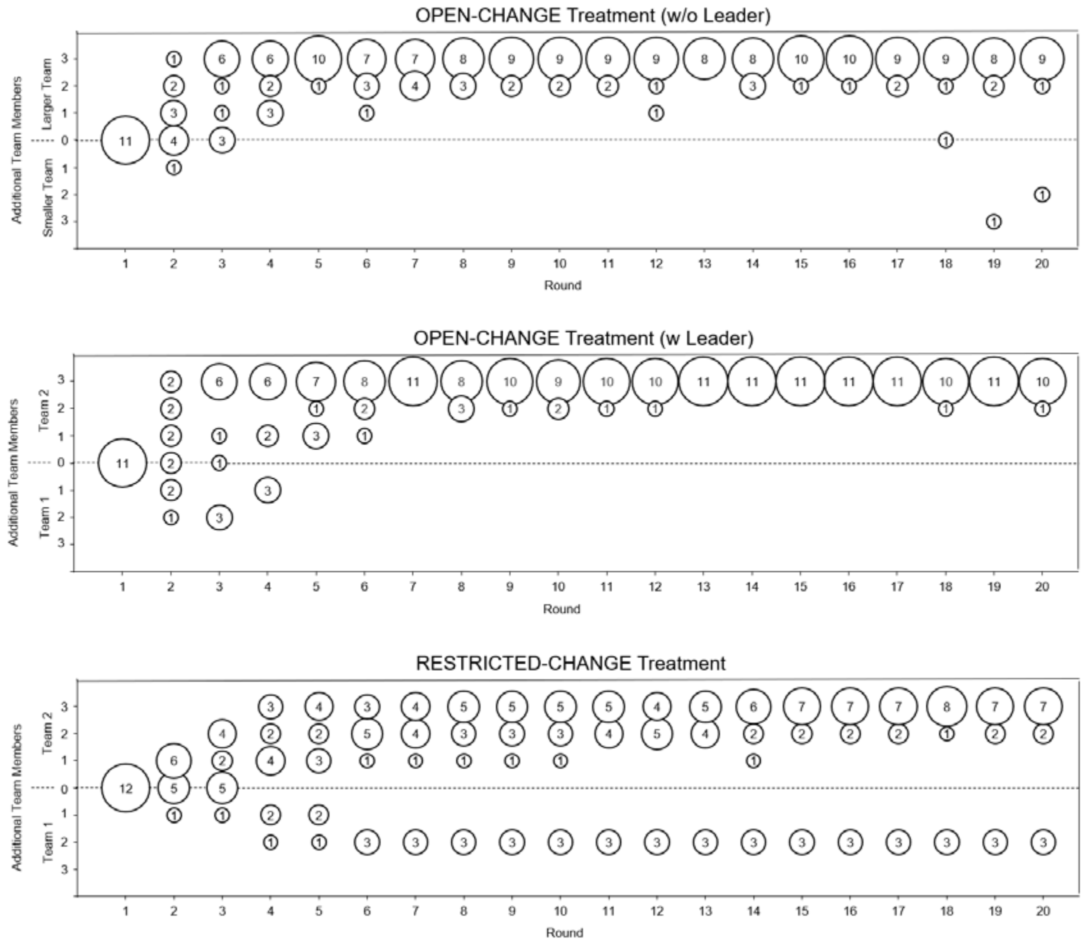
3.4.3 Endogenous team formation and cooperation

We now focus on the treatments in which participants can change the teams. Figure 3-3 depicts the moving patterns of the matching group members in the respective treatment. More specifically, the figure shows how many additional team members are in team 1 or team 2 in a corresponding round. The numbers in the bubbles (and their size) indicate the number of matching groups in which this movement is found.

For all treatments, we find that participants eventually come together to form larger teams. For teams with leadership (*OPEN-CHANGE (w Leader)* and *RESTRICTED-CHANGE Treatment*), we find that almost all participants settle in the team with leader (team 2). Only in three out of twelve matching groups in the *RESTRICTED-CHANGE Treatment*, the participants move into the leaderless team, leaving the leader alone. We remove these three matching groups from the subsequent analysis.[9] Apparently, a coincidence of high contributions in the team without leader and low contributions of the leader induced this outcome (see Appendix B.IV).

[9] Removing these three matching groups does not alter the main results. We report replications of the subsequent estimations including all matching groups in the appendix B.V.

Figure 3-3: Additional team members in both OPEN-CHANGE-Treatments and the RESTRICTED-CHANGE Treatment



Note: The numbers in the bubbles (and their size) indicate the number of matching-groups in which this movement is found. The figure contains all matching groups, including those we exclude from our subsequent analysis (see below). Note that for the Open-Change-Treatment (w/o leader), we used the average team size across all rounds to distinguish between larger and smaller teams.

The preceding observations suggest that participants form larger teams and that leaders can encourage this migration. Next, we disentangle the motives behind these changes. Table 3-5 presents results from a probit-regression. Since the change decision in the *RESTRICTED-CHANGE Treatment* depends on the upstream voting mechanism (see the experimental design in section 3.2), we focus only on both *OPEN-CHANGE Treatments (with and without leadership)*. The dependent variable is the individual decision to change the team in a given round. The variable $\bar{x}_{\text{Other}} - \bar{x}_{\text{Own}}$ indicates the relative cooperation levels of the two teams. Models I and II include both *OPEN-CHANGE Treatments*, while the *OPEN-CHANGE Treatment (w Leader)* denotes the benchmark. Both models show that a change becomes more likely the more the other team contributed compared to the own team. An interaction term reveals no differences between the treatments (model II). The remaining models focus only on the *OPEN-CHANGE Treatment (w Leader).* Models III and V confirm our previous finding, while models

IV and VI take the contribution of the leader into account. Models IV and VI show significant but opposite effects of the leader's contribution. The leader's contribution increases the likelihood of changing for team 1 members, but team 2 members are more likely to refrain from changing when the leader's contribution increases.

Table 3-5: Individual decision to change into the other team

Dep. Variable: Ind. decision to change into the other team	Both OPEN-CHANGE Treatments		Only OPEN-CHANGE (w Leader)			
	Benchmark: w Leader		Only Team 1		Only Team 2	
	I	II	III	IV	V	VI
Open-Change (w/o L) (OC)	0.198	0.222*				
	(0.125)	(0.133)				
$\bar{x}_{Other} - \bar{x}_{Own}$	0.01***	0.01***	0.02***	0.01***	0.05***	0.05***
	(0.0022)	(0.0042)	(0.00306)	(0.00358)	(0.0102)	(0.00775)
OC*$\bar{x}_{Other} - \bar{x}_{Own}$		0.00536				
		(0.00538)				
Leader contribution				0.00504*		-0.03***
				(0.00274)		(0.00652)
Team size	-0.46***	-0.47***	-0.335*	-0.368*	0.638**	1.225***
	(0.0664)	(0.0687)	(0.179)	(0.203)	(0.320)	(0.342)
Round	-0.0256	-0.0250	0.0290	0.0252	-0.36***	-0.43***
	(0.0272)	(0.0262)	(0.0444)	(0.0431)	(0.0970)	(0.0707)
Constant	0.916***	0.917***	0.954	0.808	-3.62***	-4.55***
	(0.268)	(0.258)	(0.629)	(0.737)	(1.177)	(1.241)
Observations	666	666	114	114	162	162
Subjects	132	132	43	43	55	55
Matching Groups	22	22	11	11	11	11
Pseudo R-Squared	0.305	0.308	0.290	0.297	0.488	0.538

Note: Probit-Regression. Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

We now study how this endogenous change in team size influences subsequent contributions. Table 3-6 reports results from random-effects GLS-regressions that estimate the statistical relationship between team size and individual contributions for all three change treatments (model I and II) and only for change treatments with a leader (model III – V). For all models, the *OPEN-CHANGE Treatment (w Leader)* provides the benchmark. *Open-Change (w/o Leader)* and *Restricted-Change* denote dummies for the other treatments. For the *OPEN-CHANGE (w Leader)* as well as the *RESTRICTED-CHANGE Treatment* we include only the teams with a leader (team 2). For the *OPEN-CHANGE Treatment (w/o Leader)* we include the larger team.

Models I and II indicate a positive and highly significant coefficient for the team size variable. Moreover, model I shows that contributions in the *RESTRICTED-CHANGE Treatment* are somewhat higher than in the *OPEN-CHANGE Treatment (w Leader).* However, we find nearly the same contributions for both *OPEN-CHANGE Treatments*, indicating a rather limited leadership effect. Model II controls for an interaction term between the treatments and team size. Both interaction terms enter insignificant. Thus, with restrictions in team access, we see more cooperative behavior for a given team size in comparison with a free entry. The remaining models now focus only on the change treatments with a leader. Model III shows that leader's contributions are higher in the *RESTRICTED-CHANGE Treatment* and that contributions increase with larger teams. Model IV indicates the same findings for the followers. However, controlling for the leader's contributions, model V shows that the impact of team size is lower among followers, while they follow the leader's contributions to a high degree.

Result 3.3: *In the treatments with endogenous team composition*

1. *participants migrate into the team with higher contributions. Cooperative leaders foster the migration.*
2. *per capita contributions increase in team size. Without restrictions, we do not find differences between endogenous growth with and without leadership.*

Table 3-6: Per capita contributions in the Change-Treatments

Dep. Var.: Per Capita Contribution to club good	**All Change Treatments**		**Only Change-Treatments with a leader (team 2)**		
		Benchmark: OPEN-CHANGE (w Leader)			
			Leader	**Follower**	
	I	**II**	**III**	**IV**	**V**
Open-Change (w/o Leader)	-0.393	6.776			
	(10.27)	(15.33)			
Restricted-Change	19.20*	12.82	18.84**	19.72*	10.04
	(10.12)	(18.11)	(8.548)	(10.67)	(7.961)
Team Size (T. Size)	8.209***	8.388***	8.948***	9.295***	4.715***
	(1.341)	(2.423)	(1.912)	(2.035)	(1.784)
Open-Change (w/o L)*T. Size		-1.271			
		(3.195)			
Restricted-Change*T. Size		1.184			
		(3.779)			
Leader contribution					0.530***
					(0.0872)
Round	-1.258***	-1.272***	-0.603	-1.672***	-1.313***
	(0.278)	(0.275)	(0.450)	(0.399)	(0.308)
Constant	32.06***	31.19***	30.91***	28.54***	10.67
	(8.902)	(10.61)	(9.589)	(10.33)	(10.50)
Observations	3,339	3,339	400	1,746	1,746
Subjects	185	185	20	99	99
Matching Groups	31	31	20	20	20
R-Squared	0.076	0.076	0.145	0.091	0.380

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

Next, we are interested in the acceptance of new team members. In total, 563 acceptance decisions were made with an acceptance rate of 31.79 percentage (179 out of 563). Table 3-7 presents a probit estimation on the individual decision to accept a newcomer. The dependent variable is a dummy variable which indicates whether a voter has accepted (=1) or rejected (=0) a new member. The variable *New Member* specifies whether the participant started in the team or initially came from the other team. Our models show that the contribution level of a newcomer is the key determinant for an acceptance decision. Neither the own contribution of the voting incumbent member nor the team size do influence this decision significantly. Moreover, we do not find a significant difference neither between leaders and followers nor between new and old members.

Table 3-7: Individual's decision to accept a new member in the RESTRICTED-CHANGE Treatment

Dep. Var: Indiv. decision to accept a new member	**I**	**II**	**III**	**IV**	**V**
Contribution in t (New Member)	0.022***				0.022***
	(0.00298)				(0.00302)
Contribution in t (Voter)		0.000401			0.00298
		(0.00364)			(0.00273)
Team Size (Team 1)			0.264		0.0338
			(0.217)		(0.162)
Leader				0.0310	-0.0414
				(0.234)	(0.279)
New Member	-0.210	-0.141	-0.0564	-0.130	-0.219
	(0.279)	(0.207)	(0.206)	(0.191)	(0.243)
Round	0.0253	-0.0226	0.00489	-0.0224	0.0273
	(0.0323)	(0.0268)	(0.0371)	(0.0268)	(0.0319)
Constant	-1.396***	-0.320	-0.970	-0.300	-1.693***
	(0.334)	(0.360)	(0.635)	(0.250)	(0.339)
Observations	563	563	563	563	563
Subjects	48	48	48	48	48
Matching Groups	9	9	9	9	9
Pseudo R-Squared	0.27	0.01	0.02	0.01	0.27

Note: Probit-Regression. Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

Result 3.1 and 3.2 above in the paper showed that an exogenous increase in team size leads to more cooperation. Evidence from Table 3-6 (and result 3.3) reveals the same qualitative effect for an endogenous variation. Now, we compare these two observations. Table 3-8 reports results from random-effects GLS-regression models that estimate individual contributions in teams with different team sizes across teams with a leader. The *RESTRICTED-CHANGE Treatment* provides the benchmark. All models contain control dummies for the team size. Model I focuses on the leaders. The model shows significantly lower leader contributions in the *OPEN-CHANGE Treatment (w Leader)*. However, the *RESTRICTED-CHANGE Treatment* does not differ significantly from the *LAGE-TEAM-* and *SMALL-TEAM Treatment*. Models II and III now focus on the followers. We find that follower contributions are positively and significantly influenced by leader contributions. However, model III suggests that the effects of leader's contribution differ across treatments. More specifically, the coordination effect is significantly higher in the *RESTRICTED-CHANGE Treatment* than in the *LARGE-TEAM* and *OPEN-CHANGE Treatment (w Leader)*. Consequently, the restrictive entry mechanism reduces the lower coordination effect of leaders in larger teams. Model IV includes both leaders and followers. Model IV shows significant lower contributions in the *OPEN-CHANGE Treatment (w*

Leader), but only insignificant differences in contributions in comparison with the *LARGE-TEAM* (p = .137) and the *SMALL-TEAM Treatment* (p = .778). This finding suggests that the impact of self-selection in teams on individual cooperation is rather limited. In Appendix B.VI, we provide additional regressions with a more selective dataset. In particular, we include only teams with six members. However, even in teams with the efficient team size, the results do not alter significantly.

Result 3.4:

1. *In the RESTRICTED-CHANGE Treatment, participants in the team with leaders are more likely to accept participants from the other team with relatively high contributions.*
2. *For a given size of teams with leadership, contributions are higher in the RESTRICTED-CHANGE Treatment than in the OPEN-CHANGE Treatment (w Leader). Leaders in the RESTRICTED-CHANGE Treatment elicits more coordinated responses than in the LARGE-TEAM and in the OPEN-CHANGE Treatment.*

Table 3-8: Per capita contributions across Treatments in teams with a leader

Dep. Var.: Per Capita Contribution to club good	**Only teams with a leader Benchmark: RESTRICTED-CHANGE**			
	Leader	**Follower**		**All**
	I	**II**	**III**	**IV**
Small-Team (w L)	-10.11	9.580*	19.13***	2.884
	(11.89)	(5.580)	(6.977)	(10.24)
Large-Team (w L)	-10.49	-6.621	20.02**	-11.91
	(7.322)	(5.711)	(7.898)	(8.011)
Open-Change (w L)	-18.00**	-7.925	21.69***	-17.65*
	(8.806)	(7.226)	(6.958)	(10.23)
Leader contribution		0.562***	0.788***	
		(0.0572)	(0.0194)	
Small-Team*Leader contribution			-0.0411	
			(0.113)	
Large-Team*Leader contribution			-0.300***	
			(0.0869)	
Open-Change*Leader contribution			-0.346***	
			(0.0786)	
Round	0.0494	-0.828***	-0.854***	-0.691**
	(0.361)	(0.186)	(0.186)	(0.278)
Dummies for team size are included	✓	✓	✓	✓
Constant	49.23***	34.08***	18.87***	41.22**
	(18.74)	(7.587)	(7.002)	(18.30)
Observations	800	3,206	3,206	4,006
Subjects	40	172	172	212
Matching Groups	40	40	40	40
R-Squared	0.126	0.349	0.338	0.047

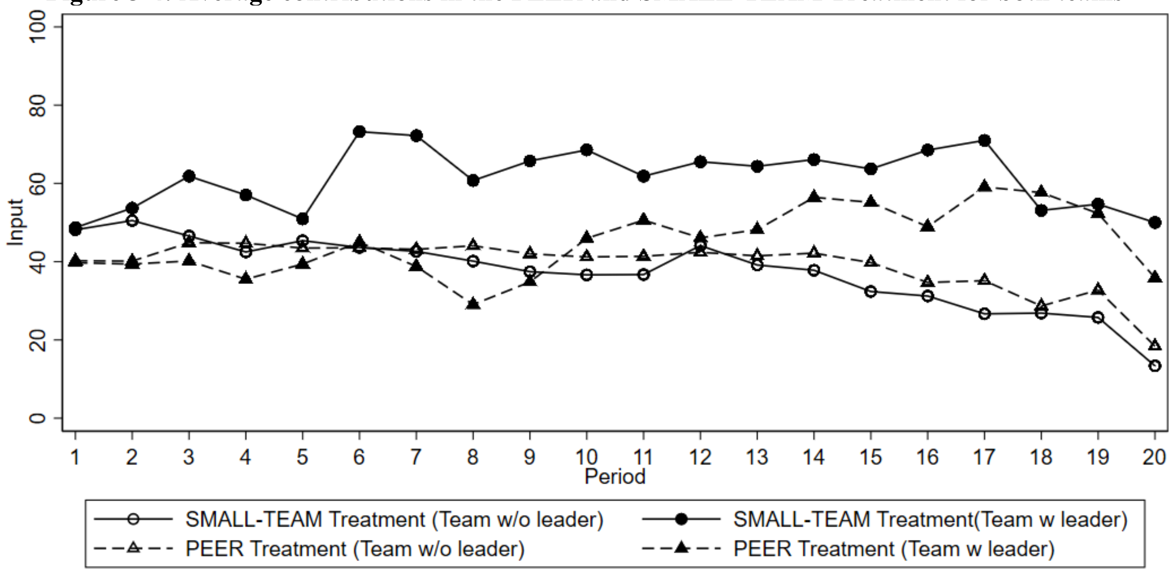
*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

3.4.4 Peer-Effects

In all three Change-treatments, participants learn about the contributions of all members from both teams in their matching group at the end of a round. In this section, we therefore examine the potential spillover or peer effects of this information on individual contributions. Figure 3-4 display the average contributions of both teams in the *PEER Treatment* and the *SMALL-TEAM Treatment* over the rounds. We find quite similar contributions over time for both teams in the *PEER Treatment*. The difference in average team contributions across the 20 rounds is insignificant (p = .29, Wilcoxon signed-rank test). Spillover effects may explain this result (see Appendix B.III). However, these estimations are not robust across model specifications.

This finding suggests that information spillovers reduce the aggregate leadership effects, but it does not imply that Leading-by-Example has no behavioral impact. Figure 3-4 shows different levels of contribution for the first and the second half of the experiment. While we find insignificant differences for the first half (p = .83, Wilcoxon signed-rank test), the differences become significant for the second half of the experiment (p < .05, Wilcoxon signed-rank test). In addition, we find that leaders contribute significantly more than their followers (p < .05, Wilcoxon signed-rank test) and that leader and follower contributions are highly correlated (r = .83, p < .01).

Figure 3-4: Average contributions in the PEER and SMALL-TEAM Treatment for both teams



Next, we analyze the impact of peer-effects on contributions in greater detail. Table 3-9 presents results from a random effects GLS-regression. The dependent variable is the per-capita contribution. The *SMALL-TEAM Treatment* denotes the benchmark. Models I and II include both teams. Model I shows that there are no significant differences in contributions between the two treatments, but that contributions are higher in teams with a leader. Model II controls for an interaction term between the treatment and the team variable. However, the interaction term enters insignificant. The three remaining models (III – V) focus only on team with leadership (team 2). Model III looks at the leaders but shows no differences between *SMALL-TEAM Treatment* and the *PEER Treatment*. Models IV and V show that followers respond strongly to the leader's contributions. However, followers in the *PEER Treatment*, on average, contribute less than followers in the *SMALL-TEAM Treatment*.

Result 3.5*: The behavioral impact of leadership persists when all contributions are revealed. However, followers, on average, contribute less.*

Table 3-9: The impact of peer effects

Dep. Variable: Per capita Contribution to club good	Benchmark: SMALL-TEAM Treatment				
	Both teams		Only Team 2		
	All		Leader	Follower	
	I	II	III	IV	V
Peer-Treatment	-7.419	1.834	-13.62	-8.083*	-7.649*
	(10.66)	(14.31)	(12.41)	(4.310)	(4.256)
Team 2 (w Leader)	14.03*	24.20**			
	(8.439)	(9.608)			
Peer-Treatment*Team 2		-18.51			
		(15.85)			
Leader contribution				0.742***	0.748***
				(0.0487)	(0.0830)
Peer-Treatment*Leader contribution					-0.00737
					(0.100)
Round	-0.263	-0.263	0.713	-0.0750	-0.0756
	(0.419)	(0.419)	(0.656)	(0.259)	(0.259)
Constant	45.23***	40.14***	54.67***	15.94***	15.60***
	(9.832)	(10.97)	(10.55)	(3.396)	(3.936)
Observations	2,400	2,400	400	800	800
Subjects	120	120	20	40	40
Matching Groups	20	20	20	20	20
R-Squared	0.04	0.05	0.04	0.67	0.67

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

3.5 Conclusion

Our paper studies how endogenous growth of successful teams' affects leadership effectiveness and cooperation. We disentangle the interaction between leadership and team growth and its impact on the dynamics of internal trust. In general, we observe a virtuous circle between leadership, team size and contributions. As predicted, we find that high contributions of leaders encourage higher per-capita contributions of their followers which foster migration into their teams. In turn, larger teams experience even more courageous leadership and higher contributions, but the coordination effect of the leader diminishes. Such evidence from anonymous interactions certainly supports the benefits of an open team culture. However, the experimental design deliberately favored this outcome because of the non-rivalrous consumption of the club good.

We had a particular interest in whether people see a trade-off between a group's openness to new members and its cooperative culture. We predicted that with a restrictive entry mechanism, incumbents are more willing to accept newcomers with high contributions to maintain a cooperative team culture. Indeed, we find that many participants reveal skepticism about new members to their team. They forego economic benefits from new team members and reject potential entrants, in particular those with a poor contribution record. Hence, incumbents favor the preservation of a rather cooperative in-group culture at the expense of overall economic gains.

While lab experiments often lack external validity, the specific design features of this experiment actually strengthen its implications. As mentioned above, we deliberately devised an experiment in which larger teams are more efficient than smaller teams and in which new team members impose no immediate costs to the incumbent members. The experimental participants were educated people who studied in a cosmopolitan city at a university with a rather liberal reputation for openness and against discrimination. Furthermore, these participants actually experienced the monetary benefits of increasing teams, but nevertheless they were reluctant to appreciate new team members who came from their own social context. Moreover, our results refer to randomly appointed leadership positions, as we are more interested in the effect of leadership itself and do not want to bias our observations by motivations or attitudes of persons who deliberately moved into that prominent position. Studies on endogenous leadership appointments suggest that such leaders contribute even more than randomly appointed leaders (Haigner and Wakolbinger, 2010; Rivas and Sutter, 2011), see themselves as role models (Arbak and Villeval, 2013) and more often put the group's will above their own interests (Kocher et al., 2013). Consequently, self-selected leaders should further strengthen our results. However, this is left for other studies.

We believe that our experimental results may shed light on some management practices that emphasizes team formations. For instance, Google states that it is'[…] less about who is on the team, and more about how the team worked together' (Google, 2014). Our results support this finding in some extends. Indeed, team effectiveness is strongly impacted by how the team worked together. In this context, the evident skepticism to new team members is a rather strong phenomenon, as it enhances the cooperation of the incumbents even as it implies economic costs. Leaders of successful teams have to take into account that many members may have rather parochial preferences.

Chapter 4: Leadership, Cooperation, and the Motivation for Tedious Tasks

4.1 Introduction

Motivating people is one of the most important leadership tasks (Chen and Kanfer, 2006; Schaffer, 2008). Encouraging pro-social motivation induces people to overcome opportunistic incentives for the benefit of their group. Higher task motivation implies that they work longer on their job. The effects of monetary and non-monetary motivational incentives on performance as well as on the interaction between these two sources of motivation have received much attention in the literature (Ariely et al., 2009; Bowles and Polanía-Reyes, 2012; Gneezy and Rustichini, 2000; Huffman and Bognanno, 2018; Ryan and Deci, 2000). However, many leaders believe that the best instrument to motivate people is the own example. John Wooden, head coach of the UCLA basketball team, for example explains that 'a leader's most powerful ally is his or her own example' (Wooden and Carty, 2009). A strong leader who leads by example can successfully increase employee motivation in the workplace by fostering trust and cooperation. For example, Google's research project 'Oxygen' examines which behaviors make great managers perform better. They argue that the leader must act as a role model, motivating and encouraging his followers by setting good examples (Google 2013).

This paper studies causal relationships between leadership, cooperation, and task motivation. More specifically, we investigate whether leadership in a social dilemma affects task motivation. We rely on an innovative design that allows us to use a methodological problem of online experiments to our advantage. Endogenous attrition in online samples usually induces a selection bias that could jeopardize the internal validity of such experiments(Zhou and Fishbach, 2016). Tedious and tiring work tasks foster such premature attrition (Horton et al., 2011; Jun et al., 2017). We turn this bug into a feature and use the attrition rate as a (noisy) signal of task motivation.

Models of social preferences typically assume that people gain utility from making choices that benefit (or harm) others at their own (short-term) expense in a social dilemma (Cox et al., 2007; Falk and Fischbacher, 2006; Fehr and Schmidt, 1999). Nevertheless, cooperation can be rewarding if others reciprocate such actions. Hence, cooperation may reduce attrition because people gain from reciprocity whereas a cooperation failure induces people to leave. Furthermore, we have a closer look at the impact of such a failure. People with a genuine (non-strategic) preference for cooperation may depart from their uncooperative fellows even if the continuation of the experiment provides them with financial rewards. In this case, a lack of cooperation crowds out task motivation. In the context of our research question, we are particular interested in the impact of leadership on task motivation. While successful leadership may enhance team work by fostering trust and encouraging employees to work towards a common goal, poor leadership can have the opposite effect (Dirks and Ferrin, 2002; Güth et al., 2007; Palanski and Yammarino, 2011).

To test this mechanism, we observe attrition in three treatments. We randomly assign participants into groups of up to four members. Each participant must complete a simple but tedious task to receive a payment. Additionally, the participants can either keep their earnings from the task in a private account or put them into a common pool. More precisely, they face a variant of the standard voluntary contribution mechanism (Isaac and Walker, 1988a, 1988b): A contribution to the common pool enhances aggregate welfare but leaves the contributor worse off financially (*COOPERATION Treatment*). Afterwards, all participants can start with a new task. However, access to this task is costly and the costs increase over time, eventually exceeding the payment. Hence, it is sensible to leave the experiment at some stage. We inform the participants that they can take their money and leave anytime, while the fellow subjects continue. Note that all participants in a group can observe the participation and earnings of the fellow group members.

The second treatment starts as described above. However, after completing the tedious work task, a randomly chosen group members decides first about the contribution. The other group members learn about the decision of the leader, before they decide for themselves (*LEADING-BY-EXAMPLE Treatment*). This type of leadership does not change the underlying economic incentives against contributions. However, studies on Leading-by-Example show that this type of leadership rather fosters cooperation (Dannenberg, 2015; Eisenkopf, 2020; Figuieres et al., 2012; Güth et al., 2007; Moxnes and Van der Heijden, 2003; Pogrebna et al., 2011). Hence, we expect substantial differences in cooperation rates between these two treatments as well as between the groups within each treatment.

Last, not least, we compare these to cooperation treatments with a treatment that does not allow for cooperation (*NO-COOPERATION Treatment)*. The *NO-COOPERATION Treatment* starts like the *COOPERATION Treatment*. Therefore, participants complete the tedious work tasks and receive a payment. However, they are no able to put their payment into a common pool.

A comparison of the *COOPERATION* and *LEADING-BY-EXAMPLE Treatments* with the *NO-COOPERATION Treatment* identifies the impact of cooperation on attrition. Voluntary cooperation can alter attrition for two reasons. First, and obviously, successful cooperation raises the economic benefit from doing additional tasks. Members in cooperating groups will continue with the experiment even once the costs of an additional task exceed its potential private returns (but not the collective). Second, cooperation may also alter the task motivation itself, i.e. the readiness to proceed with an additional task. Failed cooperation reveals this second mechanism. We take a particular interest in the initial contributions of the group members on the subsequent attrition. Low initial cooperation rates allow conditionally cooperative players to earn more money even without the help of the fellow group members, just like participants in the *NO-COOPERATION Treatment* do. However, if they drop out of the entire experiment instead, one can infer that failed cooperation diminishes task motivation. A comparison between both cooperation treatments (*LEADING-BY-EXAMPLE* Treatment vs. *COOPERATION Treatment*) reveals the impact of leadership. The role of the leader is particularly interesting here. Such a leader can act as prominent agent, whose visible actions can establish social norms of high cooperation (Acemoglu and Jackson, 2015). As argued above, in general, this type of leadership fosters cooperation. However, poor leadership has the opposite effect (Qiu et al., 2018). Consequently, we are particularly interested in the effects of (good or bad) leadership on task motivation.

Our paper provides a methodological innovation to study the impact of cooperation on endogenous attrition rates and task motivation. While most studies examine a statistical relationship between effort tasks and questionnaire responses about intrinsic motivation (Peng and Hsieh, 2012; Tauer and Harackiewicz, 2004; Walker, 2010), only few studies use an behavioral approach. Carr and Walton (2014) experimentally investigate whether the signal of cooperation can fuel intrinsic motivation even when people work alone. They show that people work longer on a challenging task when they feel they are working together. However, due to the nature of their research question, no conclusions can be drawn about the effects of actual cooperation on intrinsic motivation, which is the focus of our study. Closer to our experiment, Haeckl et al. (2018) provide an experimental measure of work motivation in a real-effort task as a deviation from the money-maximizing benchmark. They show that participants with a positive intrinsic motivation respond less to team incentives and observation by peers than those with a negative intrinsic motivation. In contrast, our design takes advantage of an online experiment in which we have behavioral measures for cooperation (contributions to a common pool) and task motivation (attrition). The cage like environment in labs typically eliminates premature attrition, but it is a (noisy) signal of task motivation in online experiments.

Moreover, our paper adds to the literature on experimental investigation of Leading-by-Example. First, we consider Leading-by-Example in social dilemmas, in which subjects interact without a known termination period. A meta-analysis by Dal Bó and Fréchette (2018) show that in such a context, high cooperation rates emerge only when the parameters of the repeated game are such that cooperation is very robust to strategic uncertainty. Our paper provides a novel approach to this literature because the participants themselves determine the end of the experiment. Hence, while our chosen experimental parameters facilitate cooperation, differences in opportunity costs and technical problems introduce a strategic uncertainty that can undermine it. In turn, Leading-by-Example can address this uncertainty to some extent. Observing a contribution of a randomly selected 'leader' reduces uncertainty and usually leads to a higher willingness to cooperate on the part of the followers (Figuieres et al., 2012; Güth et al., 2007; Moxnes and Van der Heijden, 2003; Pogrebna et al., 2011). Second, we focus on the impact of Leading-by-Example on the interaction between two sources of nonmonetary incentives. Both provide new perspectives to the literature on Leading-by-Example.

Our experimental results show that the degree of cooperation within a group is correlated with attrition. We find that successful cooperation fosters the perseverance of the participants. Moreover, our results suggest that leadership has a two-fold effect on task motivation. We find that in groups with a leader, uncooperative groups face higher premature attrition rates than cooperative groups. More precisely, participants in uncooperative groups sacrifice the economic benefit of an additional task and drop out prematurely. We do not find the same effects for leaderless groups. However, we do not find evidence that attrition rates are caused by the leader's contribution behavior.

The remainder of the paper is organized as follows. Section 4.2 explains the experimental design in greater detail. In section 4.3, we present the theoretical predictions. Section 4.4 presents our main results, while section 4.5 concludes with a discussion of the empirical results.

4.2 The experimental design and procedures

Our focus on attrition suggests that we start with the description of procedural details, and in particular how subjects could enter and leave the experiment. The next subsection will then introduce the details of the experimental game.

4.2.1 The procedure of the online experiment

The experiment was programmed using z-tree (Fischbacher 2007) and conducted with the help of z-tree unleashed (Duch et al., 2020). We contacted participants from the Laboratory for Economic Experiments Vechta (LEEV, University of Vechta) and the Laboratory for Economics Research (LAER, University of Osnabrueck) for separate sessions. Recruitment was done via hroot (Bock et al., 2014) for the participants at the University of Vechta and ORSEE (Greiner, 2015) for the participants at the University of Osnabrueck. Each participant could sign up for one session only, each consisting of 24 participants (28 in one session).

At their registration, the participants were instructed to use either a laptop or computer during the experiment in order to ensure a minimum resolution of the devices.[10] Furthermore, we standardized the virtual screens to a resolution of 1024x768. Participants were also informed to receive an e-mail about 10-15 minutes before the experiment. This e-mail from the experimenter contained a personalized internet-link and a code to access the experiment. The personalized links were generated by ztree-unleashed. Each link represented a zleaf-client.[11] The experimenter started the experimental program at the specified time. The computer program randomly allocated the zleaf-clients into groups of four and assigned a treatment (see below for more details) to each group. Then we displayed a 5-minute countdown on the screen to give time to connect and prepare the participant for the start. Once the countdown expired, the participants entered their private code on the first page. If a participant did not enter the code within a time restriction of 5 minutes, the program ejected her from the experiment. We observe substantial attrition already at this stage (see below and appendix C.II). Thus, groups became smaller if a participant failed to start the experiment correctly. Regarding the other randomly assigned group members, such a failure implies an exogenous variation in initial group size.

[10] We checked this with a control question at the beginning of the experiment. All but two participants complied with this requirement. One of these participants left the experiment after the instructions. The second person played until round 9.

[11] The zleaf-client allows the subjects to communicate with the experimental server. It thus represents the access points to the experiment. For more details see Fischbacher (2007).

The experimental instructions were provided on screen. The participants read the instructions and answered questions about their gender and study program, the device they used and their experience with experimental studies. Participants could leave the experiment at any time. Therefore, we provided a quit button on every screen of the experiment (see screenshots in appendix C.I). To ensure the continuation of the experiment, even if participants were inactive without pressing the quit button (e. g. closing the link), we also introduced a time restriction for each experimental decision.[12] Participants who pressed the quit button or did not make a decision within the time restriction could no longer participate in the experiment. Furthermore, we ruled out negative payouts but did not announce a finite number of rounds. We explained all these mechanisms in the instructions (see appendix C.I in translation from German). With this introductory procedure finished, the first round started.

4.2.2 The design of the experiment and its treatments

As described above, the computer randomly allocated the groups in each session to one of three different treatments. The participants interacted in the same groups (and treatments) until the end of the experiment. The participants played the treatment specific game repeatedly in multiple rounds. The number of rounds is endogenous since participants could leave the experiment at any time. An experimental round followed a sequence of stages, a decision stage, a working stage and, depending on the treatment, a contribution stage (see screenshots in appendix C.I).

The decision and the working stage are identical for all treatments. In the decision stage, participants decide whether they want to complete the work task. This decision is costly, the price started at 0 points in the first round and increased by 1 point per additional round. In the working stage, participants could complete a work task. Participants had to count zeros in a table of 100 randomly ordered zero and ones (Abeler et al., 2011). We restricted the work task to 180 seconds. In case of a wrong answer subjects received an error message and could try again with the same table. Successful participants received 10 points. Due to the increasing task costs, it is therefore sensible to leave the experiment at some stage. If participants did not want or failed to complete the task, they left the experiment. We chose the 0-1-task, for two reasons:

[12] For each instruction page (two pages) 300 seconds each. 180 seconds for the work task. 90 seconds for each other stage. All time restrictions were pretested to ensure a reasonable amount of time.

First, this task is easy and does not require any prior knowledge. Second, this tedious work task entails a positive cost of effort for subjects (Abeler et al., 2011).

However, access to this task is costly and the costs increase over time, eventually exceeding the payment. Hence, it is sensible to leave the experiment at some stage. We inform the participants that they can take their money and leave anytime, while the fellow subjects continue.

The contribution stage varied across the three treatments. Participants in the *NO-COOPERATION Treatment* could not make a contribution decision. For every successfully completed task, a participant received 10 points into her private account. The payoff function for individual i in the *NO-COOPERATION Treatment* in a given round is:

$$\pi_i = 10 - c_t \qquad , if\ i\ still\ participates$$
$$\pi_i = 0 \qquad , if\ i\ has\ left$$

The variable $c_t = t - 1$ captures the costs for entering the task in a given round t.

In the *COOPERATION Treatment*, the participants played a variation of the standard voluntary contribution mechanism game (Isaac and Walker, 1988a, 1988b). First, participants learned about the remaining number of group members. Then they were asked either to keep the 10 points in a private account or to contribute them to a common pool. The contribution to the common pool immediately increased the private account of each group member by 6 points, even if the member had left the experiment.[13] Hence, the social benefit of a contribution does not change with attrition. Thus, a rational and selfish player has an incentive to enhance only the own private account, while it is efficient to distribute the earnings among the group members. Note that the marginal per capita benefit from cooperation $\left(\frac{6}{10}\right)$ is rather generous. Even in groups with two players, a contribution increases aggregate welfare. At the end of each round, all members could observe the investments of all group members. The payoff function for individual i in a given round for a group of size n is:

[13] This payment did not apply to participants who failed to enter the experiment correctly.

$$\pi_i = 10 * (1 - x_i) + 6 * \sum_{j=1}^{n} x_j \ - c_t \qquad , if\ i\ still\ participates$$

$$\pi_i = 6 * \sum_{j \neq i}^{n} x_j \qquad\qquad , if\ i\ has\ left$$

The variable x_i denotes $i's$ contribution to the common pool, with $x_i \in \{0, 1\}$.

The *LEADING-BY-EXAMPLE Treatment* provided the same decision context. However, one randomly chosen group member (the 'leader') decided first. This person remained in that role until she left the experiment. The remaining group members could observe the leader's investment when they decided themselves. The comparison between the two treatments assessed the impact of leadership on cooperation. Moreover, we used this leadership device to generate more variation in cooperation rates across groups without changing the underlying economic benefits from contributions.

As mentioned above, we conducted all three treatments in each session. Note that the participants in a session had to wait until all participants had made their decisions. Only then the next round started. Thus, the time per round did not differ between the treatments within a session. Table 4-1 summarizes the design and reports the number of subjects assigned to each treatment.

We conducted fifteen sessions between May and July 2020. In total, 364 students agreed to participate in our experimental sessions. Out of this sample, 312 showed up for the experiment. In three groups with altogether eight participants in the *LEADING-BY-EXAMPLE Treatment*, the leader did not show up. Although the inclusion of observations would not significantly alter our results (neither if we included observations in *LEADING-BY-EXAMPLE Treatment* nor in the *COOPERATION Treatment*), we did not consider these observations in our subsequent analysis. Thus, our subsequent data analysis relies on observations from 304 participants. Appendix C.II summarizes the reasons for exclusions for each treatment. On average, each session lasted approximately 60 minutes (median: 57 minutes). The university paid the participants by online bank transfer (on average 20.75€ or 23.11 US Dollars). To ensure the strict separation of experimental and payment data, we collected the payment data via an external questionnaire after the experiment. We informed participants about these procedures prior to registration in a session. We linked the payoff and the account details via the participant code. We do not find any substantial attrition in this accompanying survey.

Table 4-1: Experimental Design

Treatment Stage		NO-COOP-ERATION	COOP-ERATION	LEADING-BY-EXAMPLE
Decision and Working Stage	**Step 1** Participants decides to complete the costly work task. The costs are deducted from the private account.	✓	✓	✓
	Step 2 Participants complete the work task and receive 10 points. The other group members learn about the participation and successful completion.	✓	✓	✓
Contribution Stage	**Step 3** Leader decides to keep the 10 points or to increase the private accounts of all group members. Group members learn about the decision.	---	---	✓
	Step 4 (Remaining) group members decide simultaneously.	---	✓	✓
	Step 5 Group members learn about the decisions of the other members.	---	✓	✓
Subjects		105	99	100
Groups		30	30	28

4.3 Behavioral predictions

In the *COOPERATION* and the *LEADING-BY-EXAMPLE Treatments*, participants face a social dilemma in each round. However, the undefined ending does not allow for backward induction to identify the equilibria in this context. It is well known that prisoner dilemma games with random termination yield multiple equilibria which allow for cooperation and/or defection (Dal Bó, 2005; Dal Bó and Fréchette, 2011, 2018, 2019; Normann and Wallace, 2012). Dal Bó and Fréchette (2019) identify three prominent strategies: permanent defection, tit-for-tat, and the grim trigger strategy (permanent defection after the other player has stopped cooperation once). This observation is in line with the literature on social preferences which argues that people are conditionally cooperative. They contribute to the aggregate welfare of a group if they expect others to do so as well (Fischbacher et al., 2001; Fischbacher and Gächter, 2004; Keser and Van Winden, 2000). Evidence on Leading-by-Example suggests that sequential contribution structures generally increase cooperation. In particular, several studies show that group members who are informed about the contribution of a first-mover (or leader) make significantly higher contributions than members in groups with simultaneous contributions (Dannenberg, 2015; Eisenkopf, 2020; Güth et al., 2007; Moxnes and Van der Heijden, 2003; Pogrebna et al., 2011).

Prediction 4.1*: In the LEADING-BY-EXAMPLE Treatment, contributions to the common good are higher than in the COOPERATION Treatment.*

In the working stage of the *NO-COOPERATION Treatment*, the participants face constant returns and increasing costs over time. Any payoff-maximizing person in this treatment will leave at the end of round 10. After this round, the costs are equal or higher than the reward. While technical problems (like failing internet connections) represent a random element, the date of exit provides a measure for a person's task motivation (or, in economic terms, her opportunity costs) in this treatment. A person who dislikes the task will forego earnings and leave the experiment well before round 10. If you fancy counting zeros instead you might be ready to give up some points and work beyond that round. However, for the *COOPERATION* and the *LEADING-BY-EXAMPLE Treatments*, each cooperating member increases the reward by 6 points. Thus, the scope for cooperation implies that, on average, participants in the *LEADING-BY-EXAMPLE* and the *COOPERATION Treatments* stay longer in the experiment than those in the *NO-COOPERATION-Treatment*.

Prediction 4.2*: Participants leave the experiment in later rounds if the other group members contribute more to the common good.*

For rounds 1-10, the private costs of an additional task are lower than the private earnings, even without cooperation. Since positive payoffs are possible until round 10, conditionally cooperative individuals might stop contributions (as both tit-for-tat and the grim trigger strategy suggest), but they should not leave the experiment within rounds 1-10 if they care about their payoffs. Such a departure would not even punish the fellow group members because it imposes no additional burden on them. However, as we argued in the introduction, failed cooperation may also reduce task motivation and therefore foster attrition.[14] Thus, the exit time is an indicator for a participant's task motivation.

Prediction 4.3: *Low cooperation within groups induces premature attrition.*

In the context of our research question, the impact of Leading-by-Example on (premature) attrition is particularly interesting. Leaders significantly impact the cooperation within a group through their visible contribution (see prediction 1). Due to the generous mpcr of 6 points, even one additional cooperating member is sufficient to generate added value through cooperation. Such a high marginal per capita return usually fosters the cooperation within a group (Zelmer, 2003). However, the leader may not contribute if she assumes that not enough followers will follow her example (Blanco et al., 2014; Gächter et al., 2012; Teyssier, 2012). Such uncooperative contribution behavior on the part of the leader impedes group cooperation (Güth et al., 2007; Qiu et al., 2018). As we outline in prediction 3, impeding cooperation can alter task motivation and thus foster premature attrition.

Prediction 4.4: *Bad leadership foster premature attrition.*

4.4 Results

Before we test the predictions in detail, we have a brief look at the descriptive statistics. Table 4-2 shows the average number of played rounds and the subjects' cooperation rate for all treatments, separately for groups with four or less than four members. Appendix C.III provides descriptive statistics for each initial group size. In the *LEADING-BY-EXAMPLE Treatment* we provide details for leader and follower.

In the next subsections, we will study cooperation in the *COOPERATION* and *LEADING-BY-EXAMPLE Treatment*s first before we focus on attrition across the treatments. We then turn our attention on a causal relationship between cooperation and attrition. If not noted otherwise, all subsequent nonparametric tests are two-sided and use the averages across the played rounds within a group as an independent observation.

[14] see also Peng and Hsieh (2012); Tauer and Harackiewicz (2004); Walker (2010).

Table 4-2: Descriptive statistics

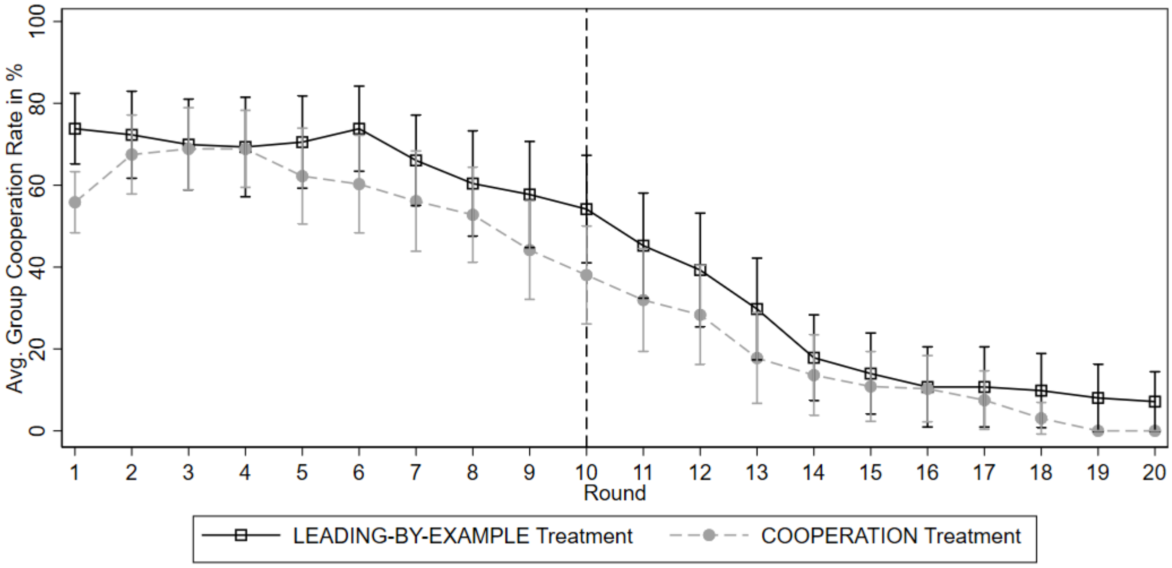
	NO-COOPERATION			COOPERATION			LEADING-BY-EXAMPLE		
Initial Group size	**Less than 4**	**4 Members**	**Both**	**Less than 4**	**4 Members**	**Both**	**Less than 4**	**4 Members**	**Both**
Subjects	37	68	105	47	52	99	32	68	100
Groups	13	17	30	17	13	30	11	17	28
Played Rounds									
Mean	10.86	9.65	10.17	10.60	12.10	11.24	11.73	12.46	12.17
	(2.92)	(1.11)	(2.14)	(3.59)	(3.67)	(3.64)	(3.51)	(4.42)	(4.04)
Leader	---	---	---	---	---	---	11.09	12.53	11.96
							(2.39)	(4.62)	(3.91)
Follower	---	---	---	---	---	---	12.05	12.43	12.28
							(4.65)	(4.42)	(4.28)
Subjects' Cooperation Rate									
All rounds	---	---	---	0.57	0.67	0.61	0.64	0.70	0.68
				(0.34)	(0.21)	(0.29)	(0.26)	(0.29)	(0.27)
Leader	---	---	---	---	---	---	0.76	0.73	0.74
							(0.22)	(0.31)	(0.28)
Follower	---	---	---	---	---	---	0.58	0.70	0.65
							(0.31)	(0.31)	(0.31)

Note: The table presents the average played rounds and the average subject's cooperation rate within a group as independent observation. Standard deviations in parentheses.

4.4.1 Analysis of cooperation

First, we focus on differences in treatments with cooperation. Figure 4-1 displays the average group cooperation rate over time. We find quite similar group cooperation rates for the *LEADING-BY-EXAMPLE Treatment* and the *COOPERATION Treatment*. The differences across all rounds are insignificant (p = .40, Mann-Whitney-U-test).

Figure 4-1: Average Group Cooperation Rate over time with 90% CIs



Note: The cooperation rate is calculated as the number of cooperation decisions within a group in a given round divided by the initial group size at the beginning of the experiment.

To disentangle treatment differences in greater detail, Table 4-3 provides a probit regression on the impact of Leading-by-Example on cooperation. Models I-III focus on all rounds, while models IV-VI consider only rounds 1-10. By design, in these early rounds' participants can generate a positive payoff even without cooperation (see section 4.2). Last, not least, model VII takes only round 1 into account The dependent variable is a subject's decision to cooperate, with standard errors clustered at the session-level. The *COOPERATION Treatment* serves as the benchmark. *Leading-by-Example* denotes the dummy variable for the other treatment. *Group Size* reflects the group size in a given round. The dummy variable *Member Exit* indicates whether at least one group member left the experiment in comparison with the previous round. *Coop Others (t-1)* denotes the relative share of how many other group members of the initial group size cooperated in the previous round.

Models I and IV study the impact of Leading-by-Example. Both models confirm our previous finding, that Leading-by-Example does not lead to higher cooperation in comparison with the simultaneous setting. Moreover, the models show a significant impact of the group size on cooperation. However, the effect vanishes in the other models. Models II and V control for the cooperation rate of the others in the previous round. Both models show that the likelihood of cooperation increases with higher cooperation rates in the previous round. Models III and VI now additional control for leaving group members. Both models indicate, that the willingness to cooperate decreases when at least one group member leaves experiment between two rounds. At this point, however, the mechanism behind such a drop out is not a clean cut. On the one hand, the incentives to cooperate decrease as soon as a group member leaves. On the other hand, however, we cannot rule out that group members stop contributing, because they just do not want to share their money with the group members who have left. To disentangle the mechanism, the attrition in round 1-10 is particularly interesting. In these rounds, participants still earn money, even without cooperation. Thus, if participants stop contributing because they do not want to share the money, they should not leave the experiment entirely. Such a departure would not punish the remaining group members or the group members who left. We focus on the premature attrition rates in section 4.4.3.

Table 4-3: Individual's decision of cooperation

Dep. Var.: Indiv. decision of cooperation	**COOPERATION and LEADING-BY-EXAMPLE Benchmark: COOPERATION**						
	All Rounds			**Rounds 1-10**			**Round 1**
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Leading-by-Example	0.17	0.05	0.06	0.12	0.02	0.02	0.47**
	(0.28)	(0.18)	(0.17)	(0.27)	(0.18)	(0.17)	(0.21)
Coop. Others (t-1)		2.1***	2.2***		2.1***	2.1***	
		(0.15)	(0.16)		(0.15)	(0.15)	
Member Exit			-0.35*			-0.4**	
			(0.19)			(0.20)	
Group Size	0.7***	0.17	0.05	0.51**	-0.00	-0.13	
	(0.14)	(0.12)	(0.15)	(0.21)	(0.17)	(0.19)	
Initial Group Size	-0.48*	-0.10	0.02	-0.34	0.04	0.17	0.15
	(0.25)	(0.14)	(0.17)	(0.31)	(0.21)	(0.24)	(0.16)
Round	0.02	-0.01	-0.01	-0.01	-0.03	-0.03	
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	
Constant	-0.39	-1.0**	-1.1**	-0.05	-0.76*	-0.78*	-0.32
	(0.70)	(0.39)	(0.39)	(0.67)	(0.43)	(0.42)	(0.52)
Observations	2,351	2,155	2,155	1,829	1,633	1,633	196
Subjects	196	194	194	196	194	194	196
Groups	58	58	58	58	58	58	58
Pseudo-R-Squared	0.080	0.291	0.294	0.031	0.240	0.243	0.03

*Probit Regression for likelihood of cooperation. Robust standard errors in parentheses. Clustered at Session-level. * p<0.1; ** p<0.05; *** p<0.01. Note: Three subjects solved the task but left the experiment during the first round without a contribution. These subjects are omitted in the models. In addition, one more subject dropped out at the end of the first round as well as at the beginning of round 2, respectively. These two subjects were omitted in models II, III, V and VI.*

Our findings do not confirm our prediction 1, but they do not imply that Leading-by-Example has no behavioral impact. Next, we focus only on the first round of the experiment (model VII). In the *COOPERATION Treatment* and the *LEADING-BY-EXAMPLE Treatment*, the first round of the experiment provides exogenous variation in cooperation as an estimator of the impact of Leading-by-Example. In later rounds, participants can update their expectations about the cooperation of the others based on the previous round. However, in the first round, expectations are based only on one's own, home-grown beliefs. Model VII presents the results. The model shows a positive and significant impact of the *LEADING-BY-EXAMPLE Treatment*. More precisely, the likelihood of cooperation increases significantly when a participant is in the *LEADING-BY-EXAMPLE Treatment* compared to the *COOPERATION Treatment*. The variable *Initial Group Size* does not affect the decision of cooperation significantly.

Our results so far show that Leading-by-Example increases cooperation at the beginning of the experiment. However, this effect vanishes in the subsequent rounds. Next, we investigate the behavioral impact of Leading-by-Example. Table 4-4 shows results from a probit regression. Again, the dependent variable is the individual decision of cooperation. The variable *Leader's coop.* indicates the cooperation decision of the leader in a given round. *Follower (Leader) Exit* takes a value of 1 if at least one follower (the leader) left the experiment in comparison with the previous round. The variable *Coop Others (t-1)* represents the average cooperation rate of the fellow group members in the previous round, while *Group Size* reflects the group size in a given round. Models I-IV focus on all rounds, while the models V-VIII consider the rounds 1-10. In Table C-III as well as Table C-IV in Appendix C.IV we provide further estimations.

Models I-III focus on the followers. The model IV considers the leader. Model I shows that the leaders strongly influence the decision of the followers, while leaving followers decreases the probability of a contribution. Model II now takes the average group contribution from the previous round into account. The variable enters highly significant. Thus, cooperation becomes more likely in cooperative groups. Moreover, model II indicates that both, leaving leaders and followers, negatively influence the decision of cooperation. Model III provides a joint model. Again, we find a strong impact of the leader's decision as well as the cooperation rate in the previous round. However, the variable *Follower Exit* turns insignificant. Model IV now focus in the leaders. The model shows that the leader's decision to cooperate strongly depends on the lagged cooperation rate.

Next, we turn to the rounds 1-10. Models V-VII focus on the followers. The models show a strong impact of the leader's decision as well as the cooperation rate in the previous round. However, for all models we do not find a significant impact of leaving members. Model VIII provides results for the leader. We find a strong impact of the others' cooperation rate in the previous round. Moreover, our model indicates that the probability of cooperation decreases with leaving followers.

Consequently, even if our findings do not provide evidence that Leading-by-Example enhance cooperation in comparison with simultaneous decisions, we find that the leaders' decisions have a strong behavioral impact on the decision of the followers.

Result 4.1: *The cooperation rate does not differ significantly between Leading-by-Example and the Cooperation Treatment. However, group members follow the example set by the leader.*

Table 4-4: Individual's decision of cooperation in the LEADING-BY-EXAMPLE Treatment

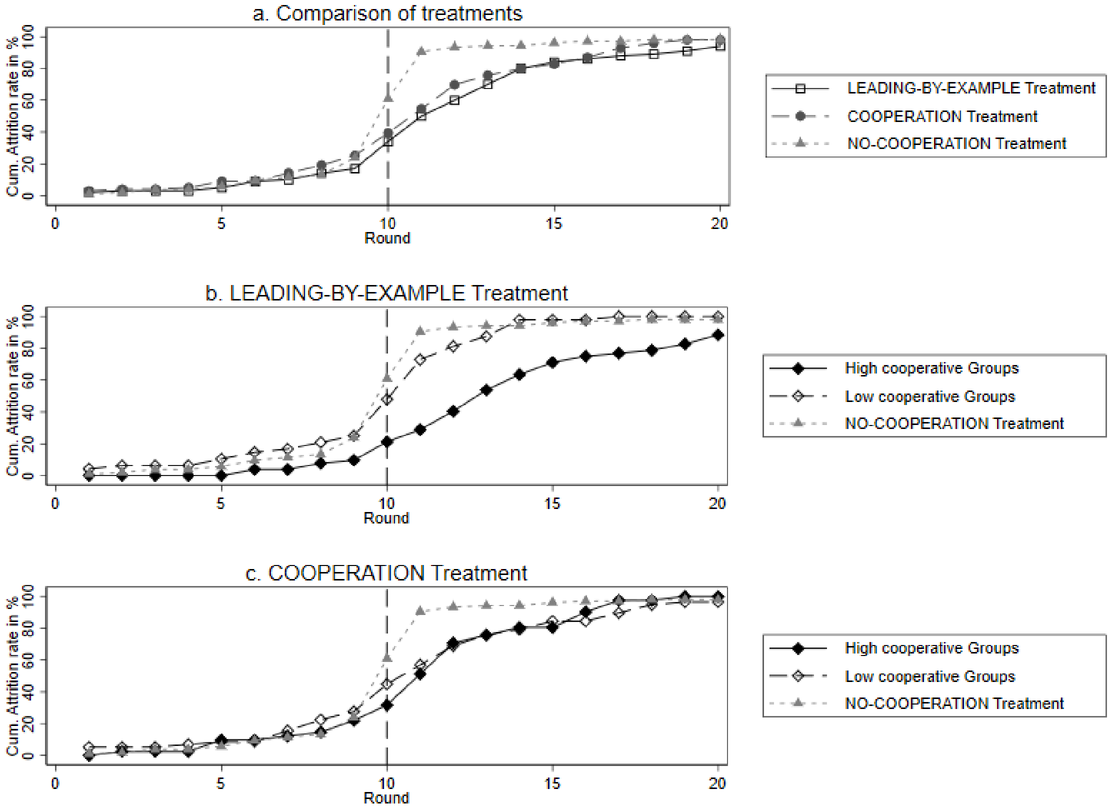
Dep. Var.: Indiv. decision of cooperation	LEADING-BY-EXAMPLE Treatment All rounds				Rounds 1-10			
	Follower			Leader	Follower			Leader
	I	II	III	IV	V	VI	VII	VIII
Leader's coop.	1.2***		0.7***		1.2***		0.7***	
	(0.241)		(0.261)		(0.252)		(0.257)	
Coop Others (t-1)		2.1***	1.7***	2.1***		2.1***	1.7***	2.1***
		(0.246)	(0.274)	(0.299)		(0.256)	(0.298)	(0.358)
Leader Exit	0.715*	-0.548			0.866	-0.377		
	(0.409)	(0.457)			(0.584)	(0.487)		
Follower Exit	0.143	-0.346	-0.271	-0.684	0.254	-0.447	-0.253	-1.51*
	(0.287)	(0.307)	(0.328)	(0.558)	(0.328)	(0.451)	(0.547)	(0.869)
Group Size	0.67**	0.149	0.165	-0.147	0.67**	0.131	0.167	-0.521
	(0.265)	(0.263)	(0.202)	(0.267)	(0.306)	(0.288)	(0.263)	(0.471)
Initial Group Size	-0.381	0.0726	0.0774	-0.021	-0.436	-0.028	0.007	0.239
	(0.360)	(0.289)	(0.254)	(0.491)	(0.400)	(0.316)	(0.310)	(0.620)
Round	0.05**	0.0213	0.0286	-0.012	0.0439	0.0155	0.0321	-0.07*
	(0.0202)	(0.0191)	(0.0203)	(0.0274)	(0.0315)	(0.0287)	(0.0318)	(0.0387)
Constant	-1.5**	-1.7***	-2.0***	0.0364	-1.281	-1.25*	-1.8**	0.773
	(0.763)	(0.527)	(0.615)	(0.980)	(0.867)	(0.657)	(0.750)	(0.922)
Observations	888	817	817	307	667	596	596	238
Subjects	71	70	70	28	71	70	70	28
Groups	28	28	28	28	28	28	28	28
Pseudo-R-Squared	0.213	0.297	0.323	0.254	0.163	0.251	0.283	0.256

*Probit Regression for likelihood of cooperation. Robust standard errors in parentheses. Clustered at Session-level. * p<0.1; ** p<0.05; *** p<0.01. Note: One subject solved the task but left the experiment during the first round without a contribution. This subject is omitted in the models. In addition, one more subject dropped out at the end of the first round. This subject is omitted in the models II, III and VI, VII.*

4.4.2 The impact of cooperation on overall attrition

Next, we focus on differences in attrition rates. Figure 4-2a shows the cumulative attrition rate for all treatments over time, while Figure 4-2b and Figure 4-2c represent the cumulative attrition rate related to the initial level of group cooperation for each cooperation treatment separately. For Figure 4-2b and Figure 4-2c we based our analyses on a median split. We considered the cooperation rate of the groups in the first three rounds. We then divided them into two groups with high or low cooperation levels. We included the attrition rates of the *NO-COOPERATION Treatment* as a benchmark for all figures. We applied the nonparametric Kaplan Meier method (Kaplan and Meier, 1958) to calculate estimated survival times. The method is used to estimate the probability that a certain event will not occur within a time interval for a test object. We also applied a log-rank test to test for differences between the populations studied in the probability of occurrence of an event at a given time point. For Figure 4-2a, we do not find differences between the *COOPERATION Treatment* and the *LEADING-BY-EXAMPLE Treatment* ($p = .22$, log-rank test), but both distributions differ significantly from the *NO-COOPERATION Treatment* ($p < .01$ for both comparisons, log-rank test). Next, we turn to the attrition rates for both cooperation treatments separately. For the *LEADING-BY-EXAMPLE Treatment* we find that the high cooperative and low cooperative groups significantly differ from each other ($p < .01$, log-rank test). Moreover, the attrition rate for the high cooperative groups differs significantly from the *NO-COOPERATION Treatment* ($p < .01$, log-rank test), while we do not find differences for the low cooperative groups ($p = .33$, log-rank test). Turning to the *COOPERATION Treatment*, we find that both high cooperative groups as well as low cooperative groups differ significantly from *NO-COOPERATION Treatment* ($p < .01$ for both comparisons with the *NO-COOPERATION Treatment*, log-rank test). However, we do not find differences between the attrition rates from the low and high cooperative groups ($p = 0.86$, log-rank test).

Figure 4-2: Cumulative Attrition rate



To address our previous finding in greater detail, we estimated the determinants on attrition by using the proportional hazards model (Jenkins, 1995) fitted to instances where participants dropped out. Table 4-5 presents the results. For models I and II the *NO-COOPERATION Treatment*, for the remaining models the *COOPERATION Treatment* denotes the benchmark. *Leading-by-Example* and *Cooperation Treatment* are dummy variables for the other treatments. The variable *Own Coop.* denotes the own decision of cooperation in a given round (=1). The variable *Coop. rate others* reflects the relative share of how many other group members of the initial group size cooperated in a given round. The dummy variable *Member Exit* indicates whether at least one group member left the experiment between two rounds. In Table C-V in Appendix C.IV we provide further estimations.

Models I-II consider all treatments, while the *NO-COOPERATION Treatment* serves as the benchmark. Model I focuses on treatment differences. The model shows that attrition is less likely if a participant is in one of the treatments that allow for cooperation. Model II shows a positive and highly significant impact of recently dropped members. Thus, it is more likely to drop out of the experiment if at least one group member left the experiment in the previous round. An interaction terms reveals that dropping members have no differential effect between treatments (see Table C-V in Appendix C.IV).

Table 4-5: Determinants of attrition (all rounds)

Dep. Var: Remaining (0) or Dropping out (1) after round t	**All treatments Benchmark: NO-COOPERATON**		**LEADING-BY-EXAMPLE and COOPERATION Benchmark: COOPERATION**				
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Leading-by-Example	-0.7***	-0.6***	-0.207	-0.48	0.00691	-0.0850	-0.0372
	(0.244)	(0.210)	(0.250)	(0.311)	(0.334)	(0.316)	(0.285)
Cooperation Treatment	-0.390*	-0.348*					
	(0.212)	(0.180)					
Coop. rate others					-1.7***		-1.2***
					(0.271)		(0.214)
Own Coop.						-1.0***	-0.306
						(0.262)	(0.199)
Member Exit		1.55***	1.56***	1.10***			1.17***
		(0.182)	(0.255)	(0.239)			(0.211)
Member Exit * LbE				0.86***			
				(0.307)			
Initial Group Size	0.0149	-0.0931	-0.261	-0.257	0.139	0.00767	0.00196
	(0.261)	(0.220)	(0.213)	(0.213)	(0.271)	(0.264)	(0.238)
Round	0.17***	0.16***	0.15***	0.15***	0.13***	0.14***	0.13***
	(0.0210)	(0.0191)	(0.0194)	(0.019)	(0.0238)	(0.0229)	(0.0208)
Constant	-3.6***	-3.4***	-3.0***	-2.9***	-3.3***	-3.2***	-3.1***
	(0.981)	(0.828)	(0.681)	(0.693)	(0.854)	(0.838)	(0.725)
Observations	3,417	3,417	2,354	2,354	2,351	2,351	2,351
Subjects	304	304	199	199	196	196	196
Groups	88	88	58	58	58	58	58

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: Three subjects completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. These three subjects are not included in the models V-VII.*

Models III-VII focus on both treatments that allow for cooperation. The *COOPERATION Treatment* serves as the benchmark. Models III and IV show a positive a highly significant impact of leaving members. However, the interaction term in model IV reveals that the effect is even stronger in the *LEADING-BY-EXAMPLE Treatment*. Models V examines the impact of the cooperation rate of the fellow group members. The model shows a negative and highly significant effect. Hence, attrition becomes more likely when the cooperation within a group is poorly addressed. Model VI controls for the own cooperation. The model shows a negative and significant impact of the own cooperation. However, the effect vanishes in a joint model (model VII), while the variables *Member Exit* and *Coop. Rate Others* remain highly significant.

Result 4.2:

1. *The attrition rate in treatments with scope for cooperation is significantly lower than in the NO-COOPERATION Treatment.*
2. *Successful cooperation within a group delays attrition. Leaving group members foster attrition.*

Results 4.2 suggest a relationship of cooperation on attrition. However, they do not rule out that low cooperation may also reduce task motivation. Hence, we must test whether the low cooperation of fellow group members causes premature attrition which we will do in the following subsection.

4.4.3 The impact of cooperation on premature attrition

We now focus on the impact of cooperation on premature attrition. By design, we can divide the experiment into two parts (see Figure 4-2). The first part covers rounds 1-10, in which participants can generate a positive payoff without cooperation. Thus, differences in attrition rates in these rounds indicate lower task motivation. However, at the end of round 10 the benefit of an additional task does not exceed its cost anymore for uncooperative groups and for the *NO-COOPERATION Treatment*. For this reason, we find an increased attrition rate at the end of round 10 (see Figure 4-2). Thus, to identify differences in premature attrition rates, we compare the survival time curves in round 1 - 9 across treatments.

Figure 4-2a shows the cumulative attrition rate for all treatments with in the first part of the experiment (round 1-9), while Figure 4-2b and Figure 4-2c represent the cumulative attrition rate related to the initial level of group cooperation for each cooperation treatment separately. We start with the treatment comparison (see Figure 4-2a). A log-rank tests reveal only insignificant differences in premature attrition rates ($p = .16$ for *LEADING-BY-EXAMPLE Treatment* vs. *COOPERATION Treatment*, $p = .27$ for *LEADING-BY-EXAMPLE Treatment* vs. *NO-COOPERATION Treatment*, $p = .74$ for *COOPERATION Treatment* vs. *NO-COOPERATION Treatment*, log-rank test). Thus, for the first part of the experiment, the survival curves do not differ between treatments. However, comparing the cumulative average attrition rates at the end of the first part (i.e., at the end of round 9), we find significant differences between the *LEADING-BY-EXAMPLE Treatment* in comparison with both other treatments ($p < .1$, Mann-Whitney-U-Test). Thus, although the survival curves do not differ, we find significantly lower average attrition rates in the *LEADING-BY-EXAMPLE Treatment* and the end of the first part.

Next, we turn to the *LEADING-BY-EXAMPLE Treatment* (see Figure 4-2b). First, we compare the survival curves of the high and the low cooperative groups with the *NO-COOPERATION Treatment*. A log-rank test states only insignificant differences for low cooperative groups with the *NO-COOPERATION Treatment* ($p > .76$, log-rank test), but significant differences for high cooperative groups with the *NO-COOPERATION Treatment* ($p < .05$, log-rank test). The comparison of the cumulative average attrition rates at the end of the first part confirms this finding and show significant differences between the highly cooperative groups with the *NO-COOPERATION Treatment* ($p < .05$, Mann-Whitney-U-Test), but no differences between the low cooperative groups with the *NO-COOPERATION Treatment* ($p = .78$, Mann-Whitney-U-Test). Next, we compare the survival curves of the high and low cooperative groups in the *LEADING-BY-EXAMPLE Treatment*. A log-rank test shows that the curves differ significantly from each other ($p < .05$, log-rank test). A Mann-Whitney-U-Test confirms that the cumulative average attrition rates are significantly higher in the low cooperative groups ($p < .1$, Mann-Whitney-U-Test).

Turning to the *COOPERATION Treatment* (see Figure 4-2b), we do not find any significant differences between the survival curves of the three populations (low cooperative groups, high cooperative groups, *NO-COOPERATION Treatment*) within the first part ($p = .51$ for low vs. high cooperative groups, $p = .52$ for low cooperative groups vs. *NO-COOPERATION Treatment*, $p = .85$ for high cooperative groups vs. *NO-COOPERATION Treatment*, log-rank test). A comparison of the cumulative average attrition rates at the end of round 9 also shows no significant differences ($p > .65$ for all comparisons, Mann-Whitney-U-Test).

Our previous findings show no treatment differences in premature attrition rates within the first part of the experiment, but the cooperation rate within a group seems to affect attrition rates. Next, we disentangle this finding and focus on the cooperation treatments in greater detail. Table 4-6 presents results from a proportional hazards model (Jenkins, 1995) fitted to instances where participants dropped out (in Table C-VI in Appendix C.IV we provide further estimations). Models I and II consider both treatments that allow for cooperation. The *COOPERATION Treatment* denotes the benchmark. However, for the first part of the experiment, we do not find significant differences between the treatments. Model II controls for the cooperation rate of the fellow group members. The coefficient of the variable enters negative and significant. Thus, attrition becomes more likely with low cooperation within a group.

Table 4-6: Determinants of attrition (round 1-9)

Dep. Var: Remaining (0) or Dropping out (1) after round t	COOPRATION and LEADING-BY-EXAMPLE						
	Benchmark: COOPERATION		Only LEADING-BY-EXAMPLE			Only COOPERATION	
	I	II	III	IV	V	VI	VII
Leading-by-Example	-0.335	-0.336					
	(0.483)	(0.514)					
Coop. rate others (CRO)	-0.897*	-0.90**	-1.112*	-0.873	-0.516	-0.510	0.223
	(0.478)	(0.429)	(0.579)	(0.607)	(0.680)	(0.502)	(0.719)
Own Coop.					-0.486		-1.112*
					(0.451)		(0.613)
Member Exit		-0.028	1.412*	2.04***	2.13***		
		(0.923)	(0.721)	(0.683)	(0.739)		
Member Exit * CRO				-2.926*	-3.137*		
				(1.589)	(1.723)		
Initial Group Size	-0.0217	-0.020	0.403	0.521	0.544	-0.240	-0.219
	(0.183)	(0.180)	(0.374)	(0.403)	(0.418)	(0.220)	(0.255)
Round	0.28***	0.28***	0.212**	0.226**	0.220**	0.32***	0.32***
	(0.064)	(0.065)	(0.097)	(0.103)	(0.104)	(0.090)	(0.092)
Constant	-4.6***	-4.6***	-6.1***	-6.8***	-6.8***	-4.3***	-4.2***
	(0.821)	(0.822)	(1.725)	(1.977)	(2.047)	(0.853)	(0.846)
Observations	1,672	1,672	850	850	850	822	822
Subjects	196	196	99	99	99	97	97
Groups	58	58	28	28	28	30	30

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: Three subjects (one in the Leading-by-Example and two in the Cooperation Treatment) completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. These three subjects are omitted in the models III-VII.*

The remaining models now focus on both cooperation treatments separately. The variable *Own Coop.* denotes the own decision of cooperation in a given round (=1). The variable *Coop. rate others* reflect the average cooperation rate of the fellow group members. Model III indicates a significant and negative effect of the other's cooperation rate in the *LEADING-BY-EXAMPLE Treatment*. Moreover, the model shows that it becomes more likely to dropping out of the experiment, when at least one group member left the experiment in the previous round. Thus, they sacrifice the economic benefits of subsequent rounds when group members leave the experiment. Model IV now controls for an interaction term of leaving members and the cooperation rate of the others. Again, the model indicates that leaving group members significantly foster attrition, but a high cooperation rate of the other group members significantly weakens this effect (model IV). Last, not least, model V controls for the own contribution. However, the coefficient of the variable enters insignificant.

Next, we turn to the *COOPERATION Treatment* (model VI and VII). Model VI shows only an insignificant impact of the other's cooperation rate (model VI) in the first part of the experiment. Model VII now controls for the own cooperation. The variable enters with a negative coefficient. Thus, while we find no effect of cooperation on attrition in the *COOPERATION Treatmen*t, the models show that participants are more likely to refrain from dropping out of the experiment if they themselves have contributed in a round. Note, that in rounds 1-9 we do not find a person who drops out after a group member dropped out in the previous round in the *COOPERATION Treatment*.

Result 4.3:

1. *There are no treatment differences in the survival time curves. However, at the aggregate level, we find that the attrition rate is lower in the LEADING-BY-EXAMPLE at the end of the first part.*
2. *For groups with a leader, successful group cooperation delays premature attrition. We do not find such a statistical relationship for leaderless groups.*

4.4.4 The impact of leadership on attrition

The previous observations indicate, that group cooperation has a significant impact for groups with a leader, but not for leaderless groups. Next, we disentangle the leader's impact on the premature attrition. Table 4-7 presents the results. Again, we estimated the determinants on attrition by using the proportional hazards model fitted to instances where participants dropped out. In Table C-VII in Appendix C.IV we provide further regressions.

Models I-II focus on the leaders, while the remaining models focus on the followers. For leaders, models I indicates that neither the cooperation rate of the others nor the own cooperation has a significant impact on attrition with in the rounds 1-9. Model II controls for the *Member Exit* variable. The model shows a strong impact of dropping members on the attrition of the leaders. This finding is particularly interesting in light of our observations from section 4.4.1. Leaving followers not only influence the leader's decision to cooperate, but leaders react also sensitively to the group's attrition and rather leave the experiment themselves. In addition, the variable *Own Coop.* turns significant. Thus, dropping out of the experiment becomes less likely, when the leader has cooperated in a given round.

Table 4-7: Determinants of attrition (round 1-9) in the LEADING-BY-EXAMPLE Treatment

Dep. Var: Remaining (0) or Dropping out (1) after round t	**Leader**		**Follower**				
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**
Coop. rate Others	-0.367	0.115	-1.34**	-1.27**			
	(0.844)	(1.013)	(0.543)	(0.537)			
Coop. rate Followers					-1.898*	-1.868*	-1.811*
					(1.050)	(1.066)	(1.054)
Coop. Leader					-0.187	-0.136	
					(0.492)	(0.447)	
Own Coop.	-1.228	-1.199*	-0.0151	-0.0433	0.00925	-0.0187	0.119
	(0.845)	(0.703)	(0.543)	(0.560)	(0.509)	(0.523)	(0.670)
Member Exit		2.98***		0.471		0.527	-0.263
		(0.945)		(0.678)		(0.638)	(0.960)
Leader Participates							-1.59**
							(0.811)
Initial Group Size	0.0423	-0.0428	0.618	0.625	0.871	0.874	0.897
	(0.661)	(0.633)	(0.599)	(0.595)	(0.665)	(0.664)	(0.668)
Round	0.40***	0.39***	0.161	0.156	0.167	0.163	0.120
	(0.0847)	(0.0935)	(0.132)	(0.131)	(0.138)	(0.137)	(0.147)
Constant	-5.513*	-5.90**	-6.40**	-6.44**	-7.41**	-7.46**	-5.985*
	(2.894)	(2.873)	(2.919)	(2.880)	(3.339)	(3.307)	(3.565)
Observations	243	243	607	607	607	607	607
Subjects	28	28	71	71	71	71	71
Groups	28	28	28	28	28	28	28

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: One subject completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. This subject is omitted in the models.*

Turning to the followers, models III and IV show a negative impact of the cooperation rate of the others on attrition, while the own cooperation does not have a significant impact. Moreover, model VI takes leaving group members into account. However, the variable enters insignificant. Models V and VI now disentangles the impact of cooperation in greater detail. More precisely, the models distinguish between the contribution rate of the other followers and the cooperation of the leader. Models V and VI show a negative and significant impact of the cooperation rate of the followers, while the coefficient for the leader’s cooperation is insignificant. Model VI controls for leaving members, but the variable enters insignificant. Consequently, our models suggest that the leader's behavior has less impact on the attrition rates in the first part of the experiment. Model VII now controls for the presence of the leader. More precisely, the variable *Leader Participates* indicates whether the leader still participates (=1) or already left the experiment (=0). The variable enters significant and negative. Thus, attrition of the followers becomes more likely when the leader left the experiment. In Table C-VIII in

Appendix C.IV we provide further estimations that show, that the leader has at least some impact on attrition of the follower. However, these estimations are not robust across model specifications.

Our results suggest that the leader does not affect the attrition rate of participants. However, our experimental design allows for further testing to examine the relationship between leader and follower. In the *LEADING-BY-EXAMPLE Treatment*, the leader's initial choice is exogenous. Therefore, the leader's initial choice for the *LEADING-BY-EXAMPLE Treatment* provides exogenous variation in cooperation as estimators of subsequent cooperation and attrition. Table 4-8 provide the results. We use different dependent variables for the models. Model I looks at the causal impact on a participant's own cooperation rate in rounds 1-9. For model II, we created a dummy variable as dependent variable indicating whether a person left the experiment within the first nine rounds. Finally, for models III, we estimated in which round a participant dropped out. We cut off this new variable after round 9 and assigned a value of 10 to all people who left after that round.

Our models show that the leader's contribution decision in round 1 does not affect the follower's willingness to cooperate, their dropout probability, or the exit round. Thus, while our results suggest a dynamic relationship between cooperation rate and attrition, our models in Table 4-8 fail to demonstrate causality related to leaders.

Result 4.4: *The leader's contribution behavior does not affect the attrition rate of the followers.*

Table 4-8: Exogenous variation in cooperation as estimators of subsequent cooperation and attrition

	LEADING-BY-EXAMPLE Treatment Followers only		
Dep. Var.	**Own Coop. Rate Round 1-9**	**Dropout in rounds 1-9 (dummy)**	**Exit round (censored at 9)**
	I	**II**	**III**
	OLS	Probit	OLS
Input Leader in t = 1	0.120	0.389	0.0763
	(0.725)	(0.522)	(0.178)
Initial Group Size	0.245	0.197	0.0387
	(0.551)	(0.339)	(0.0942)
Constant	8.268***	-2.040	0.464
	(2.023)	(1.245)	(0.287)
Observations	72	72	71
Subjects	72	72	71
Groups	28	28	28
(Pseudo-)R-Squared	0.001	0.02	0.003

*Robust standard errors in parentheses. Clustered at Session-level. * p<0.1; ** p<0.05; *** p<0.01. Note: One subject completed the task but left the experiment during the first round without a contribution. This subject is omitted in the model III.*

4.5 Conclusion

Our study investigates the relationship between leadership, intragroup cooperation, and attrition rates in an interactive online experiment. We had a particular interest in whether cooperation of others may also alter the task motivation itself, and thus foster premature attrition. Therefore, we focused on early rounds in which participants can earn money even without cooperation. Our identification strategy relied on a methodological innovation. While online experiments often have problems with internal validity due to attrition rates (Hooghe et al., 2010; Horton et al., 2011; Zhou and Fishbach, 2016), our specific design permits us to take advantage of this methodological problem. The cage like environment in labs typically eliminates premature attrition. However, dropouts in online experiments are relatively simple if you dislike the task or have more attractive alternative business at hand. To counteract this problem, participants typically do not receive any payment or other sanctions are applied if the experiment was prematurely terminated. In our design, we eliminated any negative consequences of premature dropouts. Thus, we can interpret the attrition rate as a (noisy) signal of task motivation.

Our results show that, in general, successful cooperation decreases attrition while leaving group members foster both uncooperative behavior and attrition. Moreover, we do not find evidence that successful cooperation decreases premature attrition for leaderless groups. Specifically, we find that there is little difference in attrition rates between cooperative and uncooperative groups in the first part. In both cooperative and uncooperative groups, participants forgo the economic benefit of an additional task and prefer to leave the experiment, even if positive payoffs are still possible. However, for groups with a leader, such behavior occurs only for uncooperative groups. For cooperative groups, on the other hand, we find a positive impact of cooperation on premature attrition, but we cannot attribute this effect from the leader.

Note however, that our experimental design may have favored this outcome. The marginal per capita benefit from cooperation $\left(\frac{6}{10}\right)$ is rather generous. Even in groups with two players, a contribution increases aggregate welfare. Such a high marginal per capita return usually fosters the cooperation within a group (Zelmer, 2003). Indeed, we also find a very high willingness to cooperate in our experiment, regardless of treatment. Moreover, our experimental participants were undergraduates, who have the benefits of free time and lower opportunity costs (Feltovich, 2011; Friedman and Sunder, 1994; Weimann and Brosig-Koch, 2019). Most of them are familiar with procedure and have prior experience with the monetary incentives of lab experiments.

Chapter 5: How Groups Shape Competitive Behavior

5.1 Introduction

People make astonishing sacrifices in intergroup contests. Soldiers sacrifice their lives for their country, athletes risk their health for their team, and employees work overtime to enhance their company's competitive advantage. Even in rather sterile lab experiments, contests between groups induce seemingly excessive expenditure even in comparison to contests between individuals (Abbink et al., 2010; Bornstein et al., 2008; Charness et al., 2007; Chen and Li, 2009; Sheremeta, 2018a). Explanations like joint payoff maximization (Leibbrandt and Sääksvuori, 2012) or parochial altruism (Abbink et al., 2012) suggest that the formation of groups itself promote such competitive behavior by encouraging in-group favoritism and out-group hostility.

In this paper, we test a behavioral explanation on how group formation changes competitive behavior. Our experimental treatments bridge the gap between contests of individuals and contests between groups by modifying prize sharing rules step-by-step. In a typical experimental contest between individuals, only the winner gets the prize, and all other contestants leave empty handed. In standard intergroup contests, the prize money is distributed among the members of the winning group. In many cases - and in most experimental intergroup contests (Abbink et al., 2010; Eisenkopf, 2014, 2018, 2020; Leibbrandt and Sääksvuori, 2012; Sheremeta, 2013) - all members get equal shares irrespective of their contributions, while meritocratic considerations and incentive concerns would apportion higher shares for people who provided more resources to victory. An equal distribution of prizes among group members implies that any contest expenditure generates a positive externality for the other group members. Hence, the formation of groups with egalitarian prize sharing should lead to less expenditure among risk neutral contestants without social preferences. On the other hand, the formation of groups itself seems to induce parochial altruism (Abbink et al., 2012; Bernhard et al., 2006b; Choi and Bowles, 2007) that combines hostility towards the opponents with a consideration for the benefits of fellow group members. However, there is little evidence on how groups actually elicit such a preference. Our paper addresses this concern.

To answer our research question, our design considers two benchmark treatments from the literature which we compare with two additional intermediate treatments. Our first benchmark treatment provides a standard Tullock contest (Tullock, 1980) between six individuals for an exogenously determined prize. Each participant chooses a level of expenditure, with each point invested increases the probability of winning the prize. Our second benchmark treatment provides a standard Tullock contest between groups. Two groups of three participants each compete against each other for the prize. For each group the probability of success increases in the cumulated expenditures of all three group members. Among the members of the winning group, the prize is distributed equally. These two benchmark treatments are arguably the most commonly used forms in the experimental literature (see Sheremeta (2013, 2018a) for an overview). Hence, we use these two treatments as cornerstones in our study.

We bridge the gap between these two benchmark treatments with two additional treatments that introduce groups and vary the prize sharing rules within the winning group. In each of these two additional treatments, the contest is framed as competition between two groups. Both groups compete against each other to win a prize. However, in the first of these two treatments, the prize is randomly assigned to only one member of the winning group, while the other two members leave empty-handed. More specifically, for a given amount of expenditure, the probability of winning the prize does not differ from the benchmark treatment without groups. The difference between these two treatments boils down to a difference in framing. Evidence about the minimal group paradigm (Bernhard et al., 2006a; Diehl, 1990; Efferson et al., 2008; Tajfel, 1970; Tajfel et al., 1979) suggests that group formation makes participants less competitive towards the members of their own group (Bornstein et al., 2008; Charness et al., 2007; Chen and Li, 2009). Thus, a comparison with the first benchmark treatment measures the impact of the simplest form of group formation.[15]

In the second additional treatment, the prize is distributed according to individual contribution among the members of the winning group. Thus, higher expenditures lead to higher rewards. While expected payoffs do not differ from the contest without groups, the proportional distribution ensures a rather meritocratic distribution of the prize among the members of the winning group. The sharing rule also generates an insurance effect for these group members because the actual payoffs do not differ from the expected payoffs in case of success. The literature on intergroup contests show that competition between groups increases effort when facing a common external opponent (Bornstein et al., 2008; Charness et al., 2007; Chen and Li, 2009). Abbink et al. (2010) support these findings and show in their experiment that contest expenditures exceed standard Nash predictions more strongly in intergroup contests than in contests between individuals. However, they only study contests between two parties (either

[15] Note that we deliberately did not include any particular psychological mechanism to enhance group coherence (such as communication or a highlighting of a similarity in preferences, tastes or identity). See next section.

individuals or groups) which implies a variation in the number of contestants and the aggregate value of the prizes across treatments. Our study does not vary these features across treatments.

In summary, our experiment bridges the gap between contests among individuals and standard intergroup contests in three steps: first, the framing effect of groups in contests with only one prize recipient, second, the transition from indivisible prizes to proportional prize sharing among the winners, and third, the transition from proportional to equal shares of the prize. The first and the second step are perhaps the most interesting because the consideration of different social preferences changes the direction of theoretical predictions considerably. As we show in section 5.3, the standard Nash predictions do not differ between the relevant treatments if we assume common knowledge about selfish and rational participants. However, considering additional social preferences, in particular inequality aversion (Fehr and Schmidt, 1999) and parochial altruism (Abbink et al., 2012; Bernhard et al., 2006b; Choi and Bowles, 2007), leads to contradicting behavioral predictions. A feature of our experimental design is to consider and test these competing and conflicting hypotheses within a fairly simple design.

The shift from proportional to equal sharing is more predictable because the latter division rule implies a positive externality of contest expenditure. Some studies on duopoly games show that the rules for profit distribution among members of the winning team indeed are crucial for the level of competition. More specifically, they find that prices are sustained at a higher level in competition between proportionally paid teams than in competition between equally paid teams (Bornstein et al., 2008; Bornstein and Gneezy, 2002; Kurschilgen et al., 2017). Closer to our experimental design, Kugler et al. (2009) as well Gunnthorsdottir and Rapoport (2006) studied the effect of different prize sharing rules on competitive behavior in a Tullock contest. They examine an intragroup conflict in a social dilemma embedded in an intergroup competition for an exogenously determined prize. In their design, the group's probability of wining the prize depends on how contributions to the provision of a public good by its members compare to the contributions by members of the competing group. Their results show that a proportional sharing rule elicits higher individual effort than the equal sharing rule.

At the aggregate level, our results replicate the findings from Abbink et al. (2010) and related studies. Relative to the respective standard Nash predictions, we observe excessive contest expenditure in all treatments and in particular in the standard intergroup contests. However, we do not find direct evidence for outgroup hostility or unconditional in-group favoritism, as equal sharing induces free-riding while the mere introduction of groups has no significant impact. Instead, we observe some evidence for conditional favoritism as participants in all treatments with groups increase their expenditure in the investments of the fellow group members. Furthermore, the particularly strong impact of proportional prize sharing on contest expenditure supports the notion that inequality aversion fosters intergroup conflicts. Note also, that our results do not rule out that in-group favoritism may depend on a seemingly fair distribution of the prize among the members of the winning group.

The remainder of the paper is organized as follows. Section 5.2 explains the experimental design and the procedures in greater detail. In section 5.3, we present the theoretical predictions. Section 5.4 presents our main results, while section 5.5 concludes.

5.2 The experimental design and procedures

5.2.1 The design of the experiment and its treatments

We randomly allocated participants into matching groups of six. The participants interacted in the same matching groups until the end of the experiment. The experiment last for 10 rounds. In each round, the participants competed against each other in a standard Tullock contest (Tullock, 1980). At the beginning of a round, each group member received an endowment of 100 points[16] and could invest these as an expenditure to win a prize of 300 points. Each point invested, increased the probability of winning the prize. More precisely, the winning probability of a participant was determined by her own expenditure relative to total expenditures of all six matching group members. Any point not invested went into to the participants' private account. Then, the computer randomly determined the winner based on the probability of winning. If no group member expended any points, the probability of winning was the same for each group member. At the end of a round, the participants learned about the expenditures and the payoffs of each matching group member. Such a Tullock contest is arguably the most widely used model for the experimental study of contests (Abbink et al., 2010; Dechenaux et al., 2015; Eisenkopf, 2020).

We observed expenditures in four treatments. The *NO-GROUPS Treatment* and *EQUAL-SHARE Treatment* provided our two benchmarks. The *NO-GROUPS Treatment* worked as explained above. For the *EQUAL-SHARE Treatment* we subdivided the six

[16] We exchanged 100 points into 1€, about 1.22 US-$ in late 2020.

participants into groups of three players each. The group comparison did not change between rounds. Both groups competed against each other to win the exogenously determined prize. Again, we used a Tullock contest to determine the winning group. The winning probability of the group was determined by its members' expenditures relative to total expenditure in both groups. All members of the winning group received an equal share of the prize, independent of their expenditures.

The *PROPORTIONAL-SHARE Treatment* and *NO-SHARE Treatment* provided our additional treatments. Both treatments started like the *EQUAL-SHARE Treatment*, but had different prize sharing rules. In the *NO-SHARE Treatment* only one person within the winning group received the prize of 300 points. All other group members were left empty-handed. Within the winning group, each member's winning probability was determined by the own expenditure relative to total expenditure of all three group members.

In the *PROPORTIONAL-SHARE Treatment,* each member of the winning group received a share of the prize. The share here depended on the individual chosen expenses. More precisely, the individual share of the prize was determined by her own expenditure divided by the total group expenditure.

It is noteworthy that none of the treatments includes any particular psychological mechanism that fosters group coherence (such as communication or a highlighting of a similarity in preferences, tastes or identity). While such mechanisms can play a vital role in fostering cooperation among group members, we abstained from deploying them because neither of our two benchmark treatments from the literature deploys them. Hence, such mechanisms cannot help explaining the behavioral differences between those treatments.

Belief elicitation

We are further interested in the initial effects of contest structures on participants behavior. Therefore, we elicited the participant's belief about the upcoming expenditures in round 1 after choosing the own expenditure, but before the result stage. More precisely, we asked the participants to estimate how many points the other five matching group members would expend on average. We paid participants 100 points if they estimated the average expenditure exactly (+/- 10%). Table 5-1 summarizes the design and reports the number of subjects assigned to each treatment.

Table 5-1: Experimental Design

			NO-GROUPS	NO-SHARE	PROPORTIONAL-SHARE	EQUAL-SHARE
Before Round 1		**Step 1** The matching-group is subdivided into two groups	---	✓	✓	✓
Repeated for 10 rounds	**Contest**	**Step 2** Participants choose their expenses	✓	✓	✓	✓
		Step 3 Belief elicitation (only round 1)	✓	✓	✓	✓
		Step 4 Participants learn about the expenses and the winning probabilities	✓	✓	✓	✓
		Step 5 Determination of the winning group	---	✓	✓	✓
		Step 6 Determination of the winning member	✓	✓	---	---
	Prize distribution	**Step 7** a) One group member receives the prize	✓	✓	---	---
		b) All group members receive a different share of the prize, depending on their expenses	---	---	✓	---
		c) All group members receive the same share of the prize, independent of their expenditure	---	---	---	✓
Subjects			96	102	84	84
Groups			16	17	14	14

5.2.2 Procedure

We conducted the study as an online-experiment during the covid-19 pandemic. The experiment was programmed using z-Tree (Fischbacher, 2007) and conducted with the help of z-tree unleashed (Duch et al., 2020). We contacted participants from the Laboratory for Economic Experiments Vechta (LEEV, University of Vechta) and the Laboratory for Economics Research (LAER, University of Osnabrueck). Recruitment was done via hroot (Bock et al., 2014) for the participants at the University of Vechta and ORSEE (Greiner, 2015) for the participants at the University of Osnabrueck. Each participant could sign up for only one session, each consisting of 30 participants. At their registration, the participants were instructed to use either a laptop or computer during the experiment in order to ensure a minimum resolution of the devices.[17] Furthermore, we standardized the virtual screens to a resolution of 1024x768. Participants were also informed to receive an e-mail about 10-15 minutes before the experiment. This e-mail from the experimenter contained a personalized internet-link and a code to access the experiment. The personalized links were generated by ztree-unleashed. Each link represented a zleaf-client.[18] The experimenter started the experimental program at the specified time. We displayed a countdown on the screen 5 minutes before the experiment, to prepare participants. At the start of the experiment, the participants entered their private code on the first page. If no code was entered within a time restriction of 5 minutes, the program ejected the zleaf-client and thus the respective participant from the experiment. Then, the computer randomly assigned the present participants to groups and treatments. This procedure was chosen to counteract pre-start dropouts and ensured full groups at the beginning of the experiment. Present participants in groups that were not full were not allowed to participate in the experiment. These individuals received a show-up fee as compensation.

The experimental instructions were provided on screen. The participants read the instructions and answered questions about their gender and study program, the device they used, their experience with experimental studies and answered some control questions. Then the first round started. At the end of the experiment, the university paid the participants by online bank transfer (on average 14.11€ or 17.21 US-$). To ensure the strict separation of experimental and payment data, we collected the payment data via an external questionnaire after the experiment.

---

[17] We checked this with a control question at the beginning of the experiment. All but two participants complied with this requirement. Both participants stayed in the experiment until the end.

[18] The zleaf-client allows the subjects to communicate with the experimental server. It thus represents the access points to the experiment. For more details see Fischbacher (2007).

We informed participants about these procedures prior to registration in a session. We linked the payoff and the bank-account details via the participant code.

In online experiments some participants can endogenously withdraw from the experiment. While studies show that results of online-experiments are similar to lab experiments (Arechar et al., 2018; Berinsky et al., 2012; Horton et al., 2011), tedious work task and resulting high attrition rates induces a selection bias that could jeopardize the internal validity of such experiments (Horton et al., 2011; Zhou and Fishbach, 2016). To minimize premature attrition, we chose a student sample that we recruited from existing economics lab participant pools. In addition to the well-known benefits of student samples, such as lower opportunity costs or benefits of free time (Feltovich, 2011; Friedman and Sunder, 1994; Weimann and Brosig-Koch, 2019), many of our participants had previous experience in lab experiments and were familiar with the procedures.

We conducted eighteen sessions in December 2020 (LEEV 10 sessions, LAER 8 sessions). In each session, each treatment was conducted at least once. On average, a session lasted approximately 45 minutes. In total, 540 students agreed to participate in our experimental sessions. Out of this sample, 474 showed up for the experiment, while 426 could participate.[19] Thus, we gathered data from 71 matching groups of six participants. However, in 10 matching-groups we observe dropouts after the instructions. We do not consider these groups in our analysis. Thus, our subsequent data analysis relies on observations from 366 participants in 61 groups (see Appendix D.II for a detailed overview).

5.3 Theoretical analysis

Our experimental design bridges the differences between individual and group contest in three rather simple steps. This approach allows for a set of competing and contradicting hypotheses. We start with the calculation of the standard Nash equilibria before we explain the impact of additional social preferences, in particular inequality aversion (as in Fehr and Schmidt, 1999), and parochial altruism (as in Abbink et al., 2012).

With common knowledge about uniform homo economicus preferences and risk neutrality, all players only care about their own payoff. We start our analysis with the *NO-GROUPS Treatment*. In that Tullock contest, n > 2 individuals, indexed by $i's(i \in \{1, \dots, n\})$, with an

[19] Only participants in full groups were allowed to participate (see above).

initial per-capita endowment E compete with each other for a prize $P > 0$. Let x_i denote the chosen contest expenditure and we obtain the following payoff function.

$$\pi_i = E + \frac{x_i}{\sum_{i=1}^{n} x_i} P - x_i \tag{E5.1}$$

In the *NO-SHARE Treatment* and *PROPORTIONAL-SHARE Treatment*, we divide the n players in two groups of equal size ($m = n/2$). The payoff function transforms into

$$\pi_i = E + \frac{x_i}{\sum_{i=1}^{m} x_i} * \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} P - x_i \tag{E5.2}$$

The term $\frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i}$ denotes the probability that the own group wins the prize. In the *NO-SHARE Treatment*, $\frac{x_i}{\sum_{i=1}^{m} x_i}$ indicates a player's probability of winning the entire prize, if the own group has been the winner. In the *PROPORTIONAL-SHARE Treatment*, that term indicates that player's share of the prize. Of course, equation E5.2 implies the same expected payoff for a given contest expenditure as in the *NO-GROUPS Treatment* (E5.1). Thus, all three treatments lead to the same equilibrium expenditure level of

$$x_i = \frac{n-1}{n^2} P \tag{E5.3}$$

The *EQUAL-SHARE Treatment* has a different sharing rule. Each member of the winning group gets an equal share of $\frac{P}{m}$. Hence, the payoff function turns into:

$$\pi_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} * \frac{P}{m} - x_i \tag{E5.4}$$

The winning group's members share the prize irrespective of their individual contributions which implies as equilibrium expenditure:

$$x_i = \frac{P}{n^2} \tag{E5.5}$$

Because of this externality the equilibrium expenditure is lower than in the other three treatments.

Prediction 5.1 (Homo economicus preferences and risk neutrality)

The expected expenditure levels across the treatments are ordered as follows:

NO-GROUP = NO-SHARE = PROPORTIONAL-SHARE > EQUAL-SHARE

The consideration of the aforementioned social preferences changes this prediction in different directions. First, we consider how disadvantageous inequality aversion (as in Fehr and Schmidt, 1999) influences behavior. Appendix D.III provides a formal theoretical analysis of the expected change in behavior. In case of the *NO-GROUP Treatment* and *NO-SHARE Treatment*, disadvantageous inequality aversion transforms the payoff equation (E5.1) into the following utility function:

$$\begin{aligned} U_i = E + \frac{x_i}{\sum_{i=1}^{n} x_i} P - x_i \\ - \left(1 - \frac{x_i}{\sum_{i=1}^{n} x_i}\right) \Bigg[\frac{\alpha}{n-1} \left(P - \mathrm{x}_{\mathrm{j \neq i}} + x_i\right) \\ + \frac{\alpha}{n-1} \sum_{\mathrm{j \neq i}=1}^{n-2} max\left(0; x_i - \mathrm{x}_{\mathrm{j \neq i}}\right) \Bigg] \end{aligned} \qquad \text{(E5.6)}$$

We assume common knowledge that all participants have a uniform degree $0 \leq \alpha \leq 1$ of inequality aversion towards the fellow contestants. They resent any higher payoffs of fellow contestants as the utility function in equation E5.6 shows. With an increasing number of competitors that additional marginal psychological benefit of extra expenditure (the reduced chance to be worse-off than the competitor) decreases while the extra marginal costs do not change. In multi-player contests ($n \geq 3$) this additional preference discourages expenditure relative to the standard Nash equilibrium.[20] Prize sharing within the winning group, however, has an additional impact on expenditure. More precisely, prize-sharing increases expenditure relative to contests with indivisible prizes because, in equilibrium, members of the winning group cannot be worse off than their opponents (see Appendix D.III.I) for formal theoretical analysis). Hence, the marginal psychological cost of contest expenditure (i.e. the disutility from being a bit more worse off than the others) applies only to the fellow group members if the own group. We show

[20] Note that in two player contests, this inequality aversion increases expenditure incentives (e.g. Herrmann and Orzen (2008)).

that this reduction in costs does not compensate for the adverse impact of egalitarian prize-sharing in the *EQUAL-SHARE Treatment* (see Appendix D.III.I).

Prediction 5.2 (Inequality aversion and risk neutrality)

With disadvantageous inequality aversion, contest expenditure without groups (i.e., x_i $(NO - GROUP)$) is lower than in the standard Nash equilibrium. The changes between the treatments have the following impact on the relative size of the contest expenditure:

PROPORTIONAL-SHARE >NO-GROUP = NO-SHARE > EQUAL-SHARE

We now turn to parochial altruism as a different type of social preference. Hostility towards the opponents is a key feature of parochial altruism (Abbink et al., 2012; Bernhard et al., 2006b; Choi and Bowles, 2007). It implies that a contestant's utility decreases in the payoffs of the opponents. We assume common knowledge that all participants have a uniform degree $0 \leq \lambda \leq 1$ of hostility towards the fellow contestants if they are not in the same group. More specifically, λ is a multiplier of the disutility which increases in the payoffs of the opponents. As we show in Appendix D.III.II, this parochial altruism increases expenditure in the *NO-GROUP* treatment relative to the standard Nash-equilibrium.

Parochial altruism also implies a more favorable attitude towards the members of the own group. In this context, we use the parameter$(-1) \leq \rho \leq \lambda$. At $0 < \rho < \lambda$ participants are less hostile towards their fellow group members than to the other members of the other group. In case of $-1 \leq \rho < 0$ they reveal actual in-group favoritism. Hence, we obtain the following utility function for the participants in the *NO-SHARE Treatment* and *PROPORTIONAL-SHARE Treatment*.

$$U_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} \frac{x_i}{\sum_{i=1}^{m} x_i} P - x_i - \frac{\rho}{n-1}\left(\frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} \frac{\sum_{j \neq i=1}^{m-1} x_j}{\sum_{i=1}^{m} x_i} P - \sum_{j \neq i=1}^{m-1} x_j\right) - \frac{\lambda}{n-1}\left(\frac{\sum_{m+1}^{n} x_i}{\sum_{i=1}^{n} x_i} P - \sum_{m+1}^{n} x_j\right) \quad \text{(E5.7)}$$

Unlike inequality aversion, parochial altruism implies that the formation of groups generates a reduction in contest expenditure. Contestants mind less, if their fellow group members win the prize (as in Leibbrandt and Sääksvuori, 2012 or Balliet et al., 2014). Note also that context expenditure in the *NO-SHARE Treatment* and *PROPORTIONAL-SHARE Treatment* is identical and lower than in the *NO-GROUP Treatment* because parochial altruism focuses on the expected payoffs of the others. Again, expenditure is higher than in the *EQUAL-SHARE Treatment* (see Appendix D.III.II for formal theoretical analysis).

Prediction 5.3 (Parochial altruism and risk neutrality)

With parochial altruism, contest expenditure without groups (i.e., x_i $(NO-GROUP)$) is higher than in the standard Nash equilibrium. The changes between the treatments have the following impact on the relative size of the contest expenditure:

NO-GROUP > NO-SHARE = PROPORTIONAL-SHARE > EQUAL-SHARE

5.4 Results

First, we have a brief look at the descriptive statistics. Table 5-2 shows the average per capita expenditure as well as the elicited average expenditure belief in round 1 for each treatment with standard deviation in parentheses. If not noted otherwise, all subsequent nonparametric tests use average expenditure of the relevant members in a matching group across all 10 rounds as an independent observation. All reported test results are two-sided.

Table 5-2: Descriptive statistics

	NO-GROUPS	NO-SHARE	PROPORTIONAL-SHARE	EQUAL-SHARE
Standard Nash equilibrium	41.67	41.67	41.67	8.33
Mean expenditure	51.22 (6.39)	47.55 (11.76)	60.64 (9.29)	37.22 (10.16)
Belief (before round 1)	49.44 (7.33)	45.92 (6.04)	51.63 (5.43)	43.78 (9.75)
Subjects	96	102	84	84
Matching groups	16	17	14	14

Note: Standard deviations in parentheses

The table reveals that the average per capita expenditure in each treatment is above the respective standard Nash-equilibrium. These deviations from the theoretical prediction are statistically significant ($p < .01$ for *NO-GROUPS Treatment* and *PROPORTIONAL-SHARE*

Treatment, $p < .1$ for the *NO-SHARE Treatment*, $p < .05$ for *EQUAL-SHARE Treatment*, two-sided Fisher one-sample permutation test).

> ***Result 5.1:*** *Contest expenditure exceeds the standard Nash-equilibrium in all treatments.*

Regarding treatment comparisons, we find the highest per capita expenditure in *PROPORTIONAL-SHARE Treatment* and the lowest in *EQUAL-SHARE Treatment*. The differences are significant for all comparisons with the other treatments ($p < .01$ for all comparisons with *PROPORTIONAL-SHARE Treatment* and for *NO-GROUPS Treatment vs. EQUAL-SHARE Treatment*; $p < .05$ for *NO-SHARE Treatment* vs. *EQUAL-SHARE Treatment*, Mann-Whitney-U-Test). However, we only find insignificant differences regarding *NO-GROUPS Treatment* vs. *NO-SHARE Treatment* ($p = .41$, Mann-Whitney-U-Test).

> ***Result 5.2:*** *There is no significant difference in contest expenditure between the NO-SHARE and the NO-GROUPS Treatments.*

> ***Result 5.3:*** *Average contest expenditure is higher in the PROPORTIONAL-SHARE Treatment than in the NO-SHARE Treatment.*

> ***Result 5.4:*** *Average contest expenditure is higher in the PROPORTIONAL-SHARE Treatment than in the EQUAL-SHARE Treatment.*

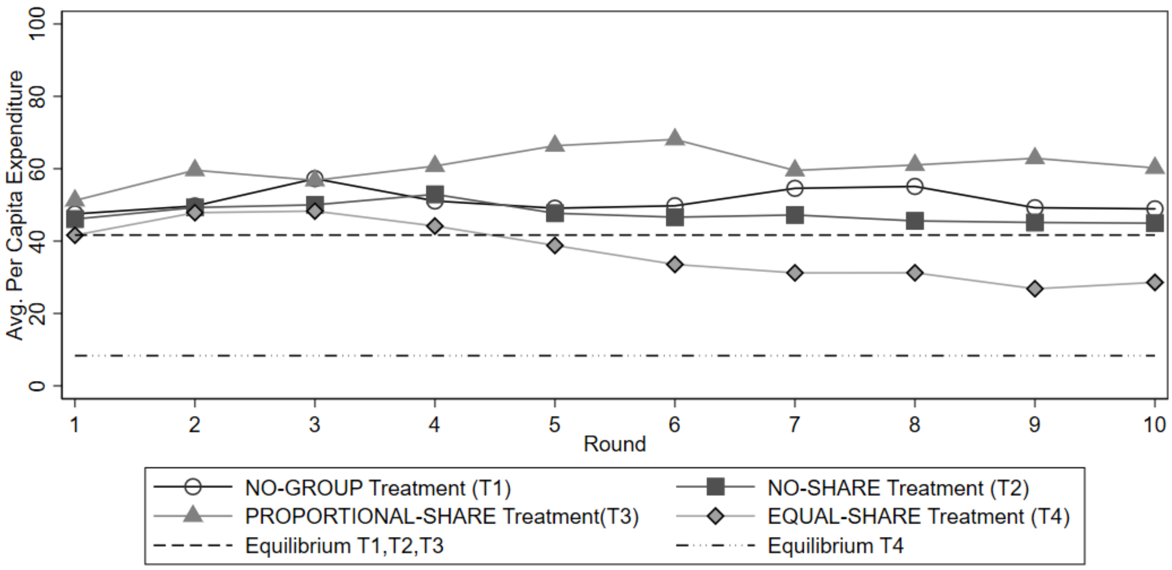
Result 5.3 in particular shows that the prize-sharing in intergroup contests has a strong impact on expenditure. We find no evidence for a particular hostility towards the opponents. Instead, the above-equilibrium expenditure suggests a general competitiveness that is not restricted to groups which is in line with the preceding literature (Mago et al., 2016; Sheremeta, 2013, 2018b). Moreover, the groups do not rule out unconditional in-group favoritism which we explore in greater detail in the following estimations in section 5.4.2 and 5.4.3 that take the intertemporal dynamics of contest behavior across the treatments into account.

5.4.1 The initial impact of the contest structure

The outcome of contests depends strongly on the expenditures of the contestants. Thus, individual expenditure levels are partly based on the expected expenditures of the others. In our experimental design, the participants learned about the lagged expenditures of the other matching group members. Hence, in later round, they can update their expectations about expenditures based on the previous round. However, the first round gives us insight into the initial impact of contest structures on expenditures, since here expectations are based only on one's own, home-grown beliefs. In this section, we consider whether different contest structures affect pre-game expectations and examine whether the design of the contest itself affects expenditures.

Table 5-2 as well as our previous results 5.2-4, indicate that the contest structure has an impact on the behavior of the participants. Figure 5-1 shows the average per capita expenditures in the respective treatment over time. Interestingly, Figure 5-1 shows that expenditures in the different contests hardly differ in the first rounds and only increase as the rounds progress.

Figure 5-1: Average per capita expenditure across the treatments



To test this finding in greater detail, Table 5-3 studies the beliefs about the average per capita expenditure of the other five matching-group members as well as the impact on the per capita expenditures in round 1. For Table 5-3 we choose two different dependent variables. Model I focuses on elicited beliefs before the start of round 1, while the dependent variable of the other three models (model II-IV) is the per capita expenditure in round 1. For all models, the standard errors are clustered at the matching group level. The *NO-GROUPS Treatment* serves as the benchmark. *No-Share*, *Proportional-Share* and *Equal-Share* are dummies for the other treatments.

Table 5-3: Beliefs and their impact on contest expenditure in round 1

	All Treatments (only round 1) Benchmark: NO-GROUP			
Dependent Variable:	**Belief in t=1**	**Per capita contest expenditure in t=1**		
	I	**II**	**III**	**IV**
No-Share	-3.520	-1.422	1.785	5.056
	(2.303)	(3.605)	(3.447)	(9.748)
Proportional-Share	2.194	3.704	1.705	4.556
	(2.288)	(3.713)	(2.951)	(7.451)
Equal-Share	-5.659*	-5.879	-0.724	17.34*
	(3.113)	(3.665)	(3.430)	(9.371)
Belief in t=1			0.911***	1.054***
			(0.0686)	(0.122)
No-Share*Belief in t=1				-0.0603
				(0.183)
Proportional-Share*Belief in t=1				-0.0613
				(0.154)
Equal-Share*Belief in t=1				-0.394**
				(0.190)
Constant	49.437***	47.51***	2.470	-4.613
	(1.797)	(2.699)	(3.880)	(6.061)
Observations	366	366	366	366
Subjects	366	366	366	366
Matching Groups	61	61	61	61
R-squared	0.022	0.011	0.342	0.353

Note: Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

Model I focuses on the elicited belief before round 1. The model shows that there are no differences in beliefs between treatments, except for the *EQUAL-SHARE Treatment*, where beliefs are slightly but significantly lower than in our benchmark treatment. Moreover, for each treatment, the beliefs exceed the predicted standard Nash-Equilibrium (see also Table 5-2). The remaining models focus on the contest expenditure in round 1. Model II shows no general treatment differences in the per capita contest expenditure in round 1. Moreover, the chosen expenditures are almost identical to the elicited beliefs (see model I). Model III controls for the beliefs and confirms the previous observation. The variable enters positive and highly significant. Thus, model III shows that individuals match their own expenditure with their belief about the expenditures of the other matching-group members. Model IV looks at treatment differences in the effect of the beliefs on the expenditures. The estimation shows that individuals in the *EQUAL-SHARE Treatment*, on average, follow their own belief to a lesser extent compared to individuals in our benchmark treatment. All other treatment differences are insignificant.

Result 5.5: *The specific incentive effects of different contests take place only after a certain period of learning.*

5.4.2 Contest behavior across treatments

Next, we compare the individual behavior in the different contest settings. Table 5-4 presents results from an OLS-regression. The table analyze to what extent subjects make their current contest expenditure decisions dependent on their own decisions and those of the other participants in the receding round. The *NO-GROUPS Treatment* serves as the benchmark. The dependent variable is the per capita contest expenditure with standard errors clustered at the level of the matching group. *No-Share*, *Proportional-Share* and *Equal-Share* denote dummy variables for the other treatments. We included the lagged average expenditure of all other matching group members and the respective round as independent variables. In Table D-II in Appendix D.IV we provide further regressions.

The first model focusses on treatment differences. The model confirms our previous results 2-4. We find that the contest expenditure is significantly higher in the *PROPORTIONAL-SHARE Treatment* and significantly lower in the *EQUAL-SHARE Treatment*. Models II and III focus on the lagged expenditures of the other five matching group members, regardless of group affiliation. Model II shows that, in general, the lagged expenditures of the other matching group members do have a small but positive effect on the own expenditure. Thus, our estimation show that individuals make adjustments to their expenditures depending on the expenditures of the other matching group members. However, model III indicates significant differences between the treatments. For all three treatments with a group structure (*NO-SHARE*, *PROPORTIONAL-SHARE* and *EQUAL-SHARE Treatments*), we find a positive and significant impact of the lagged expenditure from the other matching group members, while we find a negative impact for the *NO-GROUPS Treatment*.

Result 5.6: *Without groups, participants decrease their own expenditure in the preceding expenditure of the other contestants. With groups, the effect reverses.*

Table 5-4: Contest behavior across treatments

Dep. Var. Per capita contest expenditure	**All Treatments Benchmark: NO-GROUP**		
	I	**II**	**III**
No-Share (NS)	-3.670	-3.333	-26.67**
	(3.198)	(2.965)	(10.16)
Proportional-Share (PS)	9.423***	8.572***	-16.89
	(2.875)	(2.447)	(13.46)
Equal-Share (ES)	-14.00***	-12.76***	-47.81***
	(3.067)	(3.085)	(8.621)
Avg. Others t-1 (AO)		0.161*	-0.281**
		(0.0818)	(0.140)
NS*AO			0.454**
			(0.192)
PS*AO			0.487*
			(0.256)
ES*AO			0.764***
			(0.173)
Round	-0.495*	-0.828***	-0.616**
	(0.295)	(0.280)	(0.268)
Constant	53.94***	48.30***	69.78***
	(2.156)	(4.793)	(7.558)
Observations	3,660	3,294	3,294
Subjects	366	366	366
Matching Groups	61	61	61
R-squared	0.052	0.065	0.078

Note: Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

5.4.3 The impact of group affiliations in the contest behavior

We now consider the specific group affiliations within a matching group. We exclude the *NO-GROUPS Treatment* from our subsequent analysis. Instead, the *NO-SHARE Treatment* serves as the benchmark. Table 5-5 presents the results. In Table D-III in appendix D.IV we provide further estimations. *Proportional-Share* and *Equal-Share* denote dummy variables for the other treatments. The dependent variable is the per capita contest expenditure with standard errors clustered at the level of the matching group. The dummy variable *Winner Group t-1* indicates whether an individual was in the winner group in the previous round. The variables *Avg. Others t-1, Avg. Others (Own Group)* and *Avg. Others (Other Group)* provide information about how much the other group members spent on average in the last round. The respective variables either consider the expenditures for all matching group members or only the expenditures for those from one's own group or from the other group.

Table 5-5: The impact of group affiliations on the contest behavior

Dep. Var. Per capita contest expenditure	**Only treatments with groups**			
	Benchmark: NO-SHARE			
	I	**II**	**III**	**IV**
Proportional-Share	13.98***	10.31***	10.10***	9.958***
	(3.994)	(2.723)	(2.670)	(2.639)
Equal-Share	-10.98**	-8.230**	-8.072**	-7.962**
	(4.264)	(3.322)	(3.269)	(3.229)
Winner group t-1	9.853***			7.994***
	(2.515)			(1.527)
Avg. Others t-1		0.285***		
		(0.0866)		
Avg. Others (Own group) t-1			0.279***	0.248***
			(0.0648)	(0.0621)
Avg. Others (Other group) t-1			0.0227	0.0653
			(0.0778)	(0.0710)
Round	-1.13***	-0.95***	-0.94***	-0.93***
	(0.413)	(0.321)	(0.315)	(0.312)
Constant	49.55***	39.76***	38.91***	34.33***
	(4.105)	(5.426)	(5.229)	(5.033)
Observations	2,430	2,430	2,430	2,430
Subjects	270	270	270	270
Matching Groups	45	45	45	45
R-squared	0.099	0.098	0.119	0.130

Note: Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

Model I controls for the winning group in the previous round. The positive and highly significant coefficient indicates that expenditures increase if one was a member of the winning group in the previous round. An interaction term reveals no differences between the treatments (see Table D-III Appendix D.IV). Models II and III focus on the expenditure of the other matching group members. Model II confirms the significant effect of the lagged expenditures of all matching group members. However, as model III shows, participants react differently to the expenditures from their own and from the other group. While the expenditures of the members from the other group do not lead to any significant effect, the expenditures of the own members have a positive and highly significant impact. An interaction term reveals no treatment difference in this context (see Table D-III in Appendix D.IV). The effect remains significant even in joint model (model IV). Refining result 5.6, we can state that group affiliations lead to aligning one's expenditure especially to one's own group. The previous findings reveal at least a certain kind of *conditional* in-group favoritism.

Result 5.7: *Contest expenditure within groups reveals strategic complementarities. Group members increase their investments with higher preceding investments of the fellow group members.*

5.5 Conclusion

We have tested how group formation changes competitive behavior. We provided a novel experimental design that allowed for a clean comparison between intergroup contests and competition among individuals. We compared our evidence with theoretical predictions which included either parochial altruism or inequality aversion. We find no evidence that group formation itself changes aggregate competitive behavior. Instead, group members adjust their contributions according to their fellow members' behavior. The change in prize changing rules within groups has a stronger impact which is in line with disadvantageous inequality aversion Fehr and Schmidt (1999). Our results also show that treatments with shared prizes see particularly high contributions relative to the respective Nash equilibrium. This evidence suggests that groups generate stronger contest behavior if a certain degree of fairness is established within a group.

Of course, our design represents a rather conservative test for the impact of group formation on competitive behavior. Our group members cannot use any instrument for direct coordination like leadership or communication (Cason et al., 2012; Eisenkopf, 2020). They do not even have a common identity to rally around as in the minimal group paradigm (Diehl, 1990; Fang-Fang and Bin, 2018). We did so deliberately because neither of our benchmark treatments from the literature deployed any of these devices. Hence, we do not want to rule out that enrolment in groups with additional characteristics may have an impact.

Chapter 6: It Wasn’t Me – Unfair Decision-Making and not Standing Up for it

6.1 Introduction

A fundamental characteristic of people and organizations is the desire to see themselves in a positive light (e.g., Bénabou and Tirole, 2003). Theories of 'self signaling' or 'social signaling' are used to explain why people consciously forego information in a decision-making process (Dana et al., 2007) or change their behavior when this is publicly observable (Cueva and Dessi, 2012; Tonin and Vlassopoulos, 2013). People even delegate their decision rights to third parties to not appear as selfish and unfair or to avoid punishment (Bartling and Fischbacher, 2011; Erat, 2013; Ertac and Gurdal, 2013; Hamman et al., 2010). However, many decisions are characterized by the fact that responsibility cannot simply be delegated. Yet even in such decisions, decision-makers seem to disguise their responsibility in order to maintain a positive image. They even seem to partially twist the truth to legitimize unfair behavior. In 2017, for example, the airline Lufthansa came under heavy criticism after it raised prices for domestic flights by up to 30 percent. The airline was accused of exploiting its market power after its biggest competitor went bankrupt. While Germany's Federal Cartel Office president, Andreas Mundt, believes this was a deliberate decision by the airline[21], Lufthansa rejected the accusations, pointing out that a software is responsible for pricing.

The present paper addresses this concern. More specifically, we study whether participants in a lab experiment are willing to deceive another participant in order to disguise their own responsibility for an unfair decision. To address our research question, we designed a lab experiment with two variants of a dictator game (Kahneman et al., 1986). The motivation of giving in dictator games like fairness (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999) or altruism (Eckel and Grossman, 1996; Levine, 1998) is controversial in the literature. Studies on social signaling have questioned these interpretations (Dana et al., 2006; Dana et al., 2007; Hoffman et al., 1994). Giving seems vastly reduced when the responsibility of the dictator is diminished, e.g. if receivers' payoff is not only determined by the dictator but also by chance (Andreoni and Bernheim, 2009; Dana et al., 2007) or the anonymity of the decision-maker is increased (Alpizar et al., 2008; Franzen and Pointner, 2012). The results of these studies suggest that some dictators do not share because they care for others. Rather, people feel compelled to give, even if they prefer the outcome that maximizes their own benefit, because they do not want to appear selfish to others. Thus, the underlying motivation for fair behavior could be self-

[21] 'Such algorithms are not written in heaven by God.' (Andreas Mundt in an interview with the Süddeutsche Zeitung (28.12.2017))

interest, combined with the desire to maintain the illusion of not being selfish. This means that fair behavior disappears when individuals have an excuse for not having to give. This is the focus of our study.

Our experimental design relies on a within-subject design and is divided into two parts. In the first part, participants play a standard dictator game (Kahneman et al., 1986) and a dictator game variant of the die-roll game (Fischbacher and Föllmi-Heusi, 2013) for five rounds each. For both games, we randomly allocate participants into groups of two and rematch the group after each round. One group member is randomly chosen as the 'giver', while the other is the 'recipient'. In both games, the giver faces six alternative allocations of her endowment. In the standard dictator game, the giver directly chooses the allocation. In the die-roll variant, she rolls a die and the number rolled indicates the allocation; however, she has the option of misstating the rolled number. At the end of each round, both the allocation as well as the information whether the allocation was made by the giver (in the standard dictator game) or by the die-roll (in the die-roll variant) are displayed. The behavior in these rounds provides us with information about the giving of the subjects in both treatments.

In the second part of the study, participants play one of the two games for only a single round. While the game being played is determined randomly, the giver decides which type is communicated to the recipient. I.e., the giver can choose to report that she decided about the allocation herself or that the allocation was decided by a die-roll. With this design, the giver can take responsibility for an in fact randomly chosen allocation by informing the recipient that the game is the dictator game when in fact they play the die-roll-game. In contrast, when the game is actually the dictator game, the giver can disguise her responsibility by telling the recipient that the allocation results from a die-roll.

Disguising one's responsibility for making unfair decisions is morally costly. The literature on deception shows that individuals tend to have an aversion to liars (Eisenkopf et al., 2017; Sánchez-Pagés and Vorsatz, 2007, 2009) or lying per se (Charness and Dufwenberg, 2006; Gneezy, 2005; Gneezy et al., 2013; Hurkens and Kartik, 2009), even if no other person is harmed (Fischbacher and Föllmi-Heusi, 2013; Gorsch and Rau, 2017). Kajackaite and Gneezy (2017) show that people compare the cost of lying with the incentive to lie. Their results imply that people lie when the incentive is higher than the costs. In our experimental design, there is no monetary incentive to lie about the allocation procedure. However, it can be used to disguise selfish behavior. Both lab and field experiments show that decision-makers behave more selfishly when the anonymity of choices increases (Alpizar et al., 2008; Andreoni and Bernheim, 2009; Dana et al., 2007; Franzen and Pointner, 2012; Hoffman et al., 1994). Dana et al. (2007) as well as Larson and Capra (2009) examine how uncertainty about the recipients' payoff affects dictators' behavior. They show that a large fraction of dictators exploit the payoff uncertainty as a pretext for selfish behavior. Andreoni and Bernheim (2009) conduct a dictator game where nature may choose an unfavorable outcome for the recipient who is unable to identify what caused the outcome. They find that dictators tend to hide behind the possibility that nature chose the unfair outcome. In contrast to these studies, we use a within-subjects design to measure changes in behavior, by comparing the allocations form the first and second part.

Our experimental results show that a large proportion of participants disguise their responsibility for the decision in the standard dictator game. More specifically, over 40 percent of the givers in standard dictator game tell the recipient that the allocation was made by the die-roll. In contrast, for the die-roll variant, we find hardly any givers taking responsibility for the allocation. Moreover, the results show that more selfish givers more often disguise their responsibility in the standard dictator game. We interpret these results as indicating that givers disguise their responsibility to avoid being perceived as selfish or greedy. However, we find no evidence that the disguise itself leads to a more selfish change in allocation. Rather, givers who have already opted for more selfish allocations in the first part tend to disguise their responsibility for the selfish allocation in the second part.

The remainder of the paper is organized as follows. Section 6.2 describes our experimental design and the procedure in detail. Section 6.3 discusses the behavioral predictions. Section 6.4 presents our main results. Section 6.5 presents our conclusion.

6.2 The experimental design and procedures

The experiment had two parts modeling the same two treatments, which differ slightly in their procedures. For both treatments, we randomly allocated participants into groups of two. Each group consisted of a participant in the role of a giver and one in the role of a recipient.[22] Participants were told that the payoff for both group members would depend on the choice made by the giver. The participant in the role of the recipient had no active part.

[22] We used a neutral language in the instructions. We have referred to the giver as player A and the recipient as player B (see Appendix E.I as well as E.II).

The first treatment (*STANDARD-DICTATOR-GAME*) is a variation of the standard dictator game. The giver received an endowment of 10 tokens and was asked to choose one out of a set of six possible allocations. In the second treatment *(GARBLED-DICTATOR-GAME)* the giver rolled a six-sided digital die to determine the allocation of the 10 tokens endowment. More precisely, the giver was instructed to roll the die multiple times (to convince them that the die is fair) but to enter the first number rolled in order to determine the allocation. Lying was obviously feasible, even when it was not actively promoted. Thus, under the assumption of money maximizing individuals, the situation was equivalent to the *STANDARD-DICTATOR-GAME*. Table 6-1 shows the six allocations.

Table 6-1: Payoff structure for both allocation games

Allocation number	Giver gets (token)	Receiver gets (token)
1	0	10
2	2	8
3	4	6
4	6	4
5	8	2
6	10	0

At the beginning of the experiment, we informed the participants that there would be two parts, but that none of their decisions would have an impact on later decisions. In the first part, each participant was assigned the role of the giver for 10 rounds. To control for order effects, half of the participants started with the *STANDARD-DICTATOR-GAME* while the other half started with the *GARBLED-DICTATOR-GAME*. Participants switched to the other game after the fifth round. This procedure also allows us to check for spillovers effects. At the end of the first part, one of the ten rounds, the group assignment, the role (giver or recipient), and the respective allocation of the giver's endowment were randomly selected for payment. Participants did not receive information about their payoff in the first part until the end of the second part. However, at the end of each round the participants were shown a results screen that represented how the allocation would be communicated to the recipient if this round were randomly selected for payoff (see the screenshots in appendix E.III). More precisely, for the *STANDARD-DICTATOR-GAME* the results screen stated that the 'The giver decided the following allocation', while for the *GARBLED-DICTATOR-GAME* the screen showed 'A die was rolled, and the following allocation was chosen'. We implemented the first part for two reasons: First, playing as the giver ensures that each participant is familiar with both treatments and the statements on the results screen. Second, the allocations in this part serves as a benchmark against which we can compare the second part.

The second part lasted for only one round. In contrast to the first part, the roles (giver or receiver) were assigned before the round. The participant in the role of the recipient had no active part. Participants were instructed that one of the two treatments of the first part would be played again. However, only the participant in the role of the giver knew which treatment was actually assigned. Procedures were identical to those of the first part, with one additional choice made by the giver: After choosing the allocation, she was asked to inform the recipient about the allocation procedure. For this purpose, the giver had to choose one out of two messages to be sent to the recipient: *Message 1*: 'The giver decided the following allocation' or *Message 2:* 'A die was rolled, and the following allocation was chosen' (see screenshot Figure E- V in appendix E.III). Note that this procedure was common knowledge for all givers before they made their allocation decision. The recipient, however, was not aware of this additional decision. The two messages clearly differ in the degree of perceived responsibility. While the perceived responsibility in message 1 is entirely on the giver, the responsibility in message 2 is partially shifted to the random procedure of a die-roll. Consequently, the second message can be used as an excuse to justify selfish and greedy behavior. Table 6-2 summarizes the experimental design.

We conducted our experiment at the Laboratory for Economic and Accounting Research (Osnabrück University). A total of 234 participants were involved in ten sessions. The experiment was programmed using z-Tree (Fischbacher, 2007). Recruiting was done by ORSEE (Greiner, 2015). One session lasted approximately 45 minutes. The average income for the participants was 11.86 EUR. Each participant was placed on a randomly assigned seat. The instructions were handed out as well as read out loud. All subjects received their payment privately at the end of the session.

Table 6-2: Experimental Design

			STANDARD-DICTATOR-GAME	GARLED-DICTATOR-GAME
		Before round 1: All participants are randomly assigned to one of the two treatments and decide in the role of the giver for five rounds		
Part 1	*Repeated for five rounds, then the treatment is switched*	**Step 1** Each participant receives 10 tokens	✓	✓
		Step 2 The participants roll a die	---	✓
		Step 3 The participants enter the allocation number	✓	✓
		Step 4 Results screen	✓	✓
		After round 5: The participants decide in the other treatment for five rounds		
		Before the start of part 2: Random assignment to one of the two treatments and role (giver or receiver)		
Part 2	*One Round*	**Step 5** The giver receives 10 tokens	✓	✓
		Step 6 The giver rolls a die	---	✓
		Step 7 The giver enters the allocation number	✓	✓
		Step 8 The giver chooses a message for the recipient	✓	✓
		Step 9 The recipient receives the message and the transferred tokens	✓	✓
		Step 10 Payoff information for part 1	✓	✓

6.3 Behavioral predictions

We start with the prediction regarding the messages in the second part of the experiment. In both treatments, the givers face the same additional decision after choosing an allocation: Inform the receiver whether the allocation was made either by herself or by a die-roll. The additional decision does not affect payoffs, but can be used to disguise the giver's responsibility. If givers prefer to disguise their responsibility in the decision-making process (Andreoni and Bernheim, 2009; Dana et al., 2006; Dana et al., 2007), they have to send the message that 'A die was rolled […]'.

Prediction 6.1: *The message 'A die was rolled […]' is more frequent than the message 'The giver decided […]'.*

Experimental literature indicates that subjects tend to act more selfish when anonymity is increased (Alpizar et al., 2008; Andreoni and Bernheim, 2009; Dana et al., 2007; Franzen and Pointner, 2012; Grossmann, 2015; Hamman et al., 2010; Matthey and Regner, 2011). In both treatments, the message 'A die was rolled […]' allows the giver to hide behind an external event, which gives her some sort of anonymity in the decision-making process. Thus, we predict that this message is associated with more selfish allocations than the message 'The giver decided […]'.

Prediction 6.2: *In both treatments, givers keep more of the endowment for themselves when they send the message 'A die was rolled [...]' compared to the givers who send the message 'The giver decided […]'.*

Next, we address allocation differences between the first and second part. In both parts of the experiment, givers allocate the endowment, either by their own decisions or with the help of a die. However, misreporting the result of the die is possible. For all selfish pure money-maximizing individuals, both treatments will lead to the same payoff. However, it is well known that even in standard dictator-games, participants do not always maximize their individual payoff (Andreoni and Bernheim, 2009; Carpenter et al., 2008; Eckel and Grossman, 1998). A meta-analysis by Engel (2011) confirms these findings and states that, on average, dictators retain about 72 percent of the pie. Turning to the *GARBLED-DICTATOR-GAME*, the literature on lying aversion shows that individuals tend to have preferences for truth telling and that a large fraction of individuals foregoes the potential gains from lying (Abeler et al., 2019). More precisely, only a small number of individuals act as money maximizers, while most individuals report numbers that are close to those drawn (Abeler et al., 2019; Fischbacher and Föllmi-Heusi, 2013; Ruffle and Tobol, 2017). In addition, selfish misstatements are significantly reduced when the true number is observable to the experimenter (Abeler et al., 2019). A comparison with deception games and dictator games confirm that payoffs are higher in dictator games (Gneezy, 2005; Hurkens and Kartik, 2009). Thus, we predict that givers in the *STANDARD-DICTATOR-GAME* allocate the endowment more selfishly than givers in the *GARBLED-DICTATOR-GAME* in the first part of the experiment.

Prediction 6.3: *The allocation of the endowment is more selfish in the STANDARD-DICTATOR-GAME than in the GARBLED-DICTATOR-GAME in the first part of the experiment.*

In part 2, givers in the *STANDARD-DICTATOR-GAME* can disguise their responsibility for a selfish allocation, whereas incentives do not change in the *GARBLED-DICTATOR-GAME*. Consequently, we predict:

Prediction 6.4: *The differences between the STANDARD-DICTATOR-GAME and the GARBLED-DICTATOR-GAME are even stronger in the second part.*

6.4 Results

Before we study the predictions in detail, we have a brief look at the descriptive statistics. Table 6-3 shows the proportion of the messages in both treatments, as well as the average number of tokens retained for the first part and the second part of the experiment. Note, that the subsequent analysis relies on the observations of these participants who act also as a giver in part 2. In Appendix E.V we provide further results, including all observations. All reported test results are two-sided.

Table 6-3: Descriptive statistics

Message in the second part \ Treatment	STANDARD-DICTATOR-GAME Prop.	STANDARD-DICTATOR-GAME Retained Tokens 1st Part	STANDARD-DICTATOR-GAME Retained Tokens 2nd Part	GARBLED-DICTATOR-GAME Prop.	GARBLED-DICTATOR-GAME Retained Tokens 1st Part	GARBLED-DICTATOR-GAME Retained Tokens 2nd Part	Total Prop.
The giver decided […]	56.9%	7.41 (2.00)	8.79 (1.87)	11.9%	5.94 (1.82)	6.29 (4.07)	34.2%
A die was rolled […]	43.1%	7.55 (1.91)	9.36 (0.95)	88.1%	5.65 (2.42)	7.04 (3.30)	65.8%
Observations	58			59			117

Note: The table gives the proportion of the messages in the second part of the experiment for both treatments. In addition, the table gives the average number of tokens retained by the givers in the first and second part.

6.4.1 Analysis of the messages in the second part

Table 6-3 shows the proportion of messages for both treatments. In total, we find that 65.8 percent (77 out of 117) of all givers in both treatments choose the message 'A die was rolled […]'. A binomial probability test confirms that the message 'A die was rolled […]' is significantly more frequent than the message 'The giver decided […]' ($p < .01$, Binomial probability test). Next, we focus on both treatments in greater detail. In the *STANDARD-DICTATOR-GAME*, 43.1 percent (25 out of 58) of the givers send the message 'A die was rolled […], while 56.9 percent send the message 'The giver decided […]'. Turning to the *GARBLED-DICTATOR-GAME* the frequency of the message 'The giver decided […]' is much lower. Only 11.86 percent (7 out of 59) send this message, while the majority uses the message 'A die was rolled […]' (88.14 percent).

We now examine the messages in light of whether the corresponding or non-corresponding allocation procedure was reported. Consequently, we interpret the message 'A die was rolled [...]' as the non-corresponding message in the *STANDARD-DICTATOR-GAME* and we interpret the same message as the corresponding message in the *GARBLED-DICTATOR-GAME*. Conversely, the message 'The giver decided [...]' represents the corresponding message in the *STANDARD-DICTATOR-GAME*, while this message reflects the non-corresponding message in the *GARBLED-DICTATOR-GAME*. The comparison of both treatments indicates that participants use the non-corresponding message significantly more often in the *STANDARD-DICTATOR-GAME* than in the *GARBLED-DICTATOR-GAME* ($p < .01$, Pearson Chi-

Square-Test). Consequently, givers in the *STANDARD-DICTATOR-GAME* frequently blame the die-roll for the allocation. In contrast, only a small number of givers in the *GARBLED-DICTATOR-GAME* take responsibility for the allocation.

Result 6.1:

1. *Overall, the responsibility for allocation is significantly more often disguised behind the roll of the die than the givers themselves taking responsibility for allocation.*
2. *More than 40 percent of the givers in the STANDARD-DICTATOR-GAME deceive the recipients by stating that the allocation was decided by a die-roll.*

Next, we focus on the differences in the allocations between the two messages. Table 6-3 provides information about the number of retained tokens for both treatments, separated by the messages. For both treatments, we find a slightly higher number of retained tokens when givers sent the message 'A die was rolled […]'. More precisely, for the *STANDARD-DICTATOR-GAME* we find that givers retain on average about 8.79 tokens when they sent the message 'The giver decided […]' and about 9.36 when they sent the other message. In the *GARBLED-DICTATOR-GAME*, on the other hand, givers withheld about 6.29 tokens when they sent the message 'The giver decided […]' and about 7.04 when they sent the other message. However, both differences are statistically insignificant (p = .43 for the *STANDARD-DICTATOR-GAME*, p = .71 for the *GARBLED-DICTATOR-GAME*, Mann-Whitney-U-Test).

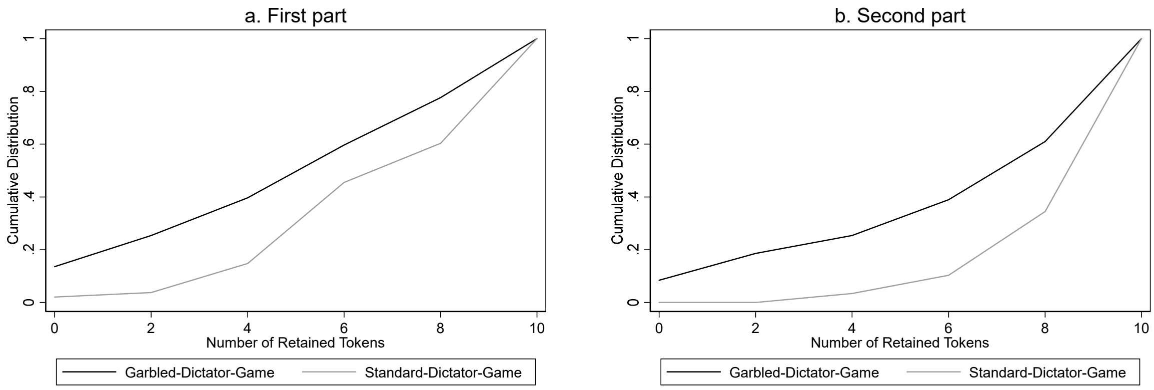
Result 6.2: *Givers do not retain more of the endowment for themselves when they send the message 'A die was rolled [...]' compared to the givers who send the message 'The giver decided […]'.*

6.4.2 Analysis of the allocation in the first and second part

Next, we compare the allocation in the first and second part of the experiment. Figure 6-1 reflects the cumulative distribution of retained tokens for the two treatments for the first (Figure 6-1a) and second part (Figure 6-1b) of the experiment. For both parts, Figure 6-1 indicates that the distribution for the *GARBLED-DICTATOR-GAME* is always above the distribution for the *STANDARD-DICTATOR-GAME*. Thus, givers in the *STANDARD-DICTATOR-*

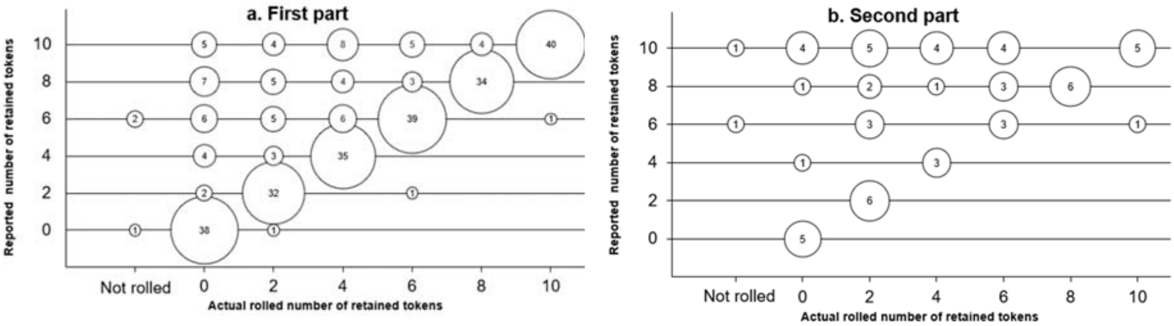
GAME retain more tokens than givers in the *GARBLED-DICTATOR-GAME*. A Mann-Whitney-U-Test confirms that the average number of retained tokens is significantly higher in the *STANDARD-DICTATOR-GAME* than in the *GARBLED-DICTATOR-GAME* for the first part ($p < .01$) as well as for the second part ($p < .01$, Mann-Whitney-U-Test). Our analysis in the Appendix E.IV shows that the order of treatments does not affect allocation.

Figure 6-1: Cumulative distribution of retained tokens



Our previous findings show that givers in the *GARBLED-DICTATOR-GAME*, on average, retain less tokens than givers in the *STANDARD-DICTATOR-GAME*, although misreporting was possible. Next, we turn to the *GARBLED-DICTATOR-GAME* in greater detail. Figure 6-2 provides information about the actual rolled and the reported number of retained tokens. For both, the first and second parts, we find that a certain proportion of givers misreport the initial allocation rolled. However, the proportion of misreporting increases sharply in part 2. More specifically, we find that in 26.1 percent of the decisions in part 1, givers report a different number than actually rolled. In the second part, however, the number of misreports increases to 52.43 percent. The differences are statistically significant ($p < .01$, Pearson chi-square test). Moreover, Figure 6-2 reveals interesting insights of misreporting. First, in most cases, givers enrich themselves through misreporting by overstating their actual numbers. In very few cases, we find that givers use their misreporting to benefit recipients. Second, we find that misreporting givers do not always maximize their own payoff.

Figure 6-2: Distribution of reported number of retained tokens in the GARBLED-DICTATOR-GAME



Note: The figure shows the distribution of the reported number of retained tokens depended on the actual rolled allocation of tokens in the first and second parts of the experiment. The size of the circles corresponds to the number of subjects.

Table 6-4 disentangles the previous observation in greater detail. Models I – III examine the relationship between the reported and the rolled allocation. The reported number of retained tokens denotes the dependent variable. The variable *Rolled Allocation* specifies how many tokens should be retained according to the die-roll. The variable *2nd Part* is a dummy variable that denotes the second part. Model I shows a strong correlation between the rolled and reported allocation. More specifically, the models suggest that the givers adhere to the rolled allocation to some extent. However, it turns out that givers, on average, retain more tokens than actually allocated by the die-roll. Model II controls for the second part. The variable enters positive and significant, while the variable *Rolled Allocation* remains highly significant. The interaction term in model III reveals that givers follow the rolled allocation to a lesser extend in the second part.

Models IV – VI now study the relationship between the rolled number and the likelihood of misreporting. The dependent variable is a dummy variable that indicates whether the number rolled was reported (=0) or whether the giver reported a different number (=1). Model IV shows that misreporting becomes less likely the higher the rolled allocation. Model V controls for the second part. The model indicates that misreporting is more likely in the second part in comparison with the first part. The variable *Rolled Allocation* remains highly significant. An interaction term reveals no differential effects between the first and second part (model VI).

Result 6.3:

1. *Participants in both treatments, on average, do not opt for the money-maximizing allocation. In the STANDARD-DICTATOR-GAME, however, the givers allocate themselves a significantly higher share of endowment than in the GARBLED-DICTATOR-GAME.*

2. *Although misreporting was possible, giver foregoes the potential gains from lying and followed the allocation by the die-roll in 74% of all decisions. For the second part, we find that misreporting becomes more frequent.*

Table 6-4: Misreporting in the GARBLED-DICTATOR-GAME

Dep. Var.:	**Tobit Number of retained tokens**			**Probit Misreporting**		
	I	**II**	**III**	**IV**	**V**	**VI**
Rolled Allocation	0.860*** (0.0560)	0.864*** (0.0546)	0.916*** (0.0589)	-0.13*** (0.0216)	-0.13*** (0.0221)	-0.14*** (0.0244)
2nd Part		1.893*** (0.474)	3.301*** (0.746)		0.691*** (0.192)	0.615** (0.310)
Rolled Allocation*2nd Part			-0.367** (0.149)			0.0198 (0.0582)
Constant	2.668*** (0.286)	2.335*** (0.292)	2.114*** (0.303)	-0.0154 (0.113)	-0.142 (0.120)	-0.128 (0.126)
Observations	*349*	*349*	*349*	*349*	*349*	*349*
Subjects	*116*	*116*	*116*	*116*	*116*	*116*
(Pseudo) R-Squared	*0.12*	*0.13*	*0.13*	*0.08*	*0.11*	*0.11*

*Note that three participants did not roll the die. These observations are omitted in the models. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

Next, we compare the allocation in the first and second part. Table 6-3 as well as Figure 6-1 suggest that allocations tend to be more selfish in the second part. Table 6-5 provides results from a tobit-regression. The number of retained tokens represent the dependent variable. The *GARBLED-DICTATOR-GAME* provides the benchmark, while the variable *Standard-Dictator-Game* denotes the other treatment. The variable *2nd Part* is a dummy variable that indicates the second part of the experiment (=1). Models I and II include all givers. Models III and IV contain only these givers who sent the message 'The giver decided […]' in the second part of the experiment, while models V and VI include the other givers.

All our models indicate the same three results: First, on average, givers in the *STANDARD-DICTATOR-GAME* retain more tokens than givers in the *GARBLED-DICTATOR-GAME*. Second, the second part generally leads to a more selfish allocation. Last, but not least, the interaction terms reveal no differential effects for both treatments. These observations are independent of the message (see models III-VI). Consequently, we find that the number of retained tokens increases significantly in the second part of both treatments. However, we cannot detect any differential effect between the messages.

Result 6.4: *In both treatments, givers choose a more selfish allocation in the second part. We find no evidence that disguise itself leads to a more selfish change in allocation.*

Table 6-5: Tobit Regression for the individual decision of token allocation

Dep. Var.: Number of retained tokens	**Benchmark: GARBLED-DICTATOR-GAME**					
	All givers		Givers who sent the message: The giver decided […]		A die was rolled […]	
	I	**II**	**III**	**IV**	**V**	**VI**
Standard-Dictator-Game	2.59***	2.41***	2.25***	1.96**	2.80***	2.57***
	(0.330)	(0.357)	(0.720)	(0.781)	(0.441)	(0.475)
2^{nd} Part	2.35***	1.80***	2.28***	0.807	2.39***	1.94***
	(0.462)	(0.618)	(0.799)	(1.765)	(0.565)	(0.662)
Standard-Dictator-Game*2^{nd} Part		1.230		1.836		1.626
		(0.931)		(1.977)		(1.278)
Constant	3.74***	6.18***	4.01***	6.52***	3.68***	6.13***
	(0.578)	(0.247)	(1.125)	(0.706)	(0.695)	(0.264)
Observations	702	702	240	240	462	462
Subjects	117	117	40	40	77	77
R-Squared	0.03	0.03	0.02	0.02	0.03	0.03

*Tobit-regression. *** p<0.01, ** p<0.05, * p<0.1*

Our previous results do not confirm prediction 6.2 and 6.4. We now disentangle the motivation for sending the message 'A die was rolled […]' in greater detail. Table 6-6 provides results form a probit regression. The dependent variable is the individual decision to send the message 'A die was rolled […]', with robust standard errors. The variable *Retained tokens (Part 2)* indicates how many tokens a giver retained in the second part, while the variable *Avg. Ret. Tokens (Part 1)* reflects how many tokens the giver retained on average in the first part in the same treatment. The variable *2^{nd} Part – 1^{st} Part* indicates the difference between the number of retained tokens in the second part and the average number of retained tokens in the first part.

Models I – III focus on *STANDARD-DICTATOR-GAME*. Model II shows that the number of retained tokens has a significant effect on the probability of sending the message 'A die was rolled […]' in the *STANDARD-DICTATOR-GAME*. More specifically, we find that givers who favor themselves more lie about the allocation procedure. Model II controls for the average number of retained tokens in the first part. However, we only find an insignificant effect, while the number of retained tokens in the second part remains significant. Finally, we are interested in whether the change in retained tokens between the two parts affects the choice of message.

Therefore, model III controls for the difference in the allocation of both parts. However, the coefficient of the variable shows only insignificant differences. Consequently, we find that unfair behavior favors sending the message 'A die was rolled […]' in the *STANDARD-DICTATOR-GAME*. However, our models do not suggest that changes in allocation behavior cause participants to disguise their responsibility. Model IV – VI now focus on the *GARBLED-DICTATOR-GAME*. However, none of our independent variables show a significant impact.

Result 6.5: *More selfish givers more often disguise their responsibility in the STANDARD-DICTATOR-GAME. We do not find that changes in allocation behavior cause participants to disguise their responsibility.*

Table 6-6: Probit Regression for likelihood of sending the message 'A die was rolled […]'

Dep. Var.: Indiv. decision to send the message 'A die was rolled […]'	**Only STANDARD-DICTATOR-GAME**			**Only GARBLED-DICTATOR-GAME**		
	I	**II**	**III**	**IV**	**V**	**VI**
Retained tokens (2nd Part)	0.172*	0.171*		-0.033	-0.052	
	(0.0986)	(0.0989)		(0.0640)	(0.0771)	
Avg. Ret. tokens (1st Part)		0.0064			0.0666	
		(0.0877)			(0.102)	
2nd Part – 1st Part			0.0525			-0.0542
			(0.0703)			(0.0749)
Constant	-1.74*	-1.78*	-0.258	-0.96**	-1.22**	-1.1***
	(0.901)	(1.066)	(0.199)	(0.473)	(0.531)	(0.223)
Observations	58	58	58	59	59	59
Subjects	58	58	58	59	59	59
Pseudo-R-Squared	0.03	0.03	0.01	0.01	0.02	0.02

*Probit regression for likelihood of sending the non-corresponding messages. Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01*

6.5 Conclusion

In this paper, we study whether participants are willing to deceive another participant in order to disguise their own responsibility for an unfair decision. Some literature on the motivation of giving suggest that the underlying motivation for fair behavior could be self-interest, combined with the desire to maintain the illusion of not being selfish (Andreoni and Bernheim, 2009; Dana et al., 2007). More precisely, fair behavior should disappear when individuals have an excuse for not having to give. We addressed this concern in our paper.

Our results show that givers in our treatments do indeed disguise their responsibility. In more than 40 percent of the decisions, givers in the *STANDARD-DICTATOR-GAME* indicate to the recipient that the allocation was made by a roll of the die. This disguise is used especially by people who have chosen a selfish allocation. In contrast, for the *GARBLED-DICTATOR-GAME* we find only a few participants who state that the allocation was made by themselves. Both supports the prediction that maintaining a positive image is the motivation for disguising responsibility. We had a particular interest in whether fair behavior disappears when individuals have an excuse for not having to give. However, our results cannot confirm this prediction. Regardless of the treatment, we find that the allocation becomes significantly more selfish in the second part. Nevertheless, we find no evidence that disguise itself leads to a more selfish change in allocation.

Our paper aims to contribute to a more comprehensive understanding of behaviors in unfair decision-making situations. While the literature shows that individuals often delegate unfair decisions (Bartling and Fischbacher, 2011; Erat, 2013) or hide behind nature (Andreoni and Bernheim, 2009; Dana et al., 2007), our participants were not able to abdicate responsibility but to disguise it. We interpret this disguise as an indication that givers do not want to be perceived as selfish or greedy. The dynamic interaction between givers and receivers in a natural setting makes it difficult to evaluate causal links of such image concerns and unfair decisions because decision-making is strongly influenced by the social distance (Bohnet and Frey, 1999; Candelo et al., 2018; Leider et al., 2010), the decision context (Tversky and Simonson, 1993; Winking and Mizer, 2013) or anonymity (Alpizar et al., 2008; Franzen and Pointner, 2012; Soetevent, 2005). Our simple lab experiment addresses these concerns. We intentionally created a one-shot, anonymous environment in which there are no incentives to disguise responsibility. However, even in such abstract, clinical decision-making situations, people are willing to deceive in order to disguise their responsibility for an unfair decision and thus prevent themselves from being perceived as selfish or greedy.

References

Abbink, K., Brandts, J., Herrmann, B., Orzen, H., 2010. Intergroup Conflict and Intra-Group Punishment in an Experimental Contest Game. American Economic Review 100 (1), 420–447.

Abbink, K., Brandts, J., Herrmann, B., Orzen, H., 2012. Parochial Altruism in Inter-Group Conflicts. Economics Letters 117 (1), 45–48.

Abeler, J., Falk, A., Goette, L., Huffman, D., 2011. Reference Points and Effort Provision. American Economic Review 101 (2), 470–492.

Abeler, J., Nosenzo, D., Raymond, C., 2019. Preferences for Truth-Telling. Econometrica 87 (4), 1115–1153.

Acemoglu, D., Jackson, M.O., 2015. History, Expectations, and Leadership in the Evolution of Social Norms. The Review of Economic Studies 82 (2), 423–456.

Ahn, T.-K., Isaac, R.M., Salmon, T.C., 2008. Endogenous Group Formation. Journal of Public Economic Theory 10 (2), 171–194.

Alan, S., Ertac, S., Kubilay, E., Loranth, G., 2019. Understanding Gender Differences in Leadership. The Economic Journal 130 (626), 263–289.

Alpizar, F., Carlsson, F., Johansson-Stenman, O., 2008. Anonymity, Reciprocity, and Conformity: Evidence from Voluntary Contributions to a National Park in Costa Rica. Journal of Public Economics 92 (5-6), 1047–1060.

Anderson, J., Burks, S.V., Carpenter, J., Götte, L., Maurer, K., Nosenzo, D., Potter, R., Rocha, K., Rustichini, A., 2013. Self-Selection and Variations in the Laboratory Measurement of Other-Regarding Preferences across Subject Pools: Evidence from one College Student and two Adult Samples. Experimental Economics 16 (2), 170–189.

Andreoni, J., Bernheim, B.D., 2009. Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects. Econometrica 77 (5), 1607–1636.

Andreoni, J., Miller, J., 2002. Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. Econometrica 70 (2), 737–753.

Antonakis, J., Bendahan, S., Jacquart, P., Lalive, R., 2010. On Making Causal Claims: A Review and Recommendations. The Leadership Quarterly 21 (6), 1086–1120.

Arbak, E., Villeval, M.-C., 2013. Voluntary Leadership: Motivation and Influence. Social Choice and Welfare 40 (3), 635–662.

Arechar, A.A., Gächter, S., Molleman, L., 2018. Conducting Interactive Experiments Online. Experimental Economics 21 (1), 99–131.

Ariely, D., Bracha, A., Meier, S., 2009. Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. American Economic Review 99 (1), 544–555.

Avolio, B.J., Gardner, W.L., 2005. Authentic Leadership Development: Getting to the Root of Positive Forms of Leadership. The Leadership Quarterly 16 (3), 315–338.

Balliet, D., Wu, J., Dreu, C.K.W. de, 2014. Ingroup Favoritism in Cooperation: A Meta-Analysis. Psychological Bulletin 140 (6), 1556–1581.

Barberis, N., Thaler, R., 2003. A Survey of Behavioral Finance. In: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.). Handbook of the Economics of Finance, 1B ed. Financial Markets and Asset Pricing: Elsevier, 1053–1128.

Barr, A., Serneels, P., 2009. Reciprocity in the Workplace. Experimental Economics 12 (1), 99–112.

Barron, K., Nurminen, T., 2020. Nudging Cooperation in Public Goods Provision. Journal of Behavioral and Experimental Economics 88, 101542.

Bartling, B., Fischbacher, U., 2011. Shifting the Blame: On Delegation and Responsibility. The Review of Economic Studies 79 (1), 67–87.

Bedi, A., Alpaslan, C.M., Green, S., 2016. A Meta-Analytic Review of Ethical Leadership Outcomes and Moderators. Journal of Business Ethics 139 (3), 517–536.

Bénabou, R., Tirole, J., 2003. Intrinsic and Extrinsic Motivation. The Review of Economic Studies 70 (3), 489–520.

Berinsky, A.J., Huber, G.A., Lenz, G.S., 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. Political Analysis 20 (3), 351–368.

Bernhard, H., Fehr, E., Fischbacher, U., 2006a. Group Affiliation and Altruistic Norm Enforcement. American Economic Review 96 (2), 217–221.

Bernhard, H., Fischbacher, U., Fehr, E., 2006b. Parochial Altruism in Humans. Nature 442 (7105), 912–915.

Black, E.L., Carnes, T.A., Richardson, V.J., 2000. The Market Valuation of Corporate Reputation. Corporate Reputation Review 3 (1), 31–42.

Blanco, M., Engelmann, D., Koch, A.K., Normann, H.-T., 2014. Preferences and beliefs in a sequential social dilemma: a within-subjects analysis. Games and Economic Behavior 87, 122–135.

Blasco, A., Jung, O.S., Lakhani, K.R., Menietti, M., 2019. Incentives for public goods inside organizations: Field experimental evidence. Journal of Economic Behavior & Organization 160, 214–229.

Bochet, O., Page, T., Putterman, L., 2006. Communication and Punishment in Voluntary Contribution Experiments. Journal of Economic Behavior & Organization 60 (1), 11–26.

Bock, O., Baethe, I., Nicklisch, A., 2014. hroot: Hamburg Registration and Organization Online Tool. European Economic Review 71, 117–120.

Boezeman, E.J., Ellemers, N., 2014. Volunteer Leadership: The Role of Pride and Respect in Organizational Identification and Leadership Satisfaction. Leadership 10 (2), 160–173.

Böhm, R., Rockenbach, B., 2013. The Inter-Group Comparison-Intra-Group Cooperation Hypothesis: Comparisons between Groups Increase Efficiency in Public Goods Provision. PLOS ONE 8 (2), e56152.

Bohnet, I., Frey, B.S., 1999. Social Distance and Other-Regarding Behavior in Dictator Games: Comment. American Economic Review 89 (1), 335–339.

Bolton, G.E., Ockenfels, A., 2000. ERC: A Theory of Equity, Reciprocity, and Competition. American Economic Review 90 (1), 166–193.

Bornstein, G., Gneezy, U., 2002. Price Competition between Teams. Experimental Economics 5 (1), 29–38.

Bornstein, G., Kugler, T., Budescu, D.V., Selten, R., 2008. Repeated Price Competition between Individuals and between Teams. Journal of Economic Behavior & Organization 66 (3-4), 808–821.

Bortolotti, S., Casari, M., Pancotto, F., 2015. Norms of Punishment: Experiments with Students and the General Population. Economic Inquiry 53 (2), 1207–1223.

Bowles, S., Polanía-Reyes, S., 2012. Economic Incentives and Social Preferences: Substitutes or Complements? Journal of Economic Literature 50 (2), 368–425.

Brekke, K.A., Hauge, K.E., Lind, J.T., Nyborg, K., 2011. Playing with the Good Guys. A Public Good Game with Endogenous Group Formation. Journal of Public Economics 95 (9-10), 1111–1118.

Brosig, J., Weimann, J., 2003. The Effect of Communication Media on Cooperation. German Economic Review (2), 217–241.

Bucciol, A., Montinari, N., Piovesan, M., 2019. It Wasn't Me! Visibility and Free Riding in Waste Disposal. Ecological Economics 157, 394–401.

Buchanan, J.M., 1965. An Economic Theory of Clubs. Economica 32 (125), 1–14.

Buckley, P.J., Carter, M.J., 2004. A Formal Analysis of Knowledge Combination in Multinational Enterprises. Journal of International Business Studies 35 (5), 371–384.

Burton-Chellew, M.N., West, S.A., 2012. Pseudocompetition Among Groups Increases Human Cooperation in a Public-Goods Game. Animal Behaviour 84 (4), 947–952.

Camerer, C.F., Hogarth, R.M., 1999. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. Journal of Risk and Uncertainty 19 (1/3), 7–42.

Candelo, N., Eckel, C., Johnson, C., 2018. Social Distance Matters in Dictator Games: Evidence from 11 Mexican Villages. Games 9 (4), 77.

Cappelen, A.W., Reme, B.-A., Sørensen, E.Ø., Tungodden, B., 2016. Leadership and Incentives. Management Science 62 (7), 1944–1953.

Carpenter, J., Connolly, C., Myers, C.K., 2008. Altruistic Behavior in a Representative Dictator Experiment. Experimental Economics 11 (3), 282–298.

Carpenter, J., Verhoogen, E., Burks, S., 2005. The effect of stakes in distribution experiments. Economics Letters 86 (3), 393–398.

Carpenter, J.P., 2007. Punishing Free-Riders: How Group Size Affects Mutual Monitoring and the Provision of Public Goods. Games and Economic Behavior 60 (1), 31–51.

Cartwright, E., Patel, A., 2010. Public Goods, Social Norms, and Naïve Beliefs. Journal of Public Economic Theory 12 (2), 199–223.

Cason, T.N., Sheremeta, R.M., Zhang, J., 2012. Communication and Efficiency in Competitive Coordination Games. Games and Economic Behavior 76 (1), 26–43.

Centorrino, S., Concina, L., 2013. A Competitive Approach to Leadership in Public Good Games. Working Paper.

Chadi, A., Jeworrek, S., Mertins, V., 2017. When the Meaning of Work Has Disappeared: Experimental Evidence on Employees' Performance and Emotions. Management Science 63 (6), 1696–1707.

Chakravarty, S., Fonseca, M.A., 2017. Discrimination via Exclusion: An Experiment on Group Identity and Club Goods. Journal of Public Economic Theory 19 (1), 244–263.

Charness, G., 2010. Laboratory Experiments: Challenges and promise: A Review of" Theory and Experiment: What are the Questions?" by Vernon Smith. Journal of Economic Behavior & Organization 73 (1), 21–23.

Charness, G., Dufwenberg, M., 2006. Promises and Partnership. Econometrica 74 (6), 1579–1601.

Charness, G., Kuhn, P., 2011. Lab Labor: What can Labor Economists Learn from the Lab? In: Ashenfelter, O., Card, D. (Eds.). Handbook of Labor Economics, 4A ed.: North-Holland, 229–330.

Charness, G., Rigotti, L., Rustichini, A., 2007. Individual Behavior and Group Membership. American Economic Review 97 (4), 1340–1352.

Charness, G., Yang, C.-L., 2014. Starting Small Toward Voluntary Formation of Efficient Large Groups in Public Goods Provision. Journal of Economic Behavior & Organization 102, 119–132.

Chaudhuri, A., 2011. Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature. Experimental Economics 14 (1), 47–83.

Chen, G., Kanfer, R., 2006. Toward a Systems Theory of Motivated Behavior in Work Teams. Research in Organizational Behavior 27, 223–267.

Chen, Y., Li, S.X., 2009. Group Identity and Social Preferences. American Economic Review 99 (1), 431–457.

Choi, J.-K., Bowles, S., 2007. The Coevolution of Parochial Altruism and War. Science 318 (5850), 636–640.

Cinyabuguma, M., Page, T., Putterman, L., 2005. Cooperation Under the Threat of Expulsion in a Public Goods Experiment. Journal of Public Economics 89 (8), 1421–1435.

Clot, S., Grolleau, G., Ibanez, L., 2018. Shall We Pay All? An Experimental Test of Random Incentivized Systems. Journal of Behavioral and Experimental Economics 73, 93–98.

Conger, J.A., Kanungo, R.N., Menon, S.T., 2000. Charismatic Leadership and Follower Effects. Journal of Organizational Behavior 21, 747–767.

Cox, J.C., Friedman, D., Gjerstad, S., 2007. A Tractable Model of Reciprocity and Fairness. Games and Economic Behavior 59 (1), 17–45.

Croson, R., Donohue, K., 2002. Experimental Economics and Supply-Chain Management. Interfaces 32 (5), 74–82.

Croson, R.T., Marks, M.B., 2000. Step Returns in Threshold Public Goods: A Meta-and Experimental Analysis. Experimental Economics 2 (3), 239–259.

Cueva, C., Dessi, R., 2012. Charitable Giving, Self-Image and Personality.

Dal Bó, P., 2005. Cooperation Under the Shadow of the Future: Experimental Evidence from Infinitely Repeated Games. American Economic Review 95 (5), 1591–1604.

Dal Bó, P., Fréchette, G.R., 2011. The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence. American Economic Review 101 (1), 411–429.

Dal Bó, P., Fréchette, G.R., 2018. On the Determinants of Cooperation in Infinitely Repeated Games: A Survey. Journal of Economic Literature 56 (1), 60–114.

Dal Bó, P., Fréchette, G.R., 2019. Strategy Choice in the Infinitely Repeated Prisoner's Dilemma. American Economic Review 109 (11), 3929–3952.

Dana, J., Cain, D.M., Dawes, R.M., 2006. What You don't Know won't hurt Me: Costly (but Quiet) Exit in Dictator Games. Organizational Behavior and Human Decision Processes 100 (2), 193–201.

Dana, J., Weber, R.A., Kuang, J.X., 2007. Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. Economic Theory 33 (1), 67–80.

Dannenberg, A., 2015. Leading by Example Versus Leading by Words in Voluntary Contribution Experiments. Social Choice and Welfare 44 (1), 71–85.

Dawes, R.M., 1980. Social Dilemmas. Annual Review of Psychology 31 (1), 169–193.

Dechenaux, E., Kovenock, D., Sheremeta, R.M., 2015. A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments. Experimental Economics 18 (4), 609–669.

Delmas, M., Keller, A., 2005. Free Riding in Voluntary Environmental Programs: The Case of the U.S. EPA WasteWise Program. Policy Sciences 38 (2-3), 91–106.

Diehl, M., 1990. The Minimal Group Paradigm: Theoretical Explanations and Empirical Findings. European Review of Social Psychology 1 (1), 263–292.

Dirks, K.T., Ferrin, D.L., 2002. Trust in Leadership: Meta-Analytic Findings and Implications for Research and Practice. Journal of Applied Psychology 87 (4), 611–628.

Drouvelis, M., Nosenzo, D., 2013. Group Identity and Leading-by-Example. Journal of Economic Psychology 39, 414–425.

Drucker, P.F., 2012. Management Challenges for the 21st Century: Foreword by Will Hutton and Linda Holbeche, Rev. ed., The Classic Drucker Collection. Hoboken: Taylor and Francis.

Duch, M., Grossmann, M.R.P., Lauer, T., 2020. z-Tree unleashed: A novel Client-Integrating Architecture for Conducting z-tree Experiments over the Internet. Journal of Behavioral and Experimental Finance (28).

Dufwenberg, M., Kirchsteiger, G., 2004. A Theory of Sequential Reciprocity. Games and Economic Behavior 47 (2), 268–298.

Eckel, C.C., Grossman, P.J., 1996. Altruism in Anonymous Dictator Games. Games and Economic Behavior 16 (2), 181–191.

Eckel, C.C., Grossman, P.J., 1998. Are Women Less Selfish Than Men?: Evidence From Dictator Experiments. The Economic Journal 108 (448), 726–735.

Efferson, C., Lalive, R., Fehr, E., 2008. The Coevolution of Cultural Groups and Ingroup Favoritism. Science 321 (5897), 1844–1849.

Eichenseer, M., 2021. Leading by Example in Public Goods Experiments: What Do We Know? Working Paper.

Eisenkopf, G., 2014. The Impact of Management Incentives in Intergroup Contests. European Economic Review 67, 42–61.

Eisenkopf, G., 2018. The Long-Run Effects of Communication as a Conflict Resolution Mechanism. Journal of Economic Behavior & Organization 154, 121–136.

Eisenkopf, G., 2020. Words and Deeds – Experimental Evidence on Leading-by-Example. The Leadership Quarterly 31 (4).

Eisenkopf, G., Gurtoviy, R., Utikal, V., 2017. Punishment Motives for Small and Big Lies. Journal of Economics & Management Strategy 26 (2), 484–498.

Eisenkopf, G., Kölpin, T., 2022. Leadership and Cooperation in Growing Teams. Working Paper.

Eisenkopf, G., Walter, C., 2021. Leadership with Imperfect Monitoring. Leadership Quarterly.

Engel, C., 2011. Dictator Games: A Meta Study. Experimental Economics 14 (4), 583–610.

Erat, S., 2013. Avoiding Lying: The Case of Delegated Deception. Journal of Economic Behavior & Organization 93, 273–278.

Ertac, S., Gurdal, M.Y., 2013. Deciding to Decide: Gender, Leadership and Risk-Taking in Groups. Journal of Economic Behavior & Organization 83 (1), 24–30.

Ewoh, A.I.E., 2013. Managing and Valuing Diversity. Public Personnel Management 42 (2), 107–122.

Falk, A., Fehr, E., Fischbacher, U., 2003. On the Nature of Fair Behavior. Economic Inquiry 41 (1), 20–26.

Falk, A., Fischbacher, U., 2006. A Theory of Reciprocity. Games and Economic Behavior 54 (2), 293–315.

Falk, A., Heckman, J.J., 2009. Lab Experiments are a Major Source of Knowledge in the Social Sciences. Science 326 (5952), 535–538.

Falk, A., Meier, S., Zehnder, C., 2013. Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples. Journal of the European Economic Association 11 (4), 839–852.

Fang-Fang, Y.U., Bin, Z.U., 2018. The Minimal Group Paradigm: Operation, Psychological Mechanism and New Application. Journal of Psychological Science (3), 713.

Fehr, E., Gächter, S., 2000. Cooperation and Punishment in Public Goods Experiments. American Economic Review 90 (4), 980–994.

Fehr, E., Kirchler, E., Weichbold, A., Gächter, S., 1998. When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets. Journal of Labor Economics 16 (2), 324–351.

Fehr, E., List, J.A., 2004. The Hidden Costs and Returns of Incentives—Trust and Trustworthiness among Ceos. Journal of the European Economic Association 2 (5), 743–771.

Fehr, E., Schmidt, K.M., 1999. A Theory of Fairness, Competition, and Cooperation. The Quarterly Journal of Economics 114 (3), 817–868.

Fehr, E., Tougareva, E., Fischbacher, U., 2014. Do High Stakes and Competition Undermine Fair Behaviour? Evidence from Russia. Journal of Economic Behavior & Organization 108, 354–363.

Feltovich, N., 2011. Whats's to Know about Laboratory Experimentation in Economics? Journal of Economic Surveys 25 (2), 371–379.

Figuieres, C., Masclet, D., Willinger, M., 2012. Vanishing Leadership and Declining Reciprocity in a Sequential Contribution Experiment. Economic Inquiry 50 (3), 567–584.

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

Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in Disguise - An Experimental Study on Cheating. Journal of the European Economic Association 11 (3), 525–547.

Fischbacher, U., Gächter, S., 2004. Heterogeneous Motivations and the Dynamics of Free Riding in Public Goods. WZB Economics Seminar Series.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. Economics Letters 71 (3), 397–404.

Frackenpohl, G., Hillenbrand, A., Kube, S., 2016. Leadership Effectiveness and Institutional Frames. Experimental Economics 19 (4), 842–863.

Franzen, A., Pointner, S., 2012. Anonymity in the Dictator Game Revisited. Journal of Economic Behavior & Organization 81 (1), 74–81.

Friedman, D., Sunder, S., 1994. Experimental Methods: A Primer for Economists. Cambridge: Cambridge Univ. Press.

Gächter, S., Fehr, E., 2002. Fairness in the Labour Market. In: Bolle, F., Lehmann-Waffenschmidt, M. (Eds.). Surveys in Experimental Economics: Bargaining, Cooperation and Election Stock Markets: Springer, 95–132.

Gächter, S., Nosenzo, D., Renner, E., Sefton, M., 2012. Who Makes a Good Leader? Cooperativeness, Optimism, and Leading-by-Example. Economic Inquiry 50 (4), 953–967.

Gächter, S., Renner, E., 2010. The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments. Experimental Economics 13 (3), 364–377.

Gächter, S., Renner, E., 2018. Leaders as Role Models and 'Belief Managers' in Social Dilemmas. Journal of Economic Behavior & Organization 154, 321–334.

Gneezy, U., 2005. The Role of Consequences. American Economic Review 95 (1), 384–394.

Gneezy, U., Rockenbach, B., Serra-Garcia, M., 2013. Measuring Lying Aversion. Journal of Economic Behavior & Organization 93, 293–300.

Gneezy, U., Rustichini, A., 2000. Pay Enough or don't Pay at All. The Quarterly Journal of Economics 115 (3), 791–810.

Goeree, J.K., Holt, C.A., Laury, S.K., 2002. Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior. Journal of Public Economics 83 (2), 255–276.

Google, 2014. Guide: Understand Team Effectiveness. https://rework.withgoogle.com/guides/understanding-team-effectiveness/steps/introduction/ (accessed 03 January 2022).

Gorsch, K., Rau, H.A., 2017. Gender Differences in Honesty: The Role of Social Value Orientation. Journal of Economic Psychology 62, 258–267.

Greiner, B., 2015. Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. Journal of the Economic Science Association 1 (1), 114–125.

Grossmann, Z., 2015. Self-Signaling Versus Social-Signaling in Giving. Journal of Economic Behavior & Organization 117, 26–39.

Gunnthorsdottir, A., Rapoport, A., 2006. Embedding Social Dilemmas in Intergroup Competition reduces Free-Riding. Organizational Behavior and Human Decision Processes 101 (2), 184–199.

Gürerk, Ö., Lauer, T., Scheuermann, M., 2018. Leadership with Individual Rewards and Punishments. Journal of Behavioral and Experimental Economics 74, 57–69.

Güth, W., Levati, M.V., Sutter, M., Van der Heijden, E., 2007. Leading by Example With and Without Exclusion Power in Voluntary Contribution Experiments. Journal of Public Economics 91 (5-6), 1023–1042.

Haeckl, S., Sausgruber, R., Tyran, J.-R., 2018. Work Motivation and Teams. Working Paper.

Haigner, S.D., Wakolbinger, F., 2010. To Lead or not to Lead: Endogenous Sequencing in Public Goods Games. Economics Letters 108 (1), 93–95.

Hamman, J.R., Loewenstein, G., Weber, R.A., 2010. Self-Interest through Delegation: An Additional Rationale for the Principal-Agent Relationship. American Economic Review 100 (4), 1826–1846.

Hassan, A., Ahmed, F., 2011. Authentic Leadership, Trust and Work Engagement. International Journal of Human and Social Sciences 6 (3), 164–170.

Heinz, M., Jeworrek, S., Mertins, V., Schumacher, H., Sutter, M., 2020. Measuring the Indirect Effects of Adverse Employer Behaviour on Worker Productivity: a field Experiment. The Economic Journal 130 (632), 2546–2568.

Herrmann, B., Orzen, H., 2008. The Appearance of Homo Rivalis: Social Preferences and the Nature of Rent Seeking. CeDEx discussion paper series.

Hoffman, E., McCabe, K., Shachat, K., Smith, V., 1994. Preferences, Property Rights, and Anonymity in Bargaining Games. Games and Economic Behavior 7 (3), 346–380.

Hooghe, M., Stolle, D., Mahéo, V.-A., Vissers, S., 2010. Why Can't a Student be More Like an Average Person?: Sampling and Attrition Effects in Social Science Field and Laboratory Experiments. The ANNALS of the American Academy of Political and Social Science 628 (1), 85–96.

Horton, J.J., Rand, D.G., Zeckhauser, R.J., 2011. The Online Laboratory: Conducting Experiments in a Real Labor Market. Experimental Economics 14 (3), 399–425.

Huffman, D., Bognanno, M., 2018. High-Powered Performance Pay and Crowding Out of Nonmonetary Motives. Management Science 64 (10), 4669–4680.

Hurkens, S., Kartik, N., 2009. Would I Lie to You? On Social Preferences and Lying Aversion. Experimental Economics 12 (2), 180–192.

Isaac, R., Walker, J.M., Williams, A.W., 1994. Group Size and the Voluntary Provision of Public Goods. Journal of Public Economics 54 (1), 1–36.

Isaac, R.M., Walker, J.M., 1988a. Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism. Economic Inquiry 26, 585–608.

Isaac, R.M., Walker, J.M., 1988b. Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. The Quarterly Journal of Economics, 179–199.

Janssens, W., Kramer, B., 2016. The Social Dilemma of Microinsurance: Free-Riding in a Framed Field Experiment. Journal of Economic Behavior & Organization 131, 47–61.

Jenkins, S.P., 1995. Easy Estimation Methods for Discrete-Time Duration Models. Oxford Bulletin of Economics and Statistics 57 (1), 129–138.

Jeworrek, S., Mertins, V., 2021. Mission, Motivation, and the Active Decision to Work for a Social Cause. Nonprofit and Voluntary Sector Quarterly.

Johnson, N.D., Mislin, A.A., 2011. Trust games: A Meta-Analysis. Journal of Economic Psychology 32 (5), 865–889.

Jong, B.A. de, Dirks, K.T., Gillespie, N., 2016. Trust and Team Performance: A Meta-Analysis of Main Effects, Moderators, and Covariates. Journal of Applied Psychology 101 (8), 1134–1150.

Jun, E., Hsieh, G., Reinecke, K., 2017. Types of Motivation Affect Study Selection, Attention, and Dropouts in Online Experiments. Proceedings of the ACM on Human-Computer Interaction 1 (CSCW), 1–15.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1986. Fairness and the Assumptions of Economics. The Journal of Business 59 (S4), S285.

Kajackaite, A., Gneezy, U., 2017. Incentives and Cheating. Games and Economic Behavior 102, 433–444.

Kaplan, E.L., Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association 53 (282), 457–481.

Keser, C., Van Winden, F., 2000. Conditional Cooperation and Voluntary Contributions to Public Goods. Scandinavian Journal of Economics 102 (1), 23–39.

Kocher, M.G., Pogrebna, G., Sutter, M., 2013. Other-Regarding Preferences and Management Styles. Journal of Economic Behavior & Organization 88, 109–132.

Komai, M., Grossman, P.J., 2009. Leadership and Group Size: An Experiment. Economics Letters 105 (1), 20–22.

Konrad, K.A., Schlesinger, H., 1997. Risk Aversion in Rent-Seeking and Rent-Augmenting Games. The Economic Journal 107 (445), 1671–1683.

Kugler, T., Rapoport, A., Pazy, A., 2009. Public Good Provision in Inter-Team Conflicts: Effects of Asymmetry and Profit-Sharing Rule. Journal of Behavioral Decision Making 23 (4), 421-438.

Kurschilgen, M., Morell, A., Weisel, O., 2017. Internal Conflict, Market Uniformity, and Transparency in Price Competition between Teams. Journal of Economic Behavior & Organization 144, 121–132.

La Macchia, S.T., Louis, W.R., Hornsey, M.J., Leonardelli, G.J., 2016. In Small We Trust: Lay Theories About Small and Large Groups. Personality and Social Psychology Bulletin 42 (10), 1321–1334.

Larney, A., Rotella, A., Barclay, P., 2019. Stake Size Effects in Ultimatum Game and Dictator Game Offers: A Meta-Analysis. Organizational Behavior and Human Decision Processes 151, 61–72.

Larson, T., Capra, M., 2009. Exploiting Moral Wiggle Room: Illusory Preference for Fairness? A Comment. Judgment and decision Making 4 (6), 467–474.

Ledyard, J.O., 1995. Public Goods: A Survey of Experimental Research. In: Kagel, J.H., Roth, A.E. (Eds.). The Handbook of Experimental Economics, Princeton: Princeton University Press, 111–194.

Leibbrandt, A., Sääksvuori, L., 2012. Communication in Intergroup Conflicts. European Economic Review 56 (6), 1136–1147.

Leider, S., Rosenblat, T., Möbius, M.M., Do, Q.-A., 2010. What do we Expect from our Friends? Journal of the European Economic Association 8 (1), 120–138.

Levine, D.K., 1998. Modeling Altruism and Spitefulness in Experiments. Review of Economic Dynamics 1 (3), 593–622.

Levitt, S.D., List, J.A., 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? Journal of Economic Perspectives 21 (2), 153–174.

List, J., 2007. Field Experiments: A Bridge between Lab and Naturally Occurring Data. The B.E. Journal of Economic Analysis & Policy 5 (2).

Mago, S.D., Samak, A.C., Sheremeta, R.M., 2016. Facing your Opponents: Social Identification and Information Feedback in Contests. Journal of Conflict Resolution 60 (3), 459–481.

Maier-Rigaud, F.P., Martinsson, P., Staffiero, G., 2010. Ostracism and the Provision of a Public Good: Experimental Evidence. Journal of Economic Behavior & Organization 73 (3), 387–395.

Matthey, A., Regner, T., 2011. Do I Really Want to Know? A Cognitive Dissonance-Based Explanation of Other-Regarding Behavior. Games 2 (1), 114–135.

McCannon, B.C., 2018. Leadership and Motivation for Public Goods Contributions. Scottish Journal of Political Economy 65 (1), 68–96.

Mertins, V., Waldner, C., 2020. Evidenzbasiertes Fundraising-Management (EFM). In: . Handbuch Fundraising: Springer Gabler, Wiesbaden, 523–542.

Mertins, V., Walter, C., 2021. In Absence of Money: A Field Experiment on Volunteer Work Motivation. Experimental Economics 24 (3), 952–984.

Moxnes, E., Van der Heijden, E., 2003. The Effect of Leadership in a Public Bad Experiment. Journal of Conflict Resolution 47 (6), 773–795.

Normann, H.-T., Wallace, B., 2012. The Impact of the Termination Rule on Cooperation in a Prisoner's Dilemma Experiment. International Journal of Game Theory 41 (3), 707–718.

Nosenzo, D., Quercia, S., Sefton, M., 2015. Cooperation in Small Groups: The Effect of Group Size. Experimental Economics 18 (1), 4–14.

Obama, B., 2007. Renewing American Leadership. Foreign Affairs 86 (4), 2–16.

Palanski, M., Yammarino, F.J., 2011. Impact of Behavioral Integrity on Follower Job Performance: A Three-Study Examination. The Leadership Quarterly 22 (4), 765–786.

Peng, W., Hsieh, G., 2012. The Influence of Competition, Cooperation, and Player Relationship in a Motor Performance Centered Computer Game. Computers in Human Behavior 28 (6), 2100–2106.

Podsakoff, P.M., Bommer, W.H., Podsakoff, N.P., MacKenzie, S.B., 2006. Relationships between Leader Reward and Punishment Behavior and Subordinate Attitudes, Perceptions, and Behaviors: A Meta-Analytic Review of Existing and New Research. Organizational Behavior and Human Decision Processes 99 (2), 113–142.

Pogrebna, G., Krantz, D.H., Schade, C., Keser, C., 2011. Words Versus Actions as a Means to Influence Cooperation in Social Dilemma Situations. Theory and Decision 71 (4), 473–502.

Qiu, H., Zhang, Y., Hou, G., Wang, Z., 2018. The Integrative Effects of Leading by Example and Follower Traits in Public Goods Game: A Multilevel Study. Frontiers in Psychology 9, 1687.

Rivas, M.F., Sutter, M., 2011. The Benefits of Voluntary Leadership in Experimental Public Goods Games. Economics Letters 112 (2), 176–178.

Rothstein, B., 2000. Trust, Social Dilemmas and Collective Memories. Journal of Theoretical Politics 12 (4), 477–501.

Rousseau, D.M., 2006. Is there Such a thing as "Evidence-Based Management"? Academy of Management Review 31 (2), 256–269.

Ruffle, B.J., Tobol, Y., 2017. Clever Enough to Tell the Truth. Experimental Economics 20 (1), 130–155.

Ryan, R.M., Deci, E.L., 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. The American psychologist 55 (1), 68–78.

Sahin, S.G., Eckel, C., Komai, M., 2015. An Experimental Study of Leadership Institutions in Collective Action Games. Journal of the Economic Science Association 1 (1), 100–113.

Sánchez-Pagés, S., Vorsatz, M., 2007. An Experimental Study of Truth-Telling in a Sender–Receiver Game. Games and Economic Behavior 61 (1), 86–112.

Sánchez-Pagés, S., Vorsatz, M., 2009. Enjoy the Silence: An Experiment on Truth-Telling. Experimental Economics 12 (2), 220–241.

Schaffer, B., 2008. Leadership and motivation. Supervision 69 (2).

Schram, A., 2005. Artificiality: The Tension between Internal and External Validity in Economic Experiments. Journal of Economic Methodology 12 (2), 225–237.

Schwaiger, M., 2004. Components and Parameters of Corporate Reputation — An Empirical Study. Schmalenbach Business Review 56 (1), 46–71.

Selten, R., 1967. Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. In: Sauermann, H. (Ed.). Beiträge zur experimentellen Wirtschaftsforschung, Tübingen: Mohr, 136–168.

Sheremeta, R.M., 2013. Overbidding and Heterogenous Behavior in Contest Experiments. Journal of Economic Surveys 27 (3), 491–514.

Sheremeta, R.M., 2018a. Behavior in Group Contests: A Review of Experimental Research. Journal of Economic Surveys 32 (3), 683–704.

Sheremeta, R.M., 2018b. Impulsive Behavior in Competition: Testing Theories of Overbidding in Rent-Seeking Contests. Working Paper.

Simons, T., Leroy, H., Collewaert, V., Masschelein, S., 2015. How Leader Alignment of Words and Deeds Affects Followers: A Meta-analysis of Behavioral Integrity Research. Journal of Business Ethics 132 (4), 831–844.

Soboroff, S.D., Kelley, C.P., Lovaglia, M.J., 2020. Group Size, Commitment, Trust, and Mutual Awareness in Task Groups. The Sociological Quarterly 61 (2), 334–346.

Soetevent, A.R., 2005. Anonymity in Giving in a Natural Context—A Field Experiment in 30 churches. Journal of Public Economics 89 (11-12), 2301–2323.

Sutter, M., Rivas, F., 2014. Leadership, Reward, and Punishment in Sequential Public Goods Experiments. In: Van Lange, P.A., Rockenbach, B., Yamagishi, T. (Eds.). Reward and Punishment in Social Dilemmas, Cambridge: Cambridge University Press, 113–160.

Swope, K.J., 2002. An Experimental Investigation of Excludable Public Goods. Experimental Economics 5 (3), 209–222.

Tajfel, H., 1970. Experiments in Intergroup Discrimination. Scientific American 223 (5), 96–102.

Tajfel, H., Turner, J., Austin, W., Worchel, S., 1979. An Integrative Theory of Intergroup Conflict. In: Hatch, M.J., Schultz, M. (Eds.). Organizational Identity: A reader. Oxford management readers, Oxford, New York: Oxford University Press.

Tan, J.H., Bolle, F., 2007. Team Competition and the Public Goods Game. Economics Letters 96 (1), 133–139.

Tauer, J.M., Harackiewicz, J.M., 2004. The Effects of Cooperation and Competition on Intrinsic Motivation and Performance. Journal of Personality and Social Psychology 86 (6), 849–861.

Teyssier, S., 2012. Inequity and Risk Aversion in Sequential Public Good Games. Public Choice 151 (1-2), 91–119.

Tonin, M., Vlassopoulos, M., 2013. Experimental Evidence of Self-Image Concerns as Motivation for Giving. Journal of Economic Behavior & Organization 90, 19–27.

Tullock, G., 1980. Efficient Rent Seeking. In: Buchanan, J.M., Tollison, R.D., Tullock, G. (Eds.). Toward a Theory of the Rent-Seeking Society, 4th ed.: Texas A & M University Press, 97–112.

Tversky, A., Simonson, I., 1993. Context-Dependent Preferences. Management Science 39 (10), 1179–1189.

Walker, C.J., 2010. Experiencing Flow: Is Doing it Together Better Than Doing it Alone? The Journal of Positive Psychology 5 (1), 3–11.

Weber, R.A., 2006. Managing Growth to Achieve Efficient Coordination in Large Groups. American Economic Review 96 (1), 114–126.

Weimann, J., Brosig-Koch, J., 2019. Methods in Experimental Economics: An Introduction. Wiesbaden: Springer.

Winking, J., Mizer, N., 2013. Natural-Field Dictator Game shows No Altruistic Giving. Evolution and Human Behavior 34 (4), 288–293.

Wooden, J., Carty, J., 2009. Coach Wooden One-on-One: Inspiring Conversations on Purpose, Passion and the Pursuit of Success: Baker Publishing Group.

Yu, J., Kocher, M., 2020. Leading by Example in a Public Goods Experiment with Benefit Heterogeneity. Working Paper.

Zelmer, J., 2003. Linear Public Goods Experiments: A Meta-Analysis. Experimental Economics 6 (3), 299–310.

Zhang, Y., Zhang, L., Liu, G., Duan, J., Xu, S., Cheung, M.W.-L., 2019. How Does Ethical Leadership Impact Employee Organizational Citizenship Behavior? Zeitschrift für Psychologie 227 (1), 18–30.

Zhou, H., Fishbach, A., 2016. The Pitfall of Experimenting on the Web: How Unattended Selective Attrition leads to Surprising (yet false) Research Conclusions. Journal of Personality and Social Psychology 111 (4), 493–504.

Appendix

Table of Contents (Appendix)

Appendix A: Appendix to Chapter 2

A.I Descriptive statistics

Table A- I: Descriptive statistics (all studies)

Study and included treatments			Rel. contributions			Public Goods Environment			
	Struc.	Groups	All	Leader	Follow.	G. Size	MPCR	Endow.	Rounds
Centorrino & Concina (2013)									
X-Treatment	LbE	24	0.407	0.458	0.390	4	0.5	30	10
			(0.173)	(0.206)	(0.187)				
Dannenberg (2015)									
Ex-Base	Sim.	10	0.349	---	---	4	0.4	25	10
			(0.250)						
Ex-Leader	LbE	10	0.519	0.621	0.486	4	0.4	25	10
			(0.220)	(0.227)	(0.230)				
Drouvelis & Nosenzo (2013)									
No-ID	LbE	16	0.531	0.621	0.486	3	0.5	20	10
			(0.249)	(0.218)	(0.275)				
Eisenkopf & Kölpin (2021)[23]									
Small-Team (w/o Leader)	Sim.	20	0.384	---	---	3	0.5	100	20
			(0.322)						
Small-Team (w. Leader)	LbE	20	0.524	0.547	0.513	3	0.5	100	20
			(0.290=	(0.288)	(0.295)				
Large-Team (w/o Leader)	Sim.	9	0.773	---	---	3	0.5	100	20
			(0.166)	--					
Large-Team (w. Leader)	LbE	11	0.727	0.846	0.704	3	0.5	100	20
			(0.205)	(0.181)	(0.216)				

[23] We pooled the data from the Small-Team and Peer treatments.

Study and included treatments			Rel. contributions			Public Goods Environment			
	Struc.	Groups	All	Leader	Follow.	G. Size	MPCR	Endow.	Rounds
Eisenkopf (2020)									
Baseline	Sim.	16	0.301	---	---	3	0.5	100	20
Leading-by-Example	LbE	8	0.607	0.673	0.567	3	0.5	100	20
			(0.295)	(0.304)	(0.317)				
Eisenkopf & Walter (2021) [24]									
Baseline-IM/GM	Sim.	40	0.240	---	---	3	0.5	100	20
Lbe-IM/GM	LbE	40	0.502	0.526	0.490	3	0.5	100	20
			(0.311)	(0.304)	(0.317)				
Frackenpohl (2016)									
Give-R	LbE	18	0.573	0.664	0.542	4	0.4	20	10
			(0.331)	(0.329)	(0.338)				
Gächter & Renner (2018)									
No-Leader-Treatment	Sim.	12	0.512	---	---	4	0.4	20	10
			(0.179)						
Leader-Treatment	LbE	12	0.482	0.543	0.462	4	0.4	20	10
			(0.251)	(0.242)	(0.267)				

[24] We pooled the data from the individual and group monitoring treatments.

Study and included treatments			Rel. contributions			Public Goods Environment			
	Struc.	Groups	All	Leader	Follow.	G. Size	MPCR	Endow.	Rounds
Gürerk et al. (2018)									
Treatment P	Sim.	12	0.563	---	---	4	0.4	20	20
			(0.231)						
Treatment L	LbE	12	0.488	0.560	0.463	4	0.4	20	20
			(0.286)	(0.282)	(0.293)				
Güth et al (2007)									
C-Control	Sim.	14	0.402	---	---	4	0.4	25	16
			(0.247)						
Lf-Leader fixed	LbE	14	0.524	0.611	0.495	4	0.4	25	16
			(0.294)	(0.284)	(0.301)				
Moxnes & v. d. Heijden (2003)									
Control-Treatment	Sim.	12	0.137	---	---	5	0.4	20	10
			(0.054)						
Leader-Treatment	LbE	12	0.366	0.847	0.246	5	0.4	20	10
			(0.142)	(0.027)	(0.176)				

Study and included treatments			Rel. contributions			Public Goods Environment			
	Struc.	Groups	All	Leader	Follow.	G. Size	MPCR	Endow.	Rounds
Rivas & Sutter (2011)									
Control	Sim.	14	0.40	---	---	4	0.4	25	16
			(0.247)						
Exognoues	LbE	14	0.350	0.473	0.310	4	0.4	25	16
			(0.187)	(0.238)	(0.177)				
Sahin et al. (2015)									
Baseline-Treatment	Sim.	8	0.682	---	---	6	0.2	9	20
			(0.067)						
Exemplar-Treatment	LbE	14	0.731	0.855	0.706	6	0.2	9	20
			(0.136)	(0.124)	(0.146)				

A.II Robustness Checks

A.II.I Replication without the additional studies

This subsection includes robustness checks for the regression based on the groups with leaders. In the main part of our paper, we included additional groups from other studies that do not allow comparison between simultaneous contributions and Leading-by-Example (see Table 2-1 in section 2.4.1). In the robustness tests presented here, we replicate the estimates from the main part but excluded the additional groups from the studies Centorrino and Concina (2013), Drouvelis and Nosenzo (2013) and Frackenpohl et al. (2016). This subsection is ordered as follows: Table A- II replicates all six models from Table 2-4. Table A- III provides replications for the models IV-VI from Table 2-6. Last, not least, Table A- IV replicates all seven models from Table 2-8.

Table A- II: Replication of Table 2-4: The impact of Leading-by-Example

Dep. Var.: Indiv. contribution in percentages of the endowment	**Leading-by-Example**					
	Leader & Follower			**Only Followers**		
	I	**II**	**III**	**IV**	**V**	**VI**
Leader	0.088***	0.114***	0.113***			
	(0.013)	(0.011)	(0.011)			
Rel. Leader contribution				0.685***	0.718***	0.723***
				(0.025)	(0.026)	(0.024)
Group Size		0.090***	0.074**		0.010	-0.008
		(0.021)	(0.029)		(0.014)	(0.015)
Exchange rate (in €)		-0.359	-1.324		-1.55***	-2.45***
		(0.549)	(0.830)		(0.468)	(0.567)
MPCR		0.325*	0.653***		0.306*	0.799***
		(0.170)	(0.127)		(0.170)	(0.158)
Endowment		-0.001	-0.010*		-0.002**	-0.01***
		(0.001)	(0.005)		(0.001)	(0.003)
Total number of Rounds		0.014**	0.027***		0.021***	0.041***
		(0.006)	(0.006)		(0.005)	(0.005)
Leader not fixed	-0.17***	-0.122**	-0.170*	-0.056**	-0.026	-0.078**
	(0.040)	(0.051)	(0.089)	(0.028)	(0.033)	(0.038)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	0.545***	-0.196	-0.149	0.085***	-0.32***	-0.47***
	(0.024)	(0.143)	(0.252)	(0.015)	(0.121)	(0.178)
Observations	12,232	12,232	12,232	9,096	9,096	9,096
Groups	190	190	190	190	190	190
R-squared	0.034	0.093	0.114	0.402	0.447	0.478

*OLS-Regression. Robust standard errors in parentheses clustered at group level. *** p<0.01, ** p<0.05, * p<0.1*

Table A- III: Replication of Table 2-6: Long-term impact of leadership (model IV-VI)

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-Example I	II	III
Leader	0.0973***	0.1108***	0.1083***
	(0.0207)	(0.0185)	(0.0183)
Leader*Round	-0.0010	0.0003	0.0005
	(0.0014)	(0.0013)	(0.0013)
Round	-0.0034*	-0.0068***	-0.0068***
	(0.0018)	(0.0015)	(0.0015)
Group Size		0.0900***	0.0740**
		(0.0214)	(0.0289)
Exchange rate (in €)		-0.3595	-0.0374
		(0.5488)	(0.4983)
MPCR		0.3251*	0.6530***
		(0.1696)	(0.1268)
Endowment		-0.0005	-0.0177***
		(0.0010)	(0.0056)
Total number of Rounds		0.0176***	0.0263***
		(0.0059)	(0.0067)
Leader not fixed	-0.1841***	-0.1218**	
	(0.0393)	(0.0506)	
Fixed Effects for Studies	---	---	✓
Constant	0.5778***	-0.1915	-0.1541
	(0.0258)	(0.1435)	(0.2504)
Observations	12,232	12,232	12,232
Groups	190	190	190
R-squared	0.0361	0.1002	0.1219

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

Table A- IV: Replication of Table 2-8: The impact of group size in Leading-by-Example treatments

Dep. Var.: Indiv. contribution in percentages of the endowment	Leading-by-Example						
	Leader			Follower			
	I	II	III	IV	V	VI	VII
Group Size	0.10***	0.11***	0.10***	0.06***	-0.013	0.041**	0.07***
	(0.016)	(0.019)	(0.027)	(0.017)	(0.011)	(0.020)	(0.025)
Rel. Leader contr. (RLC)					0.70***	0.98***	1.11***
					(0.027)	(0.099)	(0.092)
Group Size * RLC						-0.1***	-0.1***
						(0.026)	(0.025)
Exchange rate (in €)		1.64***	1.926**				-2.4***
		(0.491)	(0.852)				(0.570)
MPCR		0.63***	0.77***				0.80***
		(0.219)	(0.238)				(0.157)
Endowment		0.000	-0.004				-0.0***
		(0.001)	(0.006)				(0.003)
Total number of Rounds		0.004	0.009				0.04***
		(0.006)	(0.008)				(0.005)
Leader not fixed	-0.2***	-0.2***	-0.140	-0.2***	-0.06**	-0.07**	-0.08**
	(0.047)	(0.055)	(0.095)	(0.038)	(0.028)	(0.028)	(0.038)
Fixed Effects for Studies	---	---	✓	---	---	---	✓
Constant	0.22***	-0.214	-0.262	0.26***	0.13***	-0.070	-0.7***
	(0.074)	(0.161)	(0.302)	(0.073)	(0.039)	(0.068)	(0.190)
Observations	3,112	3,112	3,112	10,700	9,096	9,096	9,096
Groups	190	190	190	190	190	190	190
R-squared	0.101	0.120	0.130	0.050	0.403	0.406	0.485

OLS-Regression. Robust standard errors in parentheses clustered at group level.
**** p<0.01, ** p<0.05, * p<0.1*

A.II.II Replication of regressions without excluded studies

In this section, we replicate the regression given in our main part of the paper, but we exclude certain studies that induce a change in the result. This section is ordered as follows: Table A-V replicates Table 2-5. Table A-VI provides replications of model I-III of Table 2-6. Table A-VII replicates model IV-VII of Table 2-8.

Table A-V: Replication of Table 2-5: Standard deviations of individual contributions within groups

Dep. Var.: Standard deviations within groups	**Excluded study: Eisenkopf & Walter (2021)**			**Excluded study: Güth et al. (2007)**		
	Both Treatments Benchmark: Simultaneous					
	I	**II**	**III**	**IV**	**V**	**VI**
Leading-by-Example	-0.740	-0.372	-0.138	-0.017	-0.03***	-0.026**
	(1.383)	(0.956)	(0.984)	(0.012)	(0.010)	(0.010)
Group Size		3.07***	3.96***		0.04***	0.04***
		(0.990)	(1.172)		(0.010)	(0.012)
Exchange rate (in €)		-23.39**	-909.820		0.205	-32.461
		(10.622)	(927.683)		(0.210)	(32.944)
MPCR		6.36***	-0.000		0.130*	0.037
		(1.808)	(.)		(0.076)	(0.091)
Endowment		0.16***	-0.599		-0.000*	-0.033
		(0.017)	(1.118)		(0.000)	(0.033)
Total number of Rounds		-0.22***	7.059		-0.003*	0.271
		(0.053)	(7.748)		(0.002)	(0.283)
Fixed Effects for Studies	---	---	✓	---	---	✓
Constant	9.00***	-10.09**	-50.636	0.19***	0.075	-1.277
	(0.919)	(4.100)	(37.794)	(0.007)	(0.064)	(1.504)
Observations	4,526	4,526	4,526	5,678	5,678	5,678
Groups	289	289	289	341	341	341
R-squared	0.000	0.100	0.110	0.003	0.134	0.139

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
**** p<0.01, ** p<0.05, * p<0.1*

Table A-VI: Replication of Table 2-6: Long-term impact of leadership (model I-III)

Dep. Var.: Indiv. contribution in percentages of the endowment	**Excluded study: Eisenkopf & Walter (2021)**		
	Both Treatments Benchmark: Simultaneous		
	I	**II**	**III**
Leading-by-Example (LbE)	0.106***	0.069**	0.060*
	(0.037)	(0.032)	(0.034)
LbE*Round	0.002	0.002	0.003
	(0.003)	(0.003)	(0.003)
Round	-0.005**	-0.011***	-0.011***
	(0.002)	(0.002)	(0.002)
Group Size		0.110***	0.092***
		(0.016)	(0.021)
Exchange rate (in €)		-1.399***	-13.143
		(0.373)	(92.071)
MPCR		0.233*	0.685***
		(0.126)	(0.103)
Endowment		-0.001**	0.029
		(0.001)	(0.032)
Total number of Rounds		0.023***	0.100
		(0.004)	(0.688)
Leader not fixed	-0.202***	-0.106**	-0.079
	(0.041)	(0.044)	(0.061)
Fixed Effects for Studies	---	---	✓
Constant	0.508***	-0.275**	-1.617
	(0.025)	(0.113)	(5.603)
Observations	18,744	18,744	18,744
Groups	289	289	289
R-squared	0.036	0.152	0.170

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
*** p<0.01, ** p<0.05, * p<0.1*

Table A-VII: Replication of Table 2-8: The impact of group size in Leading-by-Example treatments (model IV-VII)

Dep. Var.: Indiv. contribution in percentages of the endowment	**Excluded study: Sahin et al. (2015)**			
	Leading-by-Example Treatments Only Followers			
	I	**II**	**III**	**IV**
Group Size	0.0397 (0.0249)	-0.0301* (0.0170)	0.0333 (0.0227)	0.0757** (0.0317)
Rel. Leader contribution (RLC)		0.6921*** (0.0248)	1.0097*** (0.1314)	1.1233*** (0.1214)
RLC*Group Size			-0.0835** (0.0360)	-0.1083*** (0.0340)
Exchange (in €)				-2.3494*** (0.5725)
MPCR				0.7901*** (0.1567)
Endowment				0.0010 (0.0061)
Total number of Rounds				0.0305*** (0.0064)
Leader not fixed	-0.1427*** (0.0391)	-0.0459 (0.0284)	-0.0570** (0.0282)	-0.0824** (0.0391)
Fixed Effects for Studies	---	---	---	✓
Constant	0.3445*** (0.0974)	0.1920*** (0.0613)	-0.0425 (0.0788)	-0.9082*** (0.2291)
Observations	9,300	9,276	9,276	9,276
Groups	234	234	234	234
R-squared	0.0234	0.4060	0.4098	0.4810

OLS-Regression. Robust standard errors in parentheses clustered at Group-Level
**** p<0.01, ** p<0.05, * p<0.1*

Appendix B: Appendix to Chapter 3

B.I Participant instructions (all treatments)[25]

We warmly welcome you to this economic experiment.

In this experiment, your decisions and, if appropriate, the decisions of other participants influence your payouts. It is therefore very important that you read these explanations carefully. If you have any questions, please contact us before starting the experiment.

During the experiment, you are not allowed to speak with the other participants.

Failure to comply with this rule will result in exclusion from the experiment and all payments.

In today's experiment, we are carrying out various treatments. You will now receive information that is valid for all participants in the same way. You will later receive instructions on your specific treatment on the screen.

During the experiment, we are not talking about EUR but about points. Your total income is therefore calculated first in points. The total score you obtained during the experiment will then be converted into euros at the end of the experiment using the following conversion rate:

200 Points = 1 Euro

At the end of today's experiment, you will receive your earnings in cash. The following pages explain the exact procedure of the experiment.

Structure of the experiment

In this experiment you play in a group with six participants (participants 1-6) who all receive the same instructions. These are matched into one or two teams depending on the treatment.

The experiment consists of 20 rounds. In each round, each team member receives 100 points credited to a personal account. You can keep these 100 points in your personal account, or you can use them in whole or in part. Any point you do not use will automatically remain on your personal account. The points used by each team member are added. You earn 50% of the cumulated points. The amount you spend each team member earns in the same way. Conversely, you also earn something from the input of the other team member. The income of each member of the project is determined as follows:

Earnings from own income = Sum of the income of all team members

Imagine you are in a team with three participants. For example, if the total input made by the three members is 50 points, then you and each other team member receive 50 × 0.5 = 25 points

[25] Translated from German

each. Now imagine you are in a team with six participants. Again, if the total input made by the six members is 50 points, then you and each other team member receive 50 × 0.5 = 25 points each. Your total income is calculated by adding up your income from the personal account and your income from the assignment. So:

Income from the private account (=100 – use)

+ Income from use (= 0.5 * sum of all team member input)

= total income

At the end of each round, team members learn how much they have earned in this round.

Examples for determining the payout.

The numerical values for the following examples are arbitrary:

Example 1:

Participant 1 invests 10 points, participant 2 50 points and participant 3 30 points. The total deployment of the team is therefore 90 points. Each team member receives 45 points. Thus, the individual payments amount

Participant 1: 100 + 45 - 10 = 135 points

Participant 2: 100 + 45 - 50 = 95 points

Participant 3: 100 + 45 - 30 = 115 point

Example 2:

Participant 1 invests 80 points, participant 2 70 points, participant 3 50 points, participant 4 0 points, participant 5 100 points and participant 6 100 points. The accumulated use of the team is thus 400 points. Each team member gets 200 points. Thus, the individual payouts are:

Participant 1: 100 + 200 - 80 = 220 points

Participant 2: 100 + 200 – 70= 230 points

Participant 3: 100 + 200 – 50= 250 points

Participant 4: 100 + 200 – 0 = 250 points

Participant 5: 100 + 200 – 100 = 250 points

Participant 6: 100 + 200 – 100 = 250 points

B.II Treatment Instructions[26]

Small-Team-Treatment and Peer-Treatment

Participants 1, 3 and 5 are in one team, participants 2, 4 and 6 in another team. The team composition does not change between the rounds. Participant 6 has a special role. This participant makes his decision before the other participants. The other team members – but not the members from the other team – can observe the decision of participant 6 when they decide about their own.

Large-Team-Treatment (without leadership)

All six participants are in one team. The team composition does not change between the rounds.

Large-Team-Treatment (with leadership)

All six participants are in one team. The team composition does not change between the rounds. Participant 6 has a special role in each team. This participant makes his decision before the other participants. The other team members can observe the decision of participant 6 when they decide about their own.

Open-Change-Treatment (without leadership)

In the first round, participants 1, 3 and 5 are in one team, participants 2, 4 and 6 in another team. At the beginning of each subsequent round, participants 1- 6 can decide about leaving one team for the other.

Open-Change-Treatment (with leadership)

In the first round, participants 1, 3 and 5 are in one team, participants 2, 4 and 6 in another team. Participant 6 has a special role in each team. This participant makes his decision before the other participants. The other team members – but not the members from the other team – can observe the decision of participant 6 when they decide about their own.
At the beginning of each subsequent round, participants 1- 5 can decide about leaving one team for the other. Participant 6 always remains in the same team.

[26] Translated from German

Restricted-Change-Treatment

In the first round, participants 1, 3 and 5 are in one team, participants 2, 4 and 6 in another team. Participant 6 has a special role in each team. This participant makes his decision before the other participants. The other team members – but not the members from the other team – can observe the decision of participant 6 when they decide about their own.

At the beginning of each subsequent round, participants 1- 5 can decide about leaving one team for the other. Participant 6 always remains in the same team.

At the end of a round, the members of one team can decide which participants from the other team they do not want to accept. Only if more than half of the current members of the receiving team accepted to the change, a participant can join that team. Otherwise, she had to remain in her old team.

B.III Spillover effects in the PEER Treatment

Our finding in section 3.4.4 shows that Leading-by-Example does not enhance contributions in comparison with simultaneous decision when the contributions of all matching group members are revealed. Behavioral spillovers (or peer effects) may explain the similarity to some extent. Table B-I provides results from a GLS-regression for both teams. The dependent variable is a subject's contribution to the club good, with standard errors clustered at the level of the matching group. The key explanatory variables are the average contribution in the own and other team in the round before. We observe a significant positive spillover from the team with leader to the team without (model I). The reverse spillover effect is insignificant (model IV). However, the results are not robust for all meaningful model specification (model III and model VI).

Table B-I: Spillover effects in PEER Treatment

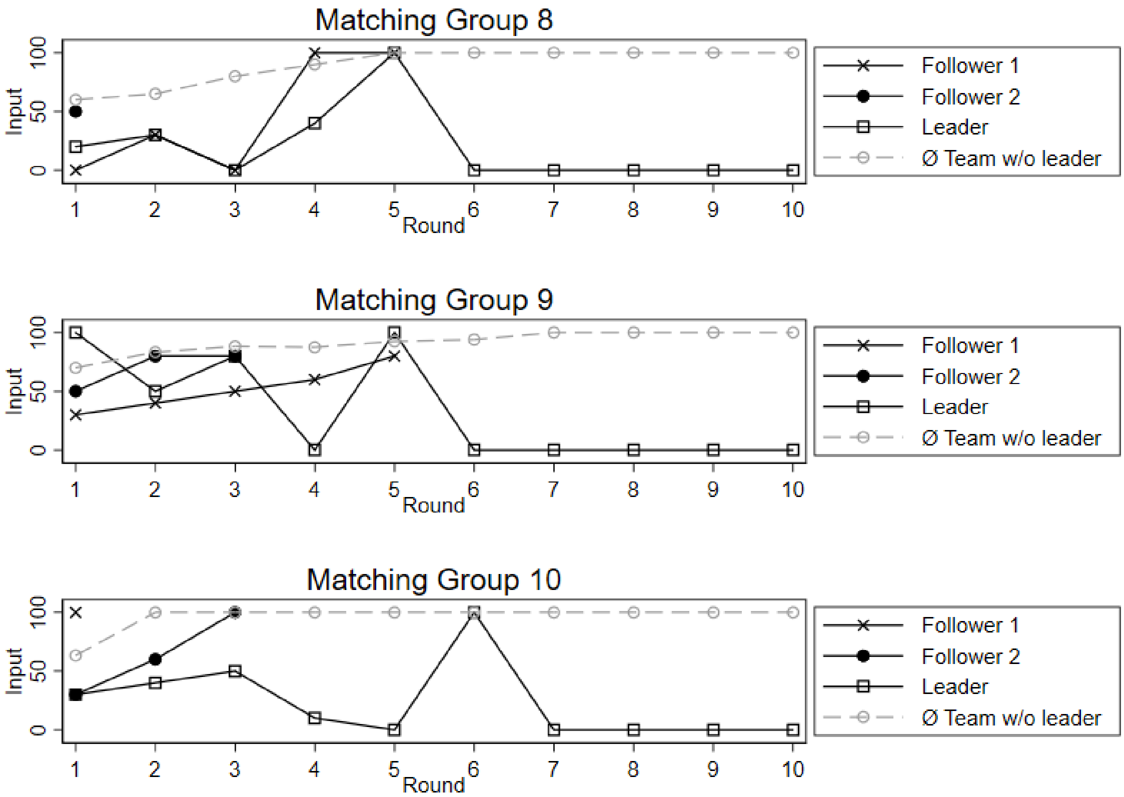
Dep. Variable: Per capita Contribution to club good	**Team 1 (w./o. leader)**			**Team 2 (w. leader)**		
	I	**II**	**III**	**IV**	**V**	**VI**
$\bar{x}_{Team1}$ in t-1		0.96***	0.95***	0.111		0.0185
		(0.0163)	(0.0203)	(0.154)		(0.0571)
$\bar{x}_{Team2}$ in t-1	0.17***		0.0289		0.84***	0.83***
	(0.0574)		(0.0277)		(0.0581)	(0.0627)
Round	-1.113*	-0.372	-0.409	0.997	-0.0571	-0.0423
	(0.608)	(0.266)	(0.255)	(0.755)	(0.186)	(0.190)
Constant	43.59***	4.701	4.008	29.75**	7.808**	7.070*
	(10.79)	(2.974)	(2.951)	(13.37)	(3.750)	(3.650)
Observations	627	627	627	627	627	627
Subjects	33	33	33	33	33	33
Matching Groups	11	11	11	11	11	11
R-Squared	0.03	0.82	0.82	0.03	0.66	0.66

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

B.IV Teams in which participants move into the leaderless team

Figure B-I displays the contributions of the team members in the three matching groups in which participants move into the leaderless team. The figure also shows the average contributions of the leaderless team. In the first rounds, the average contribution in the leaderless team is higher than in the team with a leader in these all three matching groups. Besides, in matching group 8 and 10 we find quite low and inconstant contributions of the leaders in the first rounds. These findings help explaining the migration of the followers in the first rounds.

Figure B-I: Per capita contributions in teams in which participants move into the leaderless team



B.V Estimations including all matching groups in the Change-Treatments

As in section 3.4.3 shown, in three out of twelve matching groups in *the RESTRICTED-CHANGE Treatment*, the participants move into the leaderless team. We excluded these matching groups from our subsequent analysis in the main part of our paper. The provided tables in this section provide replicates of the regressions in section 3.4.3 including all matching groups. While Table B-II is a replication of Table 3-6, Table B-III replicates Table 3-8.

Table B-II: Per capita contributions in the Change-Treatments (including all matching groups)

Dep. Var.: Per Capita Contribution to club good	**Benchmark: OPEN-CHANGE (w Leader)**				
	All Change Treatments		**Only Change-Treatments with a leader**		
			Leader	**Follower**	
	I	**II**	**III**	**IV**	**V**
Open-Change (w/o Leader)	-0.359	6.878			
	(10.24)	(15.25)			
Restricted-Change	17.33*	6.769	11.59	19.26*	10.87
	(9.774)	(15.97)	(8.369)	(10.24)	(7.616)
Team Size	8.692***	8.506***	11.78***	9.133***	4.391**
	(1.257)	(2.420)	(1.739)	(1.949)	(1.708)
Open-Change (w/o L)*Team Size		-1.289			
		(3.182)			
Restricted -Change *Team Size		2.077			
		(3.532)			
Leader contribution		0.523***			
		(0.0834)			
Round	-1.297***	-1.306***	-1.017***	-1.648***	-1.295***
	(0.268)	(0.265)	(0.375)	(0.397)	(0.307)
Constant	29.73***	30.90***	19.77**	29.20***	12.85
	(8.809)	(10.59)	(9.633)	(10.06)	(9.845)
Observations	3,419	3,419	460	1,766	1,766
Subjects	195	195	23	106	106
Matching Groups	34	34	23	23	23
R-Squared	0.102	0.106	0.448	0.090	0.374

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

Table B-III: Per capita contributions across Treatments in teams with a leader (including all matching groups)

Dep. Var.: Per Capita Contribution to club good	**Only teams with a leader Benchmark: RESTRICTED-CHANGE**			
	Leader	**Follower**		**All**
	I	**II**	**III**	**IV**
Small-Team (w L)	-4.018	7.956	12.03	3.814
	(11.66)	(5.840)	(8.060)	(9.660)
Large-Team (w L)	-4.083	-7.651	13.72	-10.97
	(7.454)	(5.751)	(9.009)	(7.831)
Open-Change (w L)	-11.14	-9.021	15.12*	-16.67*
	(8.803)	(7.104)	(8.147)	(9.848)
Leader contribution		0.560***	0.739***	
		(0.0563)	(0.0412)	
Small-Team*Leader contribution			0.00831	
			(0.118)	
Large-Team*Leader contribution			-0.250***	
			(0.0940)	
Open-Change*Leader contribution			-0.294***	
			(0.0872)	
Round	-0.00933	-0.826***	-0.854***	-0.690**
	(0.345)	(0.185)	(0.186)	(0.275)
Dummies for team size are included	✓	✓	✓	✓
Constant	21.91**	48.58***	20.56***	29.35***
	(10.87)	(7.000)	(5.733)	(9.853)
Observations	860	3,226	3,226	4,086
Subjects	43	179	179	222
Matching Groups	43	43	43	43
R-Squared	0.302	0.346	0.334	0.078

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

B.VI Additional regressions

Table B- IV replicates the estimates from Table 3-8 with a more selective dataset. More precisely, we include only teams with six team members. In our design, teams with six members are able to generate the highest social benefits. Thus, the selective data set provides information on the effects of different admission schemes for a given size of teams. However, the results do not differ significantly from our model in the main part of the paper. Only the coefficient for the *Open-Change (w Leader)* variable in model II becomes slightly insignificant (p = .104), but continues to show the same trend as in our main section.

Table B- IV: Per capita contributions across Treatments in teams with six members and a leader

Dep. Var.: Per Capita Contribution to club good	**Only teams with a leader Benchmark: RESTRICTED-CHANGE**			
	All	**Leader**	**Follower**	
	I	**II**	**III**	**IV**
Large-Team (w Leader)	-13.26 (8.289)	-9.298 (7.428)	-9.421 (6.187)	18.01** (8.542)
Open-Change (w Leader)	-19.06* (11.17)	-15.51 (9.544)	-12.01 (8.383)	21.34*** (8.201)
Leader contribution			0.488*** (0.0586)	0.782*** (0.0398)
Large-Team*Leader contribution				-0.295*** (0.0939)
Open-Change*Leader contribution				-0.372*** (0.0717)
Round	-0.862*** (0.281)	-0.286 (0.298)	-0.837*** (0.191)	-0.817*** (0.188)
Constant	95.04*** (5.832)	96.85*** (5.780)	47.33*** (7.389)	19.81*** (5.050)
Observations	2,910	485	2,425	2,425
Subjects	180	30	150	150
Matching Groups	30	30	30	30
R-Squared	0.031	0.024	0.303	0.294

*Note: Random-effects GLS-Regression. Robust standard errors in parentheses clustered at the matching group level. *** p<0.01, ** p<0.05, * p<0.1*

Appendix C: Appendix to Chapter 4

C.I Participant instructions (all treatments)[27]

Instructions (page 1)

We warmly welcome you to this economic experiment. During the experiment, we are not talking about euros but about points. Your total income is therefore calculated first in points. The total score you obtained during the experiment will then be converted into euros at the end of the experiment using the following conversion rate:

1 Point = 0.15 Euro

Notes on payout

You will automatically receive 20 (30 in Osnabrueck) points for participating in the experiment when you reach the next page. If the time (top right) runs out before you have clicked on the "Understood" button, you will not receive any points and may no longer participate in the experiment. Your further payoff depends on your decisions and the decisions of the other participants. You cannot return to previous pages. It is therefore very important that you read the explanations on the following pages carefully.

At the end of the experiment, all points in your private account (shown at the end of the page) will be paid out. The payment is made by bank transfer. After the experiment you will receive another e-mail asking you to enter your account details. The data from the experiment will not be linked to your bank account details. Your information is only needed for the purpose of the bank transfer.

In your invitation email you have received a personal code.

Please enter the code. Inserting is not possible for technical reasons. Numbers cannot be entered via the NumPad (right side of the keyboard). Use the numbers on the main keyboard (above the letters). To make entries, please click in the input field first. Afterwards you can start with the input. We need the code to be able to assign your payment.

Please enter your private code:

Detail of the experiment (page 2)

[27] Translated from German

The experiment consists of several rounds, the exact number of rounds depends on your decisions. Your group consists of maximum 4 participants. Your decision may or may not affect your group members, depending on your treatment. All group members receive the same instructions. In each round you will interact with the same participants, but the group size may decrease.

How can the group size decrease and when does the experiment end?

You can leave the experiment at any time. To do so, click on the "Exit" button (bottom left). You can also exit the experiment by letting the time (top right) run out. So, make your decisions within the time limit! You cannot re-enter the experiment after you have left. The experiment is finished for you. The remaining participants continue until all participants have left the experiment.

What do you have to do?

At the beginning of each round, each group member must decide whether they want to solve a table with zeros and ones. The task is to count the zeros. Solving the table involves a cost that increases by 1 point with each round. In the first round the table is costless. This means that you must pay 0 points in the first round, 1 point in the second, 2 points in the third and so on. Every group member who does not want to solve a table automatically quits the experiment.

(Only Control Treatment) Each group member who solves the table increases his private account by 10 points.

(Only Simultaneous Treatment) Each group member who solves the table may then increase his private account by 10 points or the private accounts of all group members by 6 points.

(Only Leading-by-Example Treatment) Each group member who solves the table may then increase his private account by 10 points or the private accounts of all group members by 6 points. **Participant 1 is in a special role. This participant makes his decision before the other participants. The other group members can observe the decision of participant 1 when they decide about their own.**

Decision Stage

You decide now whether you want to solve a table with zeros and ones. The task is to count the zeros. Solving the table involves costs that increase by 1 point with each round.

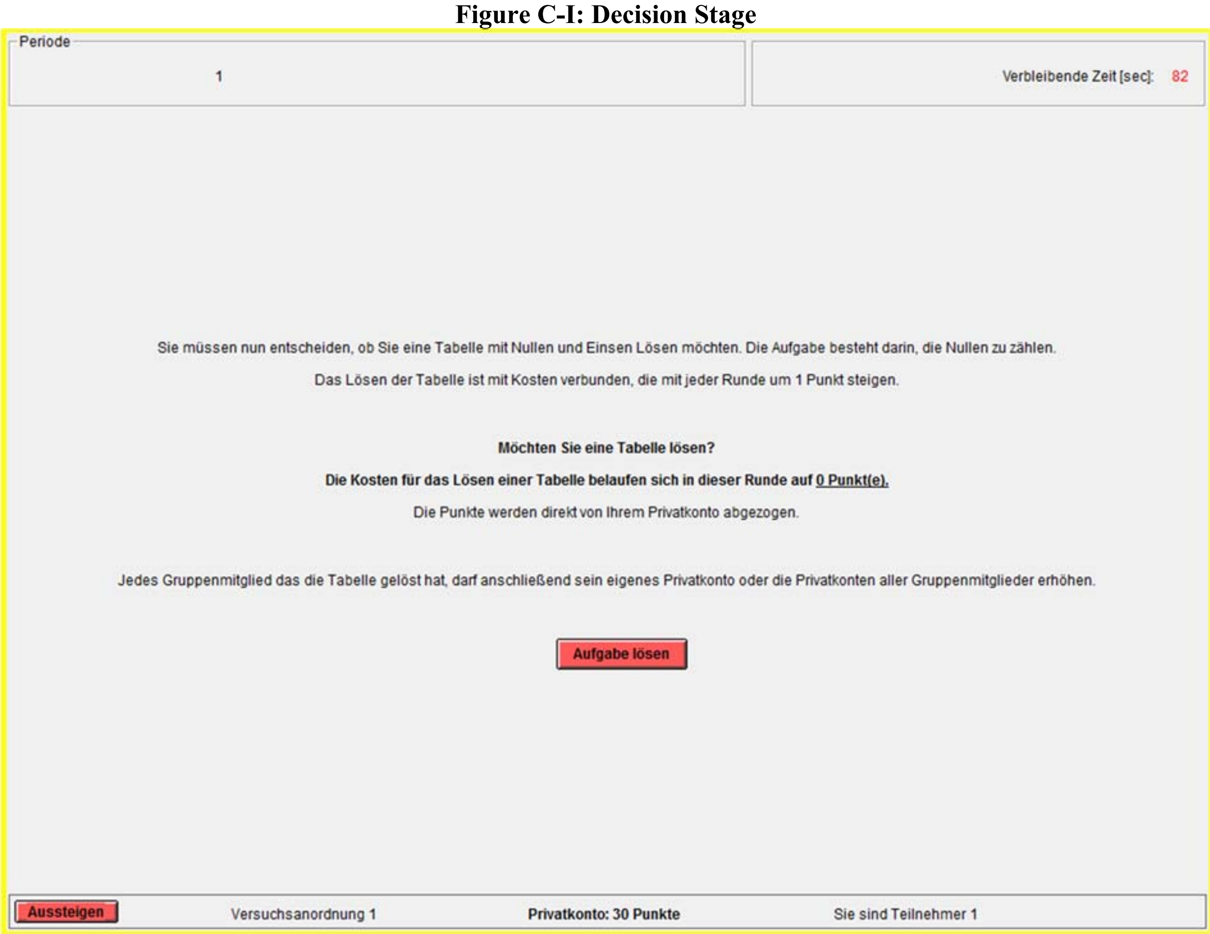
Do you want to solve a table?

The cost of solving a table in this round is 0 points.

The points are deducted directly from your private account.

Each group member who solves the table may increase his own private account or all private accounts.

Figure C-I: Decision Stage

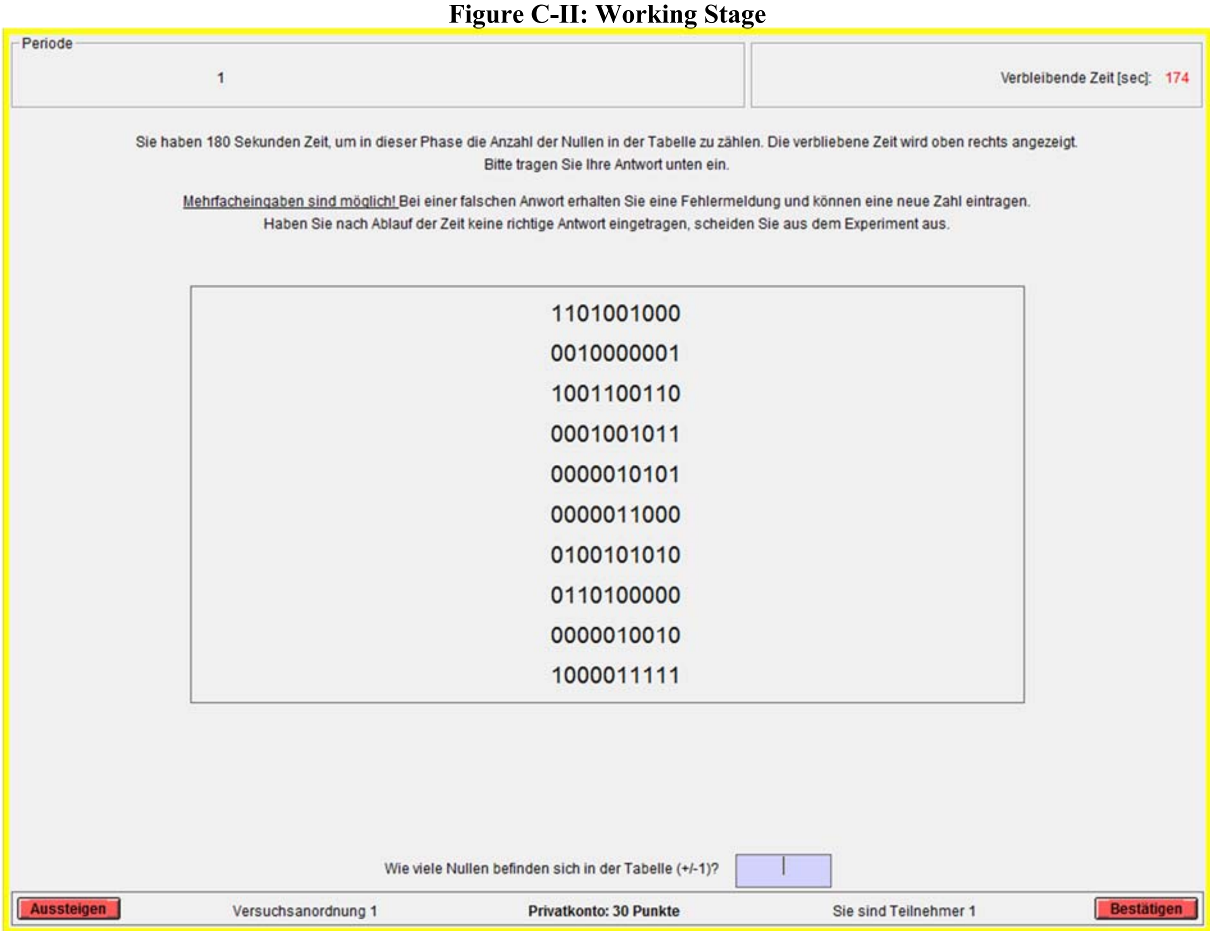


Working Stage

You have 180 seconds to count the number of zeros in the table at this stage. The time remaining is displayed in the upper right corner. Please enter your answer below.

Multiple entries are possible! If the answer is wrong, you will receive an error message and can enter a new number. If you have not entered a correct answer after the time has expired, you will be excluded from the experiment.

Figure C-II: Working Stage



Contribution Stage

~~Participant 1 left the experiment.~~

Participant 2 solved the task.

Participant 3 solved the task.

Participant 4 solved the task.

You have solved the task correctly and may make the following decision:

Decision situation:

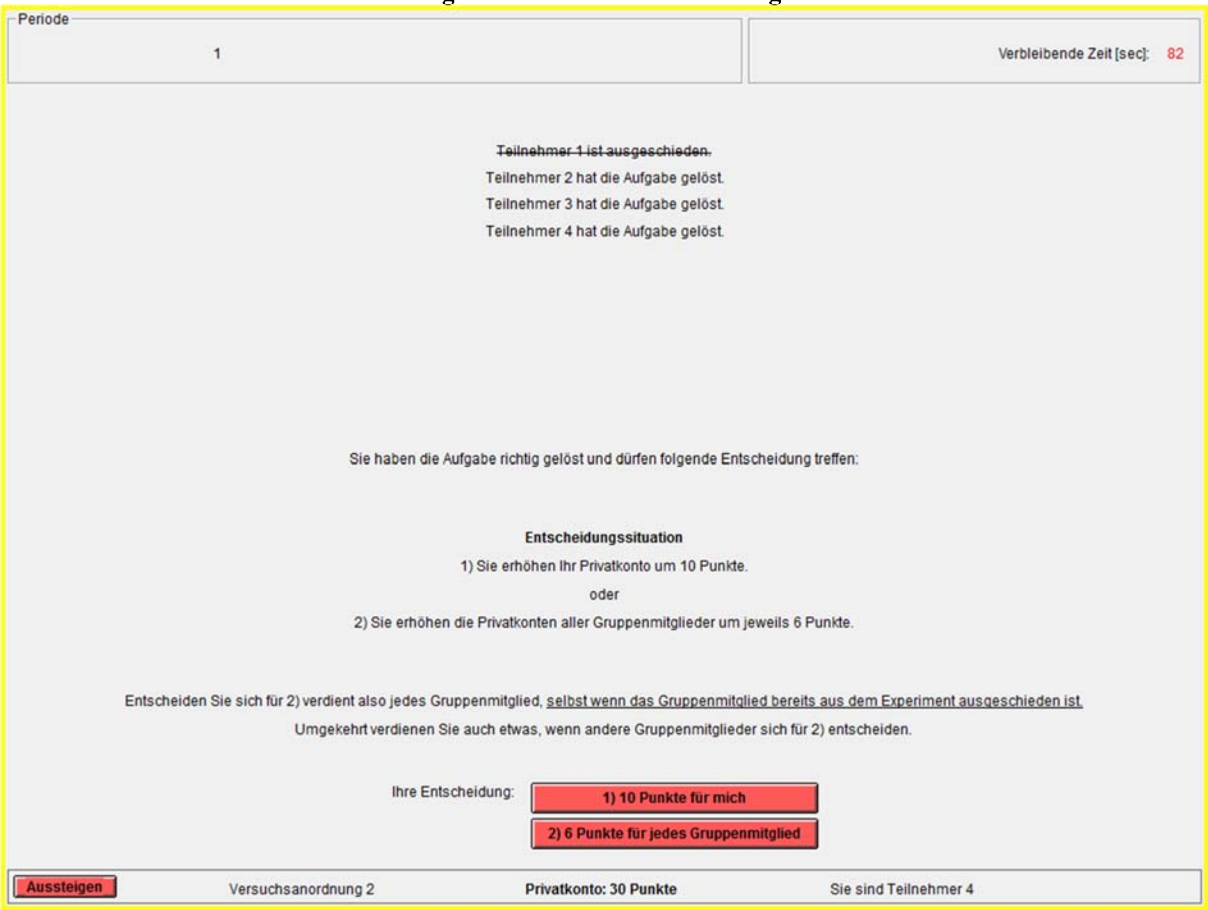
1) You increase your private account by 10 points

or

2) You increase the private accounts of all group members by 6 points

If you decide for 2) each group member earns money, even if the group member has already left the experiment. Conversely, you will also earn something if other group members decide in favor of 2).

Figure C-III: Contribution Stage



Result Stage

Results:

~~Participant 1 left the experiment.~~

Participant 2 increased the private accounts of all group members by 6 points.

Participant 3 increased her private account by 10 points.

Participant 4 increased the private accounts of all group members by 6 points.

Your payoff in this round:

You have received 0 points from participant 1.

You have received 6 points from participant 2.

You have received 0 points from participant 3.

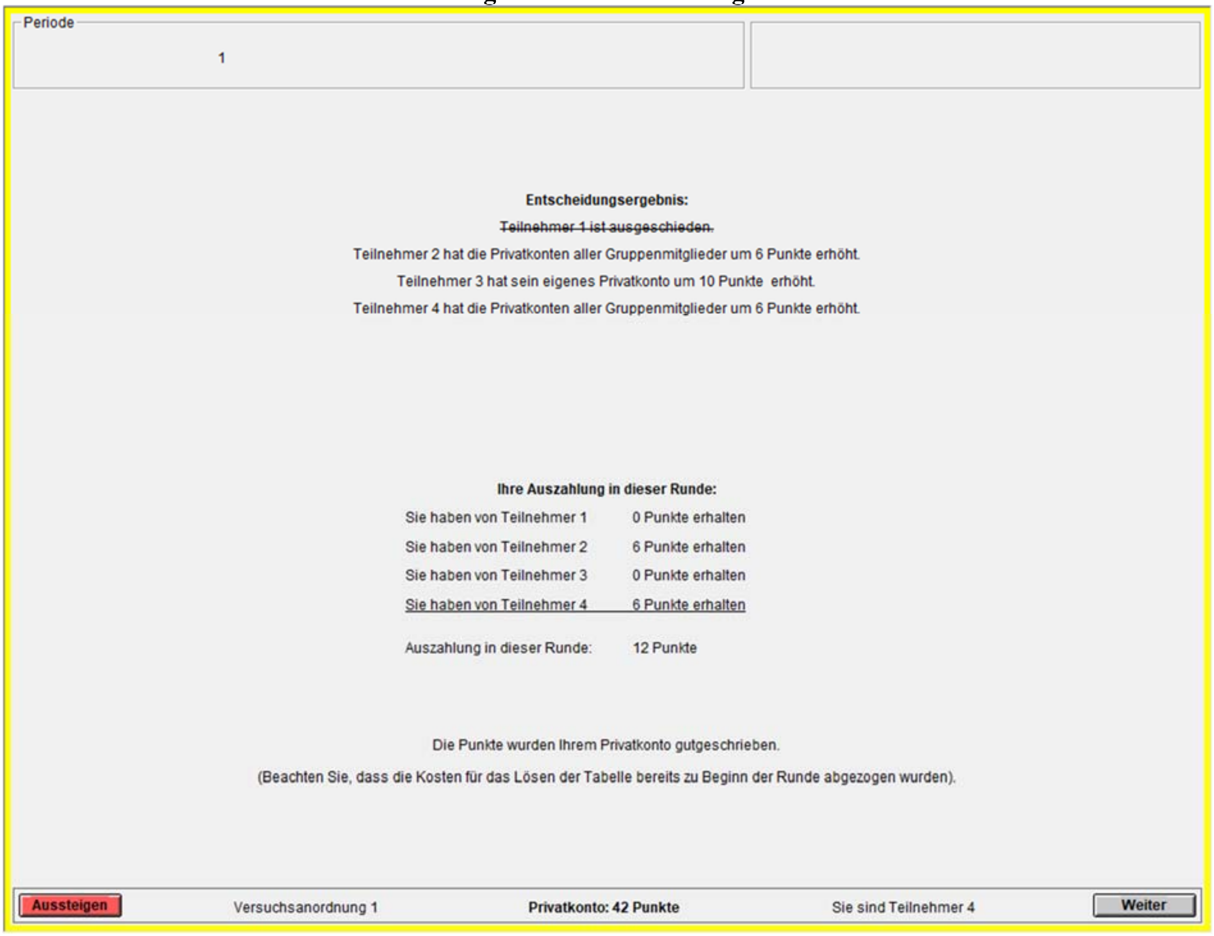
You have received 6 points from participant 4.

Your payoff in this round: 12 points

The points were credited to your private account.

(Note that the cost of solving the table was already deducted at the beginning of the round).

Figure C-IV: Result Stage



C.II Reasons for exclusion

Table C-I summarizes the number and reasons for exclusion. A total of 364 participants registered for participation. However, we observe substantial attrition (n = 52) already before the experiment started. This participant did not show up for the experiment. This leaves us with 312 observations. Moreover, we exclude three groups (8 participants) in the *LEADING-BY-EXAMPLE Treatment* because the leader did not appear. In total, we use 304 observations for our subsequent analysis.

Table C-I: Number and reasons for exclusion

	NO-COOP-ERATION	**COOP-ERATION**	**LEADING-BY-EXAMPLE**	**Total**
Agreed to participate	120	120	124	364
Not showed up	15	21	16	52
Leader has not connected	---	---	8	8
Observations	105	99	100	304

Note: The table shows how many observations were excluded and gives the respective reasons.

C.III Descriptive statistics

Table C- II: Descriptive Statistics for all initial group sizes

	NO-COOPERATION				COOPERATION				LEADING-BY-EXAMPLE			
Initial Group Size	2	3	4	All	2	3	4	All	2	3	4	All
Subjects	4	33	68	105	8	39	52	99	2	30	68	100
Groups	2	11	17	30	4	13	13	30	1	10	17	28
Played Rounds	8.25	11.33	9.65	10.17	9.88	10.82	12.10	11.21	10.00	11.90	12.46	12.14
	(0.35)	(2.93)	(1.11)	(2.14)	(3.40)	(3.75)	(3.67)	(3.64)	(0.0)	(3.66)	(4.42)	(4.04)
Subject’s Cooperation Rate	---	---	---	---	0.56	0.56	0.67	0.61	0.6	0.65	0.70	0.68
					(0.30)	(0.37)	(0.21)	(0.29)	(0.0)	(0.27)	(0.29)	(0.27)

Note: The table presents the average played rounds within a group as independent observation. Standard deviations in parentheses.

C.IV Further Regressions

Table C-III: Individual's decision of cooperation in the LEADING-BY-EXAMPLE Treatment – All rounds – (further regressions)

Dep. Var.: Indiv. decision of cooperation	LEADING-BY-EXAMPLE Treatment All rounds					
	Follower				Leader	
	I	II	III	IV	V	VI
Leader's cooperation	1.354***					
	(0.276)					
Coop Others (t-1)				2.205***	2.081***	
				(0.193)	(0.266)	
Leader Exit	0.114	-0.733**				
	(0.351)	(0.348)				
Follower Exit	-0.522*		-0.73***			-0.749**
	(0.292)		(0.258)			(0.296)
Initial Group Size	0.352*	0.345	0.386*	0.202	-0.184	0.0251
	(0.207)	(0.213)	(0.217)	(0.129)	(0.301)	(0.279)
Round	0.0220	-0.00389	-0.00159	0.00371	-0.0193	0.00867
	(0.0234)	(0.0284)	(0.0283)	(0.0176)	(0.0256)	(0.0268)
Constant	-1.760**	-0.662	-0.818	-1.67***	0.145	0.609
	(0.779)	(0.832)	(0.844)	(0.535)	(0.932)	(0.954)
Observations	888	888	888	817	307	335
Subjects	71	71	71	70	28	28
Groups	28	28	28	28	28	28
Pseudo-R-Squared	0.194	0.025	0.029	0.283	0.245	0.019

*Probit Regression for likelihood of cooperation. Robust standard errors in parentheses. Clustered at Session-level. * p<0.1; ** p<0.05; *** p<0.01. Note: One subject solved the task but left the experiment during the first round without a contribution. This subject is omitted in the models I-IV. In addition, one more subject dropped out at the end of the first round. This subject is omitted in model IV.*

Table C-IV: Individual's decision of cooperation in the LEADING-BY-EXAMPLE Treatment – rounds 1-9 – (further regressions)

Dep. Var.: Indiv. decision of cooperation	**LEADING-BY-EXAMPLE Treatment Round 1- 9**					
	Follower				**Leader**	
	I	**II**	**III**	**IV**	**V**	**VI**
Leader's cooperation	1.262***					
	(0.269)					
Coop Others (t-1)				2.108***	1.913***	
				(0.229)	(0.343)	
Leader Exit	0.225	-0.648				
	(0.583)	(0.521)				
Follower Exit	-0.394		-0.79***			-1.002**
	(0.402)		(0.270)			(0.511)
Initial Group Size	0.242	0.132	0.146	0.0933	-0.279	-0.112
	(0.238)	(0.262)	(0.263)	(0.156)	(0.270)	(0.283)
Round	0.0260	-0.0151	-0.0151	0.00785	-0.0579	-0.060**
	(0.0275)	(0.0236)	(0.0234)	(0.0261)	(0.0397)	(0.0256)
Constant	-1.337	0.147	0.107	-1.243*	0.793	1.430
	(0.866)	(0.937)	(0.951)	(0.635)	(0.897)	(1.022)
Observations	667	667	667	596	238	266
Subjects	71	71	71	70	28	28
Groups	28	28	28	28	28	28
Pseudo-R-Squared	0.151	0.006	0.011	0.244	0.232	0.032

*Probit Regression for likelihood of cooperation. Robust standard errors in parentheses. Clustered at Session-level. * p<0.1; ** p<0.05; *** p<0.01. Note: One subject solved the task but left the experiment during the first round without a contribution. This subject is omitted in the models I-IV. In addition, one more subject dropped out at the end of the first round. This subject is omitted in model IV.*

Table C-V: Determinants of attrition (all rounds) – Further regressions

Dep. Var: Remaining (0) or Dropping out (1) after round t	**All treatments Benchmark: NO-COOPERATION**	**COOPERATION AND LEADING-BY-EXAMPLE Benchmark: COOPERATION**		
	I	**II**	**III**	**IV**
Leading-by-Example (LbE)	-0.724**	-0.223	0.0126	-0.0372
	(0.324)	(0.297)	(0.339)	(0.285)
Cooperation Treatment (CT)	-0.199			
	(0.204)			
Coop. rate others			-1.605***	-1.161***
			(0.262)	(0.214)
Own Coop.			-0.186	-0.306
			(0.223)	(0.199)
Member Exit	1.594***			1.165***
	(0.275)			(0.211)
LbE*Member Exit	0.333			
	(0.484)			
CT*Member Exit	-0.535			
	(0.375)			
Initial Group Size	-0.0904	-0.179	0.147	0.00196
	(0.222)	(0.256)	(0.274)	(0.238)
Round	0.160***	0.159***	0.128***	0.127***
	(0.0188)	(0.0224)	(0.0235)	(0.0208)
Constant	-3.415***	-3.130***	-3.25***	-0.306
	(0.846)	(0.821)	(0.86)	(0.199)
Observations	3,417	2,354	2,351	2,351
Subjects	304	199	196	196
Groups	88	58	58	58

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: Three subjects completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. These three subjects are not included in the models III-IV.*

Table C-VI: Determinants of attrition in both cooperation treatments (round 1-9)

Dep. Var: Remaining (0) or Dropping out (1) after round t	LbE and Coop. Benchmark: COOPERATION			Only LEADING-BY-EXAMPLE			Only COOP.
	I	II	III	IV	V	VI	VII
Leading-by-Example (LbE)	-0.335	-0.0097	-0.385				
	(0.483)	(0.633)	(0.488)				
Coop. Rate Others (CRO)	-0.897*	-0.614		-1.298*			
	(0.478)	(0.516)		(0.682)			
LbE*CRO		-0.644					
		(0.771)					
Own Coop.						-0.897*	-1.02**
						(0.543)	(0.437)
Member Exit			0.262		1.74**		
			(0.982)		(0.813)		
Initial Group Size	-0.0217	-0.0487	-0.206	0.403	0.144	0.364	-0.192
	(0.183)	(0.178)	(0.213)	(0.412)	(0.406)	(0.459)	(0.246)
Round	0.3***	0.3***	0.2***	0.23**	0.17	0.24**	0.3***
	(0.0638)	(0.0649)	(0.0763)	(0.110)	(0.103)	(0.113)	(0.0841)
Constant	-4.6***	-4.7***	-4.0***	-6.0***	-5.5***	-6.1***	-4.2***
	(0.821)	(0.834)	(1.038)	(1.929)	(1.739)	(2.067)	(0.839)
Observations	1,672	1,672	1,675	850	851	850	822
Subjects	196	196	199	99	100	99	97
Groups	58	58	58	28	28	28	30

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: Three subjects completed the task but left the experiment during the first round without a contribution or having observed the contributions*

Table C-VII: Determinants of attrition (round 1-9) in the LEADING-BY-EXAMPLE Treatment – Further regressions

Dep. Var: Remaining (0) or Dropping out (1) after round t	Leader I	Leader II	Leader III	Follower IV	Follower V	Follower VI	Follower VII
Coop. rate others	-1.264			-1.345*			
	(1.251)			(0.660)			
Own Coop.		-1.409			-0.689		
		(1.052)			(0.605)		
Member Exit			3.18***			0.859	
			(1.100)			(0.739)	
Leader Participates							-2.0***
							(0.624)
Initial Group Size	0.173	0.00604	-0.0928	0.617	0.621	0.216	0.218
	(0.675)	(0.712)	(0.777)	(0.600)	(0.638)	(0.572)	(0.593)
Round	0.42***	0.40***	0.37***	0.161	0.184	0.108	0.0532
	(0.0881)	(0.0841)	(0.0940)	(0.132)	(0.138)	(0.141)	(0.150)
Constant	-6.31**	-5.497*	-6.27**	-6.40**	-6.92**	-5.34**	-3.196
	(2.869)	(2.969)	(2.778)	(2.885)	(3.069)	(2.685)	(3.259)
Observations	243	243	243	607	607	608	608
Subjects	28	28	28	71	71	71	71
Groups	28	28	28	28	28	28	28

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: One subject completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. This subject is omitted in the models IV-VII.*

Table C-VIII: Determinants of attrition (round 1-9) – Impact of Leadership (further regressions)

Dep. Var: Remaining (0) or Dropping out (1) after round t	Only Followers I	II	III	IV	V	VI
Coop. Leader (t-1)	-1.4**	-1.4**	-0.922			
	(0.572)	(0.543)	(0.740)			
Leader has not coop. in the last two rounds				1.38**	1.34**	1.019
				(0.618)	(0.590)	(0.715)
Member Exit		0.825			0.619	
		(0.608)			(0.625)	
Coop. Follower (in t)			-1.726			-1.357
			(1.392)			(1.098)
Initial Group Size	1.067	1.091	1.364*	0.0744	0.0858	0.679
	(0.789)	(0.786)	(0.767)	(0.568)	(0.572)	(0.680)
Round	0.186	0.180	0.181	0.0555	0.0501	0.118
	(0.157)	(0.154)	(0.167)	(0.146)	(0.145)	(0.145)
Constant	-8.4**	-8.5**	-9.1**	-4.85*	-4.88*	-7.0**
	(3.660)	(3.617)	(3.809)	(2.634)	(2.656)	(3.024)
Observations	536	536	536	608	608	607
Subjects	70	70	70	72	72	72
Groups	28	28	28	28	28	28

*Values reflect estimates from proportional hazards models fitted to binary events of participants staying (0) or dropping out (1) in a given round of the session, conditional on not having dropped out yet. Robust standard errors in parentheses. Clustered at Session-level. *** p<0.01, ** p<0.05, * p<0.1. Note: One subject completed the task but left the experiment during the first round without a contribution or having observed the contributions of the others. Another subject left the experiment at the end of round 1. Both subjects are omitted in the models I-III.*

Appendix D: Appendix to Chapter 5

D.I Participant instructions (all treatments)[28]

We warmly welcome you to this economic experiment. During the experiment, we are not talking about euros but about points. Your total income is therefore calculated first in points. The total score you obtained during the experiment will then be converted into euros at the end of the experiment using the following conversion rate:

1 Point = 0.01 Euro

Notes on payoffs

You will automatically receive 500 points for participating in the experiment. Your further payoffs depend on your decisions and the decisions of the other participants. It is therefore very important that you read the instructions on the following pages carefully.

At the end of the experiment, all points in your private account (shown at the end of the page) will be paid out. The payment is made by bank transfer. After the experiment you will receive another e-mail asking you to enter your account details. The data from the experiment will not be linked to your bank account details. Your information is only needed for the purpose of the bank transfer.

Experimental details (Treatment "No-Group")

In this experiment, you interact with five other randomly selected participants (players 1-6). These six participants are in one group. The composition remains the same throughout the experiment. The experiment consists of 10 rounds. In each round, one of the six participants can win a prize of 300 points.

At the beginning of each round, each participant receives 100 points. These points can then be used in any integer amount (i.e. from 0 to 100) to win the prize. Any points that are not used will be credited to your personal account at the end of the round. Once all six participants have placed their contest expenditure, the context expenditures within the group are added up and the prize is allocated.

[28] Translated from German

Prize allocation

The allocation of the prize will be done randomly. Each player has the possibility to increase the chance of winning with his contest expenditure. The chance of winning is calculated from the ratio of your own contest expenditure to the sum of the contest expenditures of all players.

$$\boldsymbol{Probability\ of\ winning} = \frac{Your\ contest\ expenditure}{Sum\ of\ all\ contest\ expenditures\ in\ your\ group}$$

If all group members have the same amount of contest expenditure, the chances of winning are the same for all group members. If you invest more than your group members, then your chances of winning increase. Note: The member with the highest contest expenditure does not necessarily win the prize. It just increases the probability.

Example

	Group Expenditure	
Group 1	240	
	Ind. Expenditure	**Probability of winning**
Player 1	50	20.8%
Player 2	60	25.0%
Player 3	0	0.0%
Player 4	100	41.7%
Player 5	10	4.2%
Player 6	20	8.3%

Player 1's winning probability is (50/240)*100 = 20.8%.
Player 2's winning probability is (60/240)*100 = 25.0%.
Player 3's winning probability is (0/240)*100 = 0.0%.
Player 4's winning probability is (100/240)*100 = 41.7%.
Player 5's winning probability is (10/240)*100 = 4.2%.
Player 6's winning probability is (20/240)*100 = 8.3%.

Payoff
100 points – contest expenditure + prize

Assume that player 4 has won.

This results in the following payoffs in this round:

Player 1: 100 points – 50 points + 0 points = 50 points
Player 2: 100 points – 60 points + 0 points = 40 points
Player 3: 100 points – 0 points + 0 points = 100 points
Player 4: 100 points – 100 points + 300 points = 300 points
Player 5: 100 points – 10 points + 0 points = 90 points
Player 6: 100 points – 20 points + 0 points = 80 points

Experimental details (Treatment "No-Share")

In this experiment, you interact with five other randomly selected participants (players 1-6). These six participants will be divided into two groups:

Group 1: Player 1, Player 3 and Player 5

Group 2: Player 2, Player 4 and Player 6

The composition remains the same throughout the experiment. The experiment consists of 10 rounds. In each round, one of the two groups can win a prize of 300 points. However, only one person from this winning group will receive the prize.

At the beginning of each round, each participant receives 100 points. These points can then be used in any integer amount (i.e. from 0 to 100) to win the prize. Any points that are not used will be credited to your personal account at the end of the round. Once all six participants have placed their contest expenditure, the context expenditures within the group are added up and the prize is allocated.

Prize allocation

The allocation of the prize is done randomly in two steps. First, it is determined which group will receive the prize. Then it will be determined which participant within the winning group will receive the prize of 300 points.

Determination of the winning group

Each player has the possibility to increase the winning chances of his group by his contest expenditure. The winning chance of the group is calculated from the ratio of the own group contest expenditure to the sum of the two group contest expenditures.

$$\boldsymbol{Probability\ of\ winning} = \frac{Own\ group\ contest\ expenditure}{Sum\ of\ the\ two\ group\ contest\ expenditures}$$

If both groups have the same amount of contest expenditure, the chances of winning are the same for both groups. If your group invests more than the other group, then your chances of winning increase. Note: The group with the highest contest expenditure does not necessarily win the prize. It just increases the probability.

Determination of the prize distribution

The allocation of the prize within the successful group will follow the same principle. The chance of winning is calculated from the ratio of your own context expenditure to the sum of the contest expenditures of your own group.

$$\boldsymbol{Probability\ of\ winning} = \frac{Your\ contest\ expenditure}{Sum\ of\ all\ contest\ expenditures\ in\ your\ group}$$

Example

	Group 1			Group 2	
	Group Expenditure	**Probability of winning**		**Group Expenditure**	**Probability of winning**
Group 1	60	25.0%	**Group 2**	180	75.0%
	Ind. Expenditure	**Probability of winning**		**Ind. Expenditure**	**Probability of winning**
Player 1	50	83.3%	Player 2	60	33.3%
Player 3	0	0.0%	Player 4	100	55.6%
Player 5	10	16.7%	Player 6	20	11.1%

Group 1's winning probability is (60/(60+180))*100= 25.0%
Group 2's winning probability is (180/(60+180))*100= 75.0%

Assume that group 2 has won:
Player 2's winning probability is (60/180)*100= 33.3%
Player 4's winning probability is (100/180)*100= 55.6%
Player 6's winning probability is (20/180)*100= 11.1%

Payoff
100 points – contest expenditure + prize

Assume that player 4 has won.
This results in the following payoffs in this round:
Player 1: 100 points – 50 points + 0 points = 50 points
Player 2: 100 points – 60 points + 0 points = 40 points
Player 3: 100 points – 0 points + 0 points = 100 points
Player 4: 100 points – 100 points + 300 points = 300 points
Player 5: 100 points – 10 points + 0 points = 90 points
Player 6: 100 points – 20 points + 0 points = 80 points

Experimental details (Treatment "Proportional-Share")

In this experiment, you interact with five other randomly selected participants (players 1-6). These six participants will be divided into two groups:

Group 1: Player 1, Player 3 and Player 5

Group 2: Player 2, Player 4 and Player 6

The composition remains the same throughout the experiment. The experiment consists of 10 rounds. In each round, one of the two groups can win a prize of 300 points.

At the beginning of each round, each participant receives 100 points. These points can then be used in any integer amount (i.e. from 0 to 100) to win the prize. Any points that are not used will be credited to your personal account at the end of the round. Once all six participants have placed their contest expenditure, the context expenditures within the group are added up and the prize is allocated.

Prize allocation

The allocation of the prize is done randomly in two steps. First, it is determined which group will receive the prize. Then it is determined how the 300 points are distributed among the three members of the winning group.

Determination of the winning group

Each player has the possibility to increase the winning chances of his group by his contest expenditure. The winning chance of the group is calculated from the ratio of the own group contest expenditure to the sum of the two group contest expenditures.

$$\boldsymbol{Probability\ of\ winning} = \frac{Own\ group\ contest\ expenditure}{Sum\ of\ the\ two\ group\ contest\ expenditures}$$

If both groups have the same amount of contest expenditure, the chances of winning are the same for both groups. If your group invests more than the other group, then your chances of winning increase. Note: The group with the highest contest expenditure does not necessarily win the prize. It just increases the probability.

Determination of the prize distribution

The allocation of the prize within the successful group is based on shares. The share (in %) is calculated from the ratio of your own contest expenditure to the sum of the contest expenditures in your group.

$$\boldsymbol{Your\ share} = \frac{Your\ contest\ expenditure}{Sum\ of\ all\ contest\ expenditures\ in\ your\ group}$$

The following applies here: The group member with the highest contest expenditure therefore also receives the largest share.

Example

Group 1			Group 2		
	Group Expenditure	**Probability of winning**		**Group Expenditure**	**Probability of winning**
Group 1	60	25.0%	**Group 2**	180	75.0%
	Ind. Expenditure	**Share**		**Ind. Expenditure**	**Share**
Player 1	50	83.3%	Player 2	60	33.3%
Player 3	0	0.0%	Player 4	100	55.6%
Player 5	10	16.7%	Player 6	20	11.1%

Group 1's winning probability is (60/(60+180))*100= 25.0%
Group 2's winning probability is (180/(60+180))*100= 75.0%

<u>Assume that group 2 has won:</u>
Player 2's share is (60/180)*100 = 33.3%. Hence, player 2 gets 300*0.333= 11 points.
Player 4's share is (100/180)*100= 55.6%. Hence, player 4 gets 300*0.556 = 167 points.
Player 6's share is (20/180)*100= 11.1%. Hence, player 6 gets 300*0.111 = 34 points.

Payoff

100 points – contest expenditure + prize

This results in the following payoffs in this round:
Player 1: 100 points – 50 points + 0 points = 50 points
Player 2: 100 points – 60 points + 11 points = 51 points
Player 3: 100 points – 0 points + 0 points = 100 points
Player 4: 100 points – 100 points + 167 points = 167 points
Player 5: 100 points – 10 points + 0 points = 90 points
Player 6: 100 points – 20 points + 34 points = 114 points

Experimental details (Treatment "Equal-Share")

In this experiment, you interact with five other randomly selected participants (players 1-6). These six participants will be divided into two groups:

Group 1: Player 1, Player 3 and Player 5

Group 2: Player 2, Player 4 and Player 6

The composition remains the same throughout the experiment. The experiment consists of 10 rounds. In each round, one of the two groups can win a prize of 300 points.

At the beginning of each round, each participant receives 100 points. These points can then be used in any integer amount (i.e. from 0 to 100) to win the prize. Any points that are not used will be credited to your personal account at the end of the round. Once all six participants have placed their contest expenditure, the context expenditures within the group are added up and the prize is allocated.

Prize allocation

The allocation of the prize is done randomly in two steps. First, it is determined which group will receive the prize. Then it is determined how the 300 points are distributed among the three members of the winning group.

Determination of the winning group

Each player has the possibility to increase the winning chances of his group by his contest expenditure. The winning chance of the group is calculated from the ratio of the own group contest expenditure to the sum of the two group contest expenditures.

$$\boldsymbol{Probability\ of\ winning} = \frac{Own\ group\ contest\ expenditure}{Sum\ of\ the\ two\ group\ contest\ expenditures}$$

If both groups have the same amount of contest expenditure, the chances of winning are the same for both groups. If your group invests more than the other group, then your chances of winning increase. Note: The group with the highest contest expenditure does not necessarily win the prize. It just increases the probability.

Determination of the prize distribution

The allocation of the prize within the successful group is based on shares. All members in the successful group receive the same share of the winnings regardless of their contest expenditure.

Example

Group 1			Group 2		
	Group Expenditure	**Probability of winning**		**Group Expenditure**	**Probability of winning**
Group 1	60	25.0%	**Group 2**	180	75.0%
	Ind. Expenditure	**Share**		**Ind. Expenditure**	**Share**
Player 1	50	33.3%	Player 2	60	33.3%
Player 3	0	33.3%	Player 4	100	33.3%
Player 5	10	33.3%	Player 6	20	33.3%

The probability that group 1 wins is (60/(60+180))*100= 25.0%
The probability that group 2 wins is (180/(60+180))*100= 75.0%

<u>Assume that group 2 has won:</u>
Player 2's share is 33.3%. Hence, player 2 gets 300*0.333= 100 points.
Player 4's share is 33.3%. Hence, player 4 gets 300*0.333= 100 points
Player 6's share is 33.3%. Hence, player 6 gets 300*0.333= 100 points

Payoff

100 points – contest expenditure + prize

This results in the following payoffs in this round:
Player 1: 100 points – 50 points + 0 points = 50 points
Player 2: 100 points – 60 points + 100 points = 140 points
Player 3: 100 points – 0 points + 0 points = 100 points
Player 4: 100 points – 100 points + 100 points = 100 points
Player 5: 100 points – 10 points + 0 points = 90 points
Player 6: 100 points – 20 points + 100 points = 180 points

D.II Reasons for exclusion

In total, 540 participants agreed to participate. However, 66 participants did not enter the code within our time restriction (see procedure in section 5.2). Thus, we observe substantial attrition already before the experiment started. This leaves us with 474 subjects. By a random mechanism, the subjects were divided into matching groups of six participants each. Unfortunately, we had to exclude one matching group from the experiment in each session (except one). In these 17 matching groups (48 participants) it was not possible to form a group of six participants since not enough participants were connected. The participants in these matching groups were rejected from the experiment. This leaves us with 426 participants in 71 matching groups. These matching groups were distributed to one of the four treatments. Unfortunately, in some cases (10 matching groups) one or two participants dropped out of the experiment during or immediately after the instructions. All participants (60 participants) in these matching groups were excluded from our analysis. Thus, in total we use 366 observations for our subsequent analysis. Table D-I summarizes the number and reasons for exclusion for each contest.

Table D-I: Number and reasons for exclusion

	NO-GROUPS	NO-SHARE	PROPOR-TIONAL-SHARE	EQUAL-SHARE	Total
Agreed to partici-pate	---	---	---	---	540
Not showed up	---	---	---	---	66
Matching Group could not be com-pletely occupied	---	---	---	---	48
Present participants	108	108	108	102	426
Drop out after in-structions	12	6	24	18	60
Included Subjects	96	102	84	84	366

Note: The table shows how many subjects were excluded and gives the respective reasons.

D.III Formal theoretical predictions

D.III.I Disadvantageous inequality aversion

We consider a contest of n > 2 individuals indexed by $i's(i \in \{1, \dots, n\})$, separated into two groups of equal size ($m = n/2$), with an initial per-capita endowment E and x_i denoting the chosen contest expenditure for the prize $P > 0$. Like Konrad and Schlesinger (1997), we limit ourselves to symmetric, pure-strategy Nash equilibria. We assume common knowledge that all participants have a uniform degree $0 \leq \alpha \leq 1$ of inequality aversion towards the fellow contestants. They resent any higher payoffs of fellow contestants (as in Fehr and Schmidt (1999)), as the following utility function for the *NO-GROUP* and *NO-SHARE* treatments shows

$$U_i = E + \frac{x_i}{\sum_{i=1}^{n} x_i} P - x_i - \left(1 - \frac{x_i}{\sum_{i=1}^{n} x_i}\right)\left[\frac{\alpha}{n-1}\left(P - \mathrm{x_{j \neq i}} + x_i\right) + \frac{\alpha}{n-1}\sum_{\mathrm{j \neq i = 1}}^{n-2} max\left(0; x_i - \mathrm{x_{j \neq i}}\right)\right] \tag{E.D1}$$

In case of a symmetric equilibrium, the relevant first-order condition is

$$\frac{\partial U_i}{\partial x_i} = \frac{P(n-1)}{\mathrm{n}^2 x_i}\left(1 + \frac{\alpha}{n-1}\right) - \frac{\mathrm{n}-1}{\mathrm{n}}\alpha - 1 = 0 \tag{E.D2}$$

which implies the following level of expenditure

$$x_i = \frac{P}{n^2}\left(\frac{n - 1 + \alpha}{1 + \frac{\mathrm{n}-1}{\mathrm{n}}\alpha}\right) \tag{E.D3}$$

Because of $\frac{\partial x_i}{\partial \alpha} < 0$, contest expenditure decreases in disadvantageous inequality aversion if two or more persons compete on each side. Note, that this comparative static effect increases in n, it is actually negative in a contest between two persons (Herrmann and Orzen, 2008).

We now study how proportional sharing among the members of the winning group influences the behavior. Since all members of the winning group receive a payment, the utility function of participants in that treatment transforms into

$$U_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i}\left(\frac{x_i}{\sum_{i=1}^{m} x_i} P\right) - x_i$$

$$\begin{aligned} & -\left(\frac{\alpha}{n-1}\right)\left[\left(\frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i}\right)\sum_{\mathrm{j}\neq\mathrm{i}=1}^{m-1} max\left(0; P\left(\frac{x_j}{\sum_{i=1}^{m} x_i} - \frac{x_i}{\sum_{i=1}^{m} x_i}\right) + x_i - \mathrm{x}_{\mathrm{j}\neq\mathrm{i}}\right)\right. \\ & \quad -\left(\frac{\sum_{m+1}^{n} x_j}{\sum_{i=1}^{n} x_i}\right)\left(\sum_{\mathrm{m+1}}^{n} \frac{x_j}{\sum_{m+1}^{n} x_j} P + x_i - \mathrm{x}_{\mathrm{j}\neq\mathrm{i}}\right. \\ & \quad \left.\left. + \sum_{\mathrm{j}\neq\mathrm{i}=1}^{m-1} max\left(0; x_i - \mathrm{x}_{\mathrm{j}\neq\mathrm{i}}\right)\right)\right] \end{aligned} \qquad \text{(E.D4)}$$

The relevant first order condition is

$$\frac{\partial U_i}{\partial x_i} = \frac{P}{\mathrm{n}^2 x_i}\left(n - 1 + \frac{\alpha\mathrm{m}}{n-1}\right) - 1 - \frac{\alpha}{2} = 0 \qquad \text{(E.D6)}$$

The resulting expenditure level is higher than without prize-sharing:

$$x_i = \frac{P}{n^2}\left(\frac{n - 1 + \frac{\mathrm{m}\alpha}{n-1}}{1 + \frac{\alpha}{2}}\right) \qquad \text{(E.D7)}$$

Proof: Comparison to *NO-SHARE:*

$$\frac{n - 1 + \alpha}{1 + \frac{\mathrm{n}-1}{\mathrm{n}}\alpha} < \frac{n - 1 + \frac{\mathrm{m}\alpha}{n-1}}{1 + \frac{\alpha}{2}}$$

$$(n - 1 + \alpha)\left(1 + \frac{\alpha}{2}\right) < \left(n - 1 + \frac{\mathrm{m}\alpha}{n-1}\right)\left(1 + \frac{\mathrm{n}-1}{\mathrm{n}}\alpha\right)$$

$$n - 1 + \alpha + (n-1)\frac{\alpha}{2} + \frac{\alpha^2}{2} < n - 1 + \frac{\mathrm{m}\alpha}{n-1} + (\mathrm{n}-1)\alpha - \frac{\mathrm{n}-1}{\mathrm{n}}\alpha + \frac{\alpha^2}{2}$$

$$1 + \frac{(n-1)}{2} < \frac{\mathrm{m}}{n-1} + (\mathrm{n}-1) - \frac{\mathrm{n}-1}{\mathrm{n}}$$

$$\frac{(n-1)}{2} < \frac{\mathrm{m}}{n-1} + n - 2 - \frac{\mathrm{n}-1}{\mathrm{n}} \tag{E.D8}$$

With equal prize sharing each member of the winning group gets $\frac{P}{m}$ as share of the prize. The utility function transforms into

$$\begin{aligned} U_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i}\left(\frac{\mathrm{P}}{\mathrm{m}}\right) - x_i - \frac{\alpha}{n-1}\left(\sum_{\mathrm{j}\neq\mathrm{i}=1}^{m-1} max\left(0; x_i - \mathrm{x}_{\mathrm{j}\neq\mathrm{i}}\right)\right) \\ - \left(\frac{\sum_{\mathrm{m}+1}^{n} x_j}{\sum_{i=1}^{n} x_i}\right)\frac{\alpha}{n-1}\left(\sum_{\mathrm{m}+1}^{n} \frac{P}{m} + x_i - \mathrm{x}_{\mathrm{j}\neq\mathrm{i}}\right) \end{aligned} \tag{E.D9}$$

The relevant first order condition in a symmetric equilibrium is

$$\begin{aligned} \frac{\mathrm{mP}}{n^2 x_i} - 1 - \frac{\alpha(m-1)}{n-1} + \left(\frac{\mathrm{m}}{n^2 x_i}\right)\frac{\alpha\mathrm{P}}{n-1} - \frac{1}{2}\frac{\alpha\mathrm{m}}{n-1} \\ \frac{\partial U_i}{\partial x_i} = \frac{\mathrm{P}}{n^2 x_i}\left(1 + \frac{\alpha\mathrm{m}}{n-1}\right) - 1 - \frac{1}{2}\left[\frac{\alpha\mathrm{m} + 2\alpha(m-1)}{n-1}\right] = 0 \end{aligned} \tag{E.D10}$$

The resulting expenditure turns into

$$x_i = \frac{P}{n^2}\left(\frac{1 + \frac{\alpha\mathrm{m}}{n-1}}{1 + \frac{2\alpha(m-1) + \alpha\mathrm{m}}{2(n-1)}}\right) \tag{E.D11}$$

This expenditure level is obviously lower than in the *PROPORTIONAL-SHARE* treatment. The equal division of the prize generates a two-fold free-rider effect that induces lower expenditure. First, it reduces the immediate marginal benefit of winning the prize. Second, it generates inequality within the winning group. A comparison with equation E.D3 also shows that it is even lower than in the *NO-SHARE* and *NO-GROUP* treatments:

$$\left(\frac{n-1+\alpha}{1 + \frac{\mathrm{n}-1}{\mathrm{n}}\alpha}\right) > \frac{1 + \frac{\alpha\mathrm{m}}{n-1}}{1 + \frac{2\alpha(m-1) + \alpha\mathrm{m}}{2(n-1)}}$$

$$(n-1+\alpha)\left(1 + \frac{2\alpha(m-1) + \alpha\mathrm{m}}{2(n-1)}\right) > \left(1 + \frac{\alpha\mathrm{m}}{n-1}\right)\left(1 + \frac{\mathrm{n}-1}{\mathrm{n}}\alpha\right)$$

$$n-1+\alpha+\alpha\frac{2(m-1)+\mathrm{m}}{2}+\alpha^2\frac{2(m-1)+\mathrm{m}}{2(n-1)}>1+\frac{\alpha\mathrm{m}}{n-1}+\frac{\mathrm{n}-1}{\mathrm{n}}\alpha+\frac{1}{2}\alpha^2$$

$$n-2+\alpha\left(1+\frac{2(m-1)+\mathrm{m}}{2}-\frac{\mathrm{m}}{n-1}-\frac{\mathrm{n}-1}{\mathrm{n}}\right)+\alpha^2\left(\frac{2(m-1)+\mathrm{m}}{2(n-1)}-\frac{1}{2}\right)>0$$

D.III.II Hostility towards the fellow contestants

We now switch to a different type of social preferences and assume common knowledge that all participants have a uniform degree $0 \leq \lambda \leq 1$ of hostility towards the fellow contestants. More precisely, they resent the utility of the other participants. Again, we focus only on symmetric equilibria.

Formally, the utility function of player i $\in \{1, \ldots, n\}$ in the *NO-GROUP* treatment is given by:

$$U_i = E + \frac{x_i}{\sum_{i=1}^{n} x_i} P - x_i - \frac{\lambda}{n-1}\left(\frac{\sum_{j\neq i=1}^{n-1} x_j}{\sum_{i=1}^{n} x_i} P - \sum_{j\neq i=1}^{n-1} x_j\right) \qquad \text{(E.D12)}$$

The resulting first-order condition yields $\left(1+\frac{\lambda}{n-1}\right)\frac{\sum_{j\neq i=1}^{n-1} x_j}{\left(\sum_{i=1}^{n} x_i\right)^2} P = 1$ which implies the following expenditure:

$$x_i = \frac{n-1+\lambda}{n^2} P \qquad \text{(E.D13)}$$

Hence, with $\lambda > 0$, participants make higher expenditure than in the payoff-maximizing Nash-Equilibrium in the *NO-GROUP* Treatment. In the *NO-SHARE* treatment, the n participants are divided into two groups of equal size ($m = n/2$). Evidence about the minimal group paradigm (Bernhard et al., 2006a; Diehl, 1990; Efferson et al., 2008; Tajfel, 1970; Tajfel et al., 1979) suggests that simple group assignment makes participants less competitive towards the members of their own group. We denote such a preference with the parameter $-1 \leq \rho < \lambda$. At $0 < \rho < \lambda$ participants are less hostile towards their fellow group members than to the other members of the other group. In case of $-1 \leq \rho < 0$ they reveal in-group favoritism. Again, we assume common knowledge about a uniform ρ. This additional parameter transforms the objective function for a member of group 1 into:

$$U_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} \frac{x_i}{\sum_{i=1}^{m} x_i} P - x_i - \frac{\rho}{n-1}\left(\frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{n} x_i} \frac{\sum_{j\neq i=1}^{m-1} x_j}{\sum_{i=1}^{m} x_i} P - \sum_{j\neq i=1}^{m-1} x_j\right) - \frac{\lambda}{n-1}\left(\frac{\sum_{m+1}^{n} x_i}{\sum_{i=1}^{n} x_i} P - \sum_{m+1}^{n} x_j\right) \quad \text{(E.D14)}$$

The first part of equation E.D14 represents the utility from one's own expenditures as well as the utility gain/loss from the payoffs of the own group members. The second part of equation E.D14 measures the utility loss from the generated payoff of members from the other group (out-group hostility). Equation E.D14 implies the following optimality condition with respect to x_i:

$$\frac{\partial U_i}{\partial x_i} = \frac{\sum_{j\neq i=1}^{n-1} x_j}{(\sum_{i=1}^{n} x_i)^2} P + \frac{\rho}{n-1}\left(\frac{\sum_{j\neq i=1}^{m-1} x_j}{(\sum_{i=1}^{n} x_i)^2} P\right) + \frac{\lambda}{n-1}\left(\frac{\sum_{m+1}^{n} x_i}{(\sum_{i=1}^{n} x_i)^2} P\right) - 1 = 0 \quad \text{(E.D15)}$$

Again, in equilibrium all contestants make the same expenditure $\left(x_i = x_{j\neq i}\right)$. A transformation of the first order condition with respect to x_i yields:

$$\frac{n-1+\frac{\rho(m-1)}{n-1}+\frac{\lambda m}{n-1}}{n^2} P = x_i \quad \text{(E.D16)}$$

With $\rho < \lambda$, the decline of hostility towards the members of the in-group has a negative impact on contest expenditure relative to the contest without any groups. Hence, contest expenditure is lower in the *NO-SHARE* than in the *NO-GROUPS* treatment.

Unlike in the *NO-SHARE* treatment, members of the winning group in the *PROPORTIONAL-SHARE* treatment share the prize according to their expenditure. Since the expected payoffs do not differ, hostility towards the fellow contestants does not predict any behavioral differences between the two treatments.

The *EQUAL-SHARE* treatment has a different sharing rule than the *PROPORTIONAL-SHARE* treatment. Each member of the winning group gets a prize $\frac{P}{m}$. Hence, we obtain the following utility function if we take considerations of in-group favoritism and opponents' hostility into account as well.

$$U_i = E + \frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{m} x_i + \sum_{m+1}^{n} x_i} \frac{P}{m} - x_i - \frac{\rho(m-1)}{n-1}\left(\frac{\sum_{i=1}^{m} x_i}{\sum_{i=1}^{m} x_i + \sum_{m+1}^{n} x_i} \frac{P}{m} - \sum_{j \neq i=1}^{m-1} x_j\right) - \frac{\lambda m}{n-1}\left(\frac{\sum_{m+1}^{n} x_i}{\sum_{i=1}^{m} x_i + \sum_{m+1}^{n} x_i} \frac{P}{m} - \sum_{m+1}^{n} x_i\right) \quad \text{(E.D17)}$$

The resulting first-order condition yields:

$$\frac{P}{xn^2} - P\left(\frac{(m-1)\rho + m\lambda}{(n-1)xn^2}\right) = 1 \quad \text{(E.D18)}$$

A transformation of the resulting first order condition with respect to x_i and a symmetric equilibrium yield the following average expenditure level per group member

$$\frac{P}{n^2} - P\left(\frac{(m-1)\rho + m\lambda}{(n-1)n^2}\right) = x \quad \text{(E.D19)}$$

Note that individual contributions within a group may vary (they are perfect strategic substitutes). A comparison with equation E.D16 implies that average contest expenditure is lower in the *EQUAL-SHARE* treatment than in the *PROPORTIONAL-SHARE* treatment. The proportional sharing internalizes the external effects of contest expenditure on fellow group members which increases the expenditure incentives. As in the other treatments, contest expenditure decreases without in-group preferences $(\rho > 0)$. However, a preference for the fellow group members $(\rho < 0)$ has a positive rather than a negative impact on expenditure because any investment in the conflict generates a positive externality for the members of their own group.

D.IV Further regressions

In this section, we provide additional estimations that extend our models in the main part of paper. Table D-II provide additional regressions for Table 5-4, while Table D-III for Table 5-5 in section 5.4.3.

Table D-II: Contest behavior across treatments (further regressions)

Dep. Var. Per capita contest expenditure	**All Treatments** Benchmark: NO-GROUP				
	I	**II**	**III**	**IV**	**V**
No-Share (NS)	-3.670	-2.086	-3.333	-26.67**	-19.29***
	(3.198)	(1.923)	(2.965)	(10.16)	(5.131)
Proportional-Share (PS)	9.423***	5.417***	8.572***	-16.89	-9.637
	(2.875)	(1.738)	(2.447)	(13.46)	(7.348)
Equal-Share (ES)	-14.00***	-8.203***	-12.76***	-47.81***	-25.65***
	(3.067)	(1.996)	(3.085)	(8.621)	(5.038)
Own expenditure t-1		0.504***			0.495***
		(0.0272)			(0.0275)
Avg. Others t-1 (AO)			0.161*	-0.281**	-0.147*
			(0.0818)	(0.140)	(0.0751)
NS*AO				0.454**	0.348***
				(0.192)	(0.0942)
PS*AO				0.487*	0.272*
				(0.256)	(0.140)
ES*AO				0.764***	0.403***
				(0.173)	(0.102)
Round	-0.495*	-0.692***	-0.828***	-0.616**	-0.540***
	(0.295)	(0.200)	(0.280)	(0.268)	(0.193)
Constant	53.94***	29.85***	48.30***	69.78***	36.96***
	(2.156)	(2.131)	(4.793)	(7.558)	(4.454)
Observations	3,660	3,294	3,294	3,294	3,294
Subjects	366	366	366	366	366
Matching Groups	61	61	61	61	61
R-squared	0.052	0.290	0.065	0.078	0.297

Note: Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

Table D-III: The impact of group affiliations on the contest behavior (further regressions)

Dep. Var. Per capita contest expenditure	**Only treatments with groups** Benchmark: NO-SHARE					
	I	**II**	**III**	**IV**	**V**	**VI**
Proportional-Share	8.762	10.4***	-1.729	13.7***	23.34*	5.424
	(6.267)	(3.010)	(9.116)	(3.747)	(13.47)	(11.80)
Equal-Share	-8.640*	-8.294**	-14.99*	-10.81**	-21.79**	-20.70**
	(4.744)	(3.337)	(7.698)	(4.262)	(8.870)	(8.312)
Winner group t-1 (WG)	8.07***					
	(2.978)					
Proportional-Share*WG	10.43					
	(6.699)					
Equal-Share*WG	-4.687					
	(4.043)					
Avg. Others (Own group) t-1		0.28***	0.170*			0.171*
		(0.0645)	(0.0902)			(0.0894)
Avg. Others (Other group) t-1				0.0183	-0.00918	0.00574
				(0.107)	(0.130)	(0.105)
Proportional-Share *AOOwn			0.223			0.207
			(0.148)			(0.141)
Equal-Share *AOOwn			0.148			0.110
			(0.144)			(0.161)
Proportional-Share *AOOther					-0.152	-0.103
					(0.261)	(0.180)
Equal-Share *AOOther					0.281	0.189
					(0.196)	(0.166)
Round	-1.13***	-0.95***	-0.98***	-1.116**	-0.860**	-0.811**
	(0.413)	(0.316)	(0.308)	(0.418)	(0.412)	(0.305)
Constant	50.5***	40.1***	45.5***	53.5***	53.3***	44.1***
	(4.242)	(4.550)	(5.740)	(6.664)	(7.066)	(7.123)
Observations	2,430	2,430	2,430	2,430	2,430	2,430
Subjects	243	243	243	243	243	243
Matching Groups	45	45	45	45	45	45
R-squared	0.107	0.107	0.124	0.081	0.091	0.128

Note: Robust standard errors in parentheses clustered at the matching group level.
**** p<0.01, ** p<0.05, * p<0.1*

Appendix E: Appendix to Chapter 6

E.I Participant instructions (First part)[29]

The first experiment consists of 10 rounds in which you will make decisions in two decision situations. Each situation will be present in 5 of the 10 rounds. In both situations there will be two players A and B, to whom an amount of 10 points will be divided. In the following, the situations will be explained first. You will then be informed about your role in these situations.

First allocation game 'Splitting':

The first game is called 'Splitting'. Player A decides on the allocation of 10 tokens between player B and himself. Player A can choose between six possible allocations:

Aufteilungs-nummer	1	2	3	4	5	6
A erhält	0 Punkte	2 Punkte	4 Punkte	6 Punkte	8 Punkte	10 Punkte
B erhält	10 Punkte	8 Punkte	6 Punkte	4 Punkte	2 Punkte	0 Punkte

Player A can choose one out of the allocations (1) to (6). The chosen allocation determines how many tokens player A receives and how many tokens are transferred to player B. Another allocation than the six is not possible.

Second allocation game: 'Rolling and splitting'

The second game is called 'rolling and splitting'. As in the first game, player A has to divide 10 tokens between herself and player B. As well as in the first game, there are six possible allocations:

Aufteilungs-nummer	1	2	3	4	5	6
	⚀	⚁	⚂	⚃	⚄	⚅
A erhält	0 Punkte	2 Punkte	4 Punkte	6 Punkte	8 Punkte	10 Punkte
B erhält	10 Punkte	8 Punkte	6 Punkte	4 Punkte	2 Punkte	0 Punkte

Unlike in the first situation, player A does not simply choose one of the six allocations. Instead, player A first rolls a die at least twice, whereby the die is being displayed on the computer. The

[29] Translated from German

first roll decides which allocation is chosen and thus how many tokens player A received and how many tokens are transferred to player B. The second roll is just to make sure that the die is working properly. Of course, you can roll the die more than twice. However, only the first roll counts. After player A has seen the result of the die, she enters the number (1, 2, 3, 4, 5 or 6) for the allocation on the computer.

In the following, each of you will first decide in the role of A. This means that each participant will allocate the 10 tokens ten times - five times in the ‘Splitting’ situation and five times in the ‘Rolling the Splitting’ situation. Which situation you start with will be decided by chance.

After these ten rounds, your reward will be determined as follows: First, all participants in this room will be randomly assigned to groups of two. In each group of two, one person assumes the role of A and the other person the role of B. This assignment is random. Then one of the 10 rounds is drawn as the "pay round" for each group of twos formed in this way. The decision of A in this pay round determines the pay of A and B in the group of two. Note, however, that you will be informed about their payoff from this first part only at the end of the second part.

Example (the drawn round and the distribution number are chosen arbitrarily)

For a group of two, the sixth round was drawn as a pay round. The situation in this round was ‘Splitting’. The participant who was assigned the role of A chose no. 3 in this round. Therefore, he receives 4 points, and the participant who was assigned the role of B in the same group receives 6 points as a reward for this first part.

E.II Participant instructions (Part 2)[30]

The second part is similar to the first part. It again deals with the two decision situations ‘Splitting’ and ‘Rolling and Splitting’, which you already know. However, only one round will be played.

In the following, pairs of two participants will be formed. One participant takes the role of player A, the other the role of player B. The participant in the role of player A makes his decision either in the situation ‘splitting’ or in the situation ‘rolling and splitting’. The situation is randomly determined by the computer system. The participant in the role of player B does not

[30] Translated from German

make a decision in this second part. By her decision, player A determines the payoff of player A and player B in this second part.

On screen Instructions (only player A, player B received no further instructions)

You are player A.

As explained, the computer system randomly determines whether you play the game ‘Splitting’ or the game ‘Rolling and Splitting’. Only you will know which game you actually played. Once you have learned about the allocation game you are playing, you determine the allocation as explained before.

Please note: After you have entered your allocation number (1, 2, 3, 4, 5 or 6), you have to make another decision. You decide in which form the allocation will be communicated to player B. You have two options:

Message 1: ‘Player A decided the following allocation’

Message 2: ‘A die was rolled, and the following allocation was chosen’.

Therefore, you can determine whether the message to player B corresponds to the game you have actual played or to the other allocation game that did not exist. For example, if you play the allocation game to ‘Splitting’, you can choose whether player B should receive the message corresponding to the situation (‘Player A decided the following allocation’: ...’), or whether player B should receive the message ‘A die was rolled, and the following allocation was chosen: ...’. In the same way, if you actually played the game ‘Rolling and Splitting’, you can choose whether your message corresponds to this situation or to the ‘Splitting’ game.

E.III Screenshots

Figure E- I: Decision-Screen in the Standard-Dictator-Game (part 1 and 2)

Situation 'Splitting'

You now determine the allocation. Enter your selected split number below.

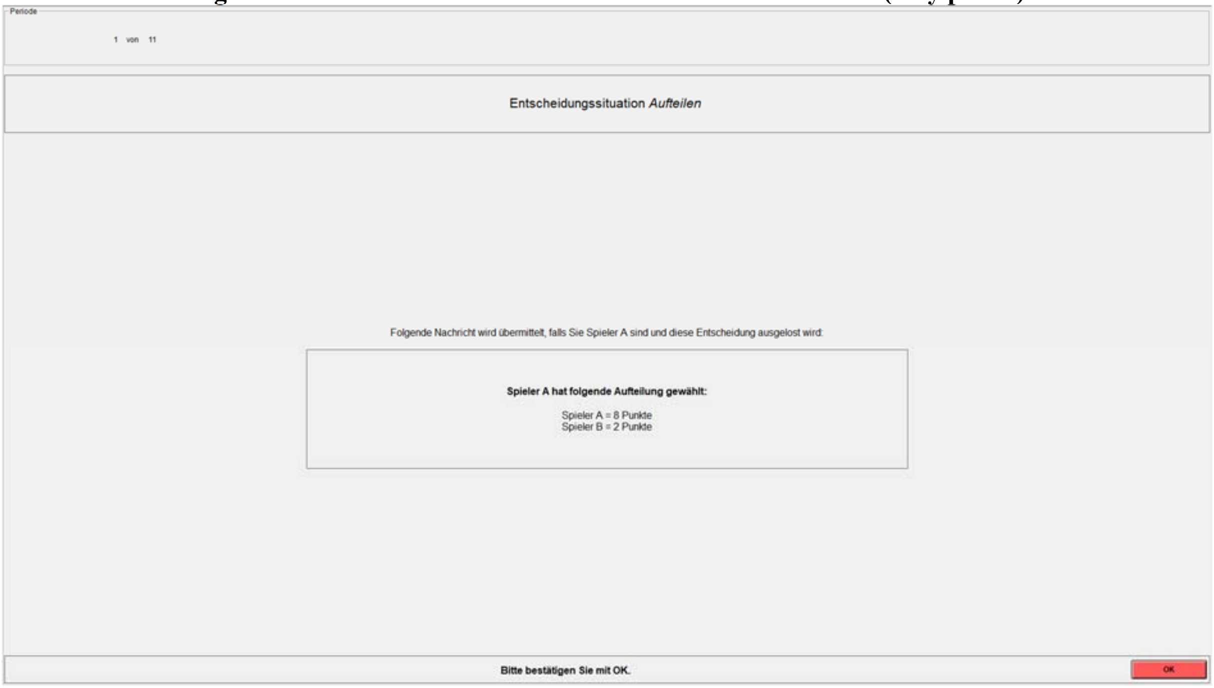
Allocation number	**1**	**2**	**3**	**4**	**5**	**6**
A gets	0 points	2 points	4 points	6 points	8 points	10 points
B gets	10 points	8 points	6 points	4 points	2 points	0 points

Please enter your allocation number.

Allocation number:

Enter 1 and the allocation number 1 will be selected, enter 2 and the allocation number 2 will be selected, etc.

Figure E- II: Results-Screen in the Standard-Dictator-Game (only part 1)



Situation 'Splitting'

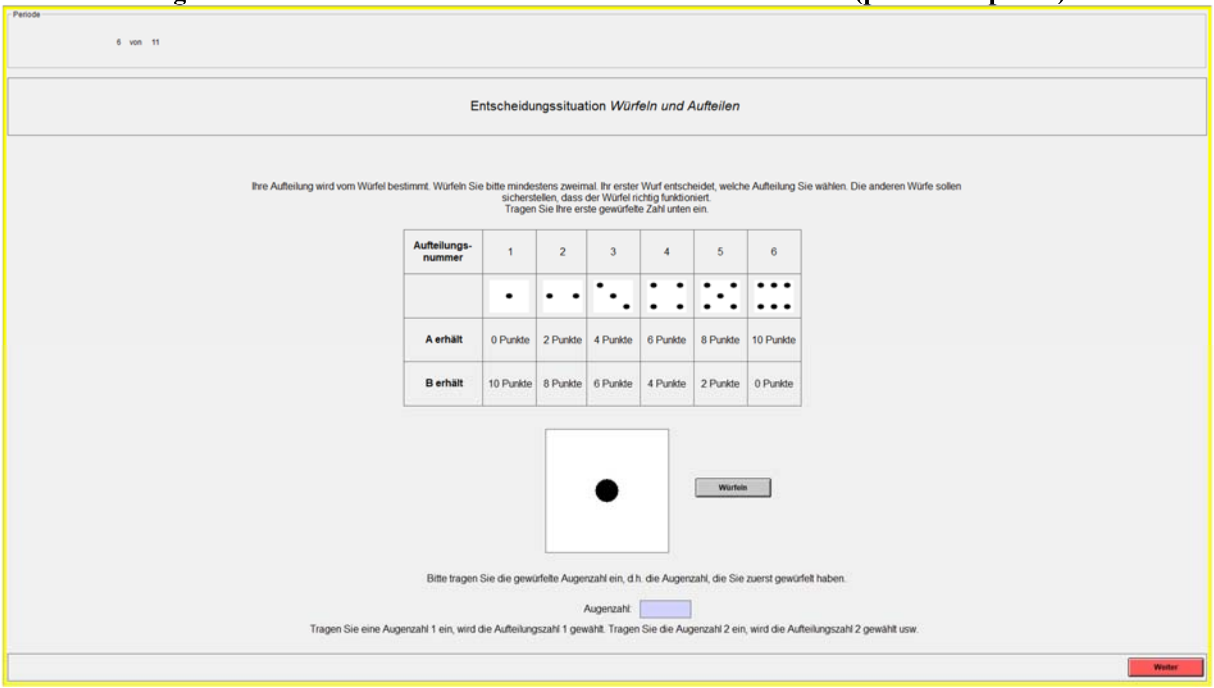
The following message will be sent if you are player A and this decision is drawn.

Player A has chosen the following allocation:

Player A: 8 tokens

Player B: 2 tokens

Figure E- III: Decision-Screen in the Garbled-Dictator-Game (part 1 and part 2)



Situation 'Rolling and Splitting'

Your allocation is determined by a die. Roll the die at least twice. However, your first roll decides which allocation you choose. You other rolls are to make sure that the die works correctly. Enter your first rolled number below.

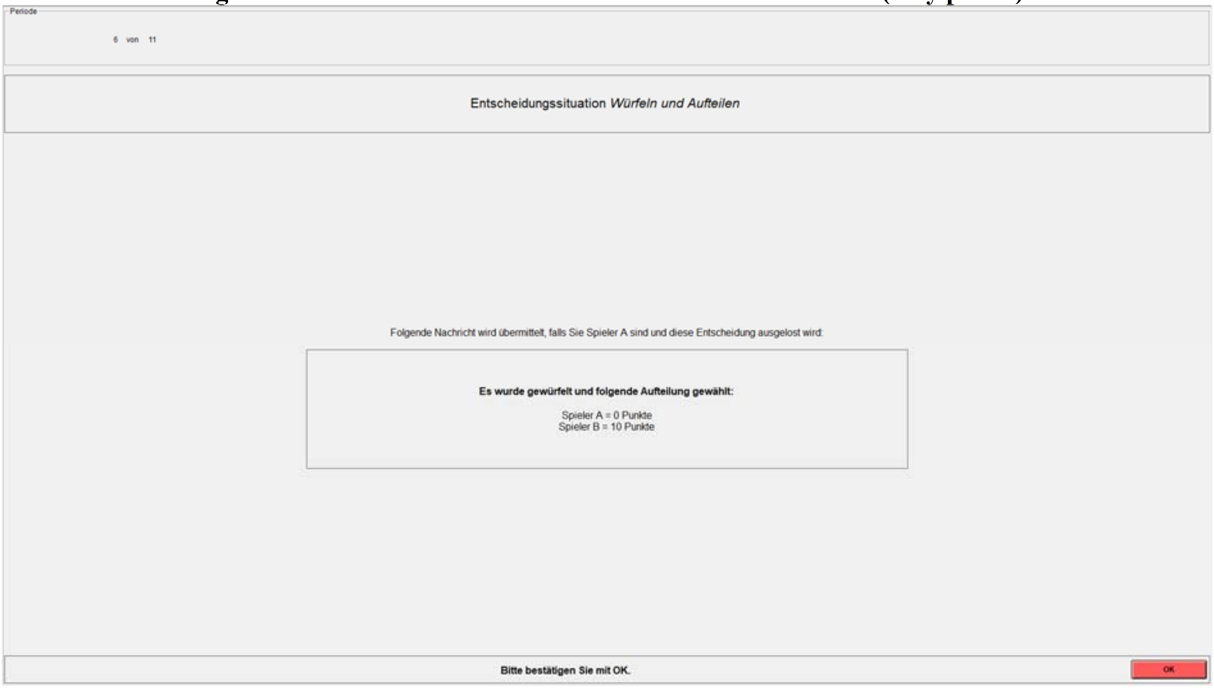
Allocation number	1	2	3	4	5	6
A gets	0 points	2 points	4 points	6 points	8 points	10 points
B gets	10 points	8 points	6 points	4 points	2 points	0 points

Enter the number you rolled first.

Rolled number:

Enter 1 and the allocation number 1 will be selected, enter 2 and the allocation number 2 will be selected, etc.

Figure E- IV: Results-Screen in the Garbled-Dictator-Game (only part 1)



Situation 'Rolling and Splitting'

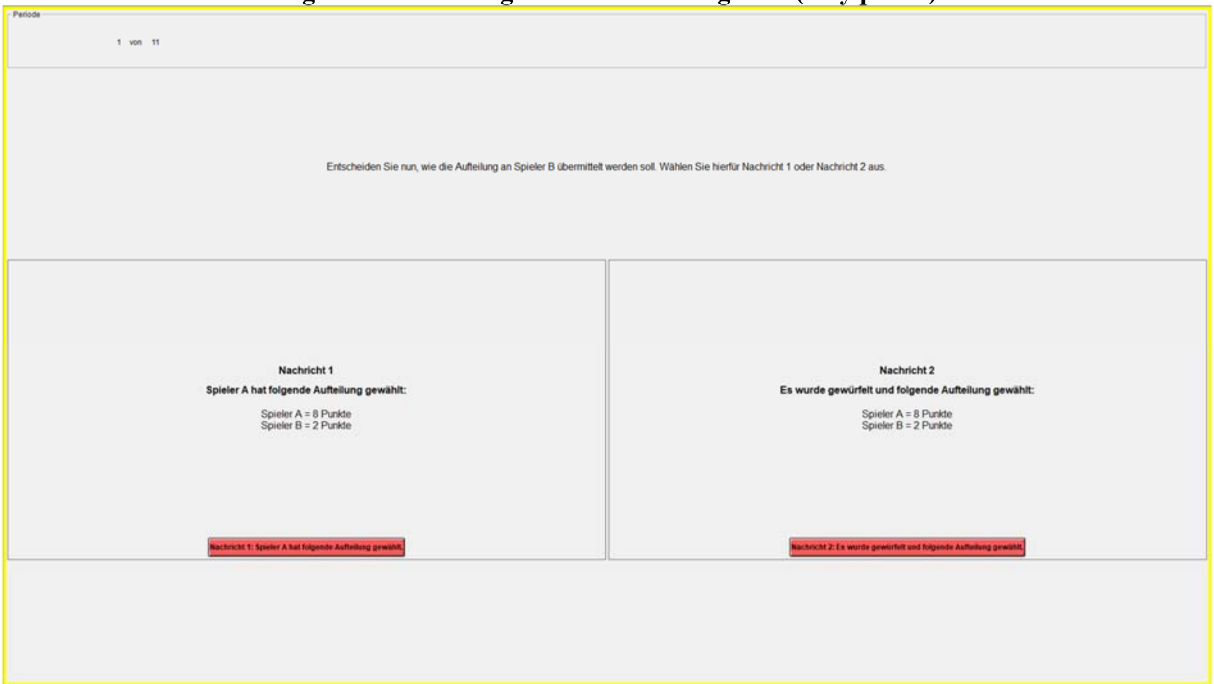
The following message will be sent if you are player A and this decision is drawn.

A die was rolled and the following allocation was chosen:

Player A: 0 tokens

Player B: 10 tokens

Figure E- V: Message-Screen for both games (only part 2)



Now decide how the allocation should be communicated to player B. Please select message 1 or message 2 for this.

Message 1	**Message 2**
Player A has chosen the following allocation:	**A die was rolled, and the following allocation was chosen:**
Player A: 8 tokens	Player A: 8 tokens
Player B: 2 tokens	Player B: 2 tokens
Message 1: Player A has chosen the following allocation	**Message 2: A die was rolled, and the following allocation was chosen**

E.IV Analysis of order effects in the first part of the experiment

This section provides analysis regarding order effect in the first part of the experiment. Table E-I shows that average number of retained tokens for both treatments, separated by the initial treatment played. A Mann-Whitney-U-Test shows only insignificant differences in the order of treatments for both the *Standard-Dictator-Game* (.69, Mann-Whitney-U-Test) and the *Garbled-Dictator-Game* (.36, Mann-Whitney-U-Test).

Table E-I: Order effects

Started with \ Treatment	Standard-Dictator-Game	Garbled-Dictator-Game
Standard-Dictator-Game	7.56 (1.85)	5.99 (2.34)
Garbled-Dictator-Game	7.39 (2.06)	5.43 (2.35)

Note: Standard deviations in parentheses

E.V Further statistics including all participants

This section provides further statistics for the first part of the experiment, including all participants. Table E-II shows the number of retained tokens for all participants, while Figure E-VI presents the cumulative distribution of retained tokens. Last, not least, Figure E-VII focus on the reported numbers in the *Garbled-Dictator-Game* in the first part.

Table E-II: Descriptive statistics including all participants (first part)

	Standard-Dictator Game	Garbled-Dictator Game
Number of retained tokens	6.91 (1.87)	5.79 (2.00)
Observations	234	234

Note: Standard deviation in parentheses

Figure E-VI: Cumulative distribution of retained tokens including all participants

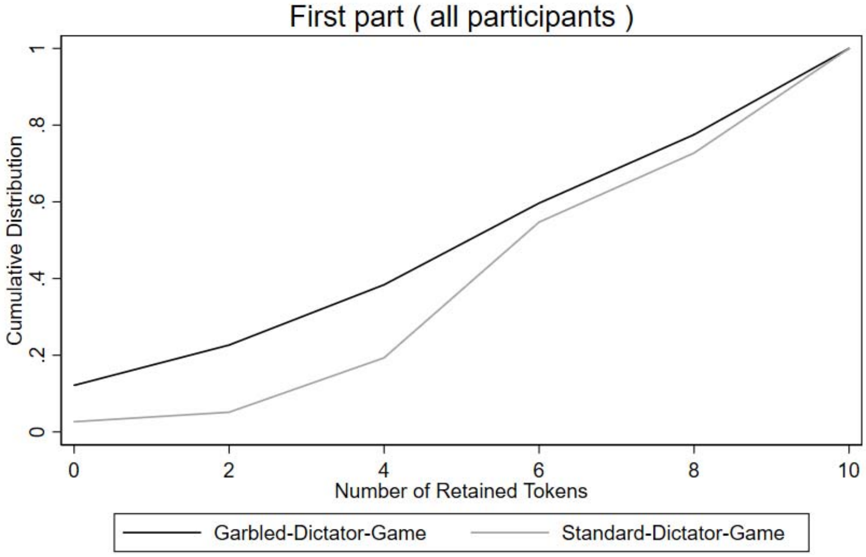


Figure E-VII: Distribution of reported number of retained tokens in the Garbled-Dictator-Game