Social Context and Decisions

Essays in Experimental Economics

NIVES DELLA VALLE

School of Social Sciences
UNIVERSITY OF TRENTO

A dissertation submitted to the Doctoral School of Social Sciences in partial fulfilment of the requirements of the DOCTOR OF PHILOSOPHY (Ph.D) degree in Economics and Management

APRIL 2017

The assessment committee consists of Prof. Maria Vittoria Levati from the University of Verona and Prof. Shaul Shalvi from the University of Amsterdam.
I am extremely grateful to them for their careful reading of my Thesis and their constructive comments and suggestions.

Abstract

Humans are social animals who evolved to live in societies. They are "encultured" actors as their preferences, perceptions and values are shaped by the social context to which they are exposed. Part of economic failures is due to suboptimal social contexts which determine individuals' decisions. These social contexts can be better designed by organizations and governments.

The ultimate goal of this research is to emphasize that social context can be detrimental for individual decisions, providing empirically-based behavioral insights for policy makers who wish to implement regulatory policies on corruption, gender gap and injustice.

Behavioral and Experimental Economics provides a clean tool to keep the internal validity necessary to disentangle complex behavioral aspects that cannot be easily observed in the field, such as those related to the influence of social environment.

This Doctoral Thesis is a collection of three laboratory experimental essays about the interplay between suboptimal social contexts and decisions. The first Chapter investigates the role of group identity in unethical decisions motivated by unfairness. The second Chapter provides evidence of gender stereotype in perceptions of others' risk attitudes. The third Chapter shows that small contextual changes can promote the diffusion of corruption while others inhibit it.

Keywords: Laboratory Experiment; Behavioral Economics; Experimental Economics; Group Identity; Fairness; Dishonesty; Risk Preferences; Gender Stereotype; Group Dishonesty; Punishment; Conflict of Interest; Loss Aversion

JEL classification: C91; C92; D03; D63; D81; Z1

Acknowledgements

This Doctoral Thesis represents the end of a journey that dates back farther than the time I entered the PhD programme in November 2013. Since I arrived in the enchanting town of Trento, I have met many people that guided me along an incredible growth process and enabled me to produce this work.

I am deeply thankful to my supervisor Dr. Matteo Ploner, who has been teaching me that every goal is within reach if we just keep trying (climbing) while being mentally balanced.

I am grateful to prof. Luigi Mittone for introducing me to the world of applied research and believing in me.

I am thankful to prof. Marco Piovesan, who taught me the passion for Behavioral Economics and doing research.

I am also thankful to prof. Morten Lau, who taught me how to be a rigorous researcher.

I am grateful to prof. Enrico Zaninotto, for always supporting me and making me meet prof. Joseph Stiglitz.

Special thanks go to Marco Tecilla, Dominique Cappelletti, Marco Faillo, M.Laura Frigotto, Sandro Casal, Marianna Baggio, Natalia Montinari, Amalia Di Girolamo, all the members of the CEEL, and the research brothers Ali, Chiara, Kaku, Sara, Piero, Tam, Mario, Lucio, Maria, Maggie, Francesco, Alice, Costanza and Tatiana, who constantly supported (and tolerated) me: I would not be the researcher I am now without you.

I am grateful to Caleb and Ella Cox, Valeria, Orsola, and my officemates Olga and Helene for making my stay in Durham and Copenhagen colorful.

It would have been impossible to grow during these years without my second family: Alessia, Carlotta, Azzurra, Ananya, Rosaria, Julia, the family Moser and all the friends from the College of Merit.

I want to thank the members of the Ciurma, Luigi, Anne, Anna Maria, Claudia, Matteo and Lucìa for their friendship no matter the distance. I am grateful to my favorite mathematicians Laura and Valeria for teaching me more than climbing. It would be impossible to mention all the other friends and people met during these years. I gratefully thank all of them.

I am deeply thankful to Chiara, the sister that I have never had. For their unconditional love, for always making me believe in my dreams, and for accompanying me throughout my life, I thank from the bottom of my heart my brothers Orlando and Pierpaolo, and my parents: you made me who I am today. Finally Eugenio: with your love you make my wings free to fly ever higher.

Table of Contents

Page

List of Tables

Table **Page**

List of Figures

Figure **Page**

Introduction

"[it] is not the consciousness of men that determines their being, but on the contrary it is their social being that determines their consciousness" (Marx, 1904)

"I did not know myself at all, I did not have any reality as my own, I was in a state like continuous illusion, almost fluid and malleable; the others knew me, each according to which reality they gave me" (Pirandello, 1926)

Humans are social animals who live in communities of individuals. Every individual's starting point in society is the outcome of a lottery, which is a matter of good or bad "fortune" (Rawls, 1971). In particular, every individual is exposed to a random social environment, such as the family into which she is born and the city in which she spent most of her life. Part of individuals' goals and values are not innate, but evolved in the social context to which they are exposed.

A well-educated family might raise the offsprings to pursue long-term goals more than a poor family. The latter might, instead, be forced to promote an education that centralizes short-term benefits: entering the workforce soon might allow the whole family to earn the money necessary to satisfy basic needs, while going to school might lead the whole family to losses. Being exposed to an environment in which resources are scarce shifts preferences toward immediate rewards (Sturge-Apple *et al.*, 2016). Living in a neighbourhood characterized by a culture of organized crime affects individuals' everyday life and their interactions with others: the threat of being punished for not conforming to the norms of the community has a detrimental effect on pro-social behavior, like trust and cooperativeness (Meier *et al.*, 2016; Nese *et al.*, 2013). Being surrounded by pervasive corruption shapes individuals' perceptions of ethical values, as individuals might view dishonesty as normal and justifiable (Gächter & Schulz, 2016). Being denied the access to the workforce shapes perceptions of fairness: to avoid cognitive dissonance from being unable to find a job, individuals change the way they acknowledge earned entitlement (Barr *et al.*, 2016).

These examples suggest that preferences, perceptions and values are not exogenously given. They are, instead, endogenously shaped by the surrounding social environment. Social patterns alter who a person is, and those who have been exposed to a specific social context may further sustain those social patterns (Bowles, 1998; Hoff & Stiglitz, 2016).

Economists have traditionally assumed that individuals have stable preferences, associating any change in behavior to changes in predictable factors, like information and prices. Behavioral Economics challenges this view by engaging with the application of psychological insights to economic analysis, and unveiling a realistic picture of individual preferences. Extensive investigation in Behavioral Economics has shown that individual preferences are not stable, but distorted by universally shared biases, such as anchoring, framing and endowment effects, that the context of the moment of decision makes salient (Tversky & Kahneman, 1975). Only recently Behavioral Economics has started incorporating insights also from sociology and anthropology, wherein decision makers are seen as "encultured actors" (Hoff & Stiglitz, 2016). This strand of Behavioral Economics acknowledges the durable influence that social context has on behavior, preferences and perceptions. Not only social context influences behavior by making certain cultural mental models [1] or reference points salient , but it also shapes who people are. Some societies are trapped into social and economic failures because, on the one hand, they make salient suboptimal mental models and, on the other hand, they are sustained by individuals' conformity to those suboptimal mental models.

The ultimate goal of this work is to emphasize that exposure to social contexts that make suboptimal mental models salient directly affects individual preferences, perceptions and values. This research aims to provide empirically-based behavioural insights for policy makers and organizations who wish to regulate corruption, gender gap and injustice. Experiencing episodes of unfairness, being surrounded by an environment that does not allow to increase welfare other than through corruption, and living in a gender-unequal society are the three examples of exposure to suboptimal social contexts that are investigated in this Doctoral Thesis.

The first Chapter is based on the working paper "Reacting to Unfairness: Group Identity and Dishonest Behavior" with Matteo Ploner. It investigates a social context in which individuals experience unfairness and the role that group identity plays in mediating unethical decisions. Experimental evidence suggests that experiencing unfairness motivates individuals to engage in unethical behavior. We conduct a laboratory experiment where we investigate how group identity changes perceptions of unfairness and, thus, unethical decisions motivated by experiences of unfairness. We define unfairness as deviation from proportionality in a real-effort dictator game. We base our hypotheses on the Social Identity Theory (Tajfel *et al.*, 1971) and evidence from neuroscience (Baumgartner *et al.*, 2013). Group Identity induces individuals to interpret unfairness made by an in-group differently from that made by an out-group. While individuals tolerate ingroup-unfairness, they react to out-group unfairness more than when unfairness is made by an individual without a salient group identity. Our findings show that sharing a group identity makes salient a mental model that changes perceptions of unfairness and prevents individuals to engage in unethical behavior motivated by unfairness. On the

[1] "(Cultural mental models) shape the way we attend to, interpret, remember, and respond emotionally to the information we encounter and possess" (DiMaggio (1997), p.274).

other hand, when the context allows for conflicting group identities to exist, individuals become more sensitive to experiences of unfairness and, in turn, more prone to engage in fairness-restorative unethical behavior. The implications of these findings for organizations and institutions are that failing to build a shared identity is likely to promote negative externalities on society: experiencing unfairness from someone with different values leads individuals to react by engaging in unethical behavior at cost of society.

I started working on the project for the second Chapter "Perceptions of others' Risk Attitude and Gender Stereotype", co-authored with Morten I. Lau and Caleb Cox, during my stay at Durham Business School. We conducted an incentivized laboratory experiment to investigate the role that mental models, such as gender stereotype, play on individual risk attitudes and perceptions of others' risk attitudes. The stereotype that women are more risk averse than men is cued by the salience of a stereotyped information, such as gender, and helps conceptualizing a situation, like that in which individuals have to make decisions based on their perceptions of others' risk attitudes. Whether women (men) are in fact more risk averse (taking) than men (women) is still an open question in the literature, and decisions based on this stereotype might provide the basis for inaccurate statistical discrimination. To this aim, we investigate gender difference in risk attitudes and gender stereotype in perceptions of others' risk attitudes. By structurally estimating the risk attitude coefficients underlying two alternative latent models of choice, we find no gender difference in risk attitudes. On the other hand, we find that males use gender salience to stereotype others' risk attitude: females are perceived to be more risk averse than they actually are. This finding is of practical relevance for job employers who wish to give promotions, as they might base their choices on stereotypes that create statistical discrimination.

During my stay at the University of Copenhagen, I started working on the project "Honesty Under Threat", co-authored with Marco Piovesan and Francesca Gino. The third Chapter presents experimental evidence that small changes to a suboptimal social context can promote the diffusion of corruption while others inhibit it. The starting point of our suboptimal social context is the group setting of Weisel & Shalvi (2015): individuals are exposed to an environment that gives the incentive to corrupt to earn some benefits. This setting creates a situation in which individuals are forced to dishonestly cooperate, even though this is against ethical values. In three experiments, we investigate how introducing a norm-enforcement mechanism -such as second party and third party punishment- a conflict of interest, and negative frame affects perceptions of corruption. We find that being exposed to a suboptimal social context distorts the typical efficacy of punishment mechanisms to enforce positive norms: by increasing dark cooperation, punishment lowers the moral cost associated to corruption and changes perceptions of corruption. Second, we find that introducing a conflict of interest slows down the diffusion of corruption by making salient inequality aversion concerns. Finally, we do not find that loss aversion has an effect on the diffusion of corruption when the group setting is suboptimally designed. These findings are of practical relevance for institutions and organizations that narrow the access to fair opportunities to earn benefits, while promoting incentives that equally reward individuals in the group. In these settings, corruption can spread in the presence of punishment among peers and non-credible monitoring. On the other hand, corruption slows down if the incentive scheme that rewards individuals in the group embeds inequality.

It is my hope that the evidence collected in these three experimental essays will contribute to the understanding of how to design the right social contexts which can eventually be used to inform policy makers and organizations.

CHAPTER 1

Reacting to Unfairness: Group Identity and Dishonest Behavior

with Matteo Ploner - University of Trento

Employees' misbehavior can be attributed to experiences of unfairness. Does this dishonest reaction change when employees identify with the whole organization or with a subunit only? We experimentally investigate whether individuals are more likely to engage in dishonest behavior after having experienced unfairness perpetrated by an individual with a salient group identity. Two individuals generate an endowment together, but only one can decide how to share it. They either share the same group identity or have distinct group identities. Then, they approach a task in which they can opportunistically engage in dishonest behavior. Our results show that when individuals share the same group identity, unfair distributive decisions do not trigger a dishonest reaction. In contrast, when different group identities coexist, dishonest behavior is observed as a reaction to unfairness.

Keywords: Group Identity, Fairness, Dishonesty
JEL classification: C91; C92; D03; D63

1.1 Introduction

The general press often reports episodes of dishonest behavior: overstated value of claims to insurance companies, consumers led to fraud through misadvice, intellectual property theft, tax evasion. These are only a few examples of dishonest actions which cost huge losses to society (Mazar & Ariely, 2006). Given the economic relevance of the consequences of such a behavior, scholars from psychology and economics have extensively examined their driving factors and the potential tools for limiting them. Individuals might decide to engage in dishonest behavior when they are given the opportunity. They might decide to take advantage of contexts in which the probability to be detected is low, but they might also prefer to give up the additional profit when this is associated to a violation of moral rules. These decisions have been widely investigated in the literature, as we review in the dedicated session. But individuals can also opt to behave dishonestly as a reaction to decisions made by other individuals. When employees perceive decisions to be unfair, they become more willing to engage in dishonest behavior, viewed as a way to balance perceived unfairness (Hollinger & Clark, 1983). An employee might be unable to sanction an unfair supervisor and attempt to restore fairness by increasing dishonest behavior outside organizational borders.

Organizations are communities of individuals that rely on the establishment of a culture that regulates social interactions. Managers can promote the diffusion of a corporate culture with the aim of making employees share the same goals and values. Prior research shows that corporate culture is crucial to foster employees' sense of identification in the organization and to boost employees' productivity and commitment toward the organization (Akerlof & Kranton, 2005). But there are also potential drawbacks to high levels of organizational identification. First, the establishment of a strong group identity might promote tolerance towards unfair behavior to not undermine the positive perception of the group to which belonging. Second, modularity in the organizational structure might encourage the development of "local" identities with different values and goals. Therefore, between-units decisions might not be equally embraced across the organization. The establishment of diverse group identities within the same organization might promote the diffusion of negative externalities associated to unfair decisions made by a supervisor belonging to a different unit. It is hard to justify unpopular decisions made by a supervisor with different values and goals; therefore, dishonest behavior is likely to emerge to restore fairness outside organizational borders.

As these example suggest, organizational identification might promote dishonest behavior at cost of society when corporate culture is not equally shared within subunits and divisions. In our study, we specifically investigate the decision to react dishonestly after experiencing a supervisor's unfair action, and the extent to which this decision is mediated by her group identity. We depict the typical organizational context in which employees cannot react against the unfair supervisor, but they can engage in dishonest behavior against society, such as by not buying the bus ticket. We address how one's decision to behave dishonestly (e.g., taking advantage from misreporting) is driven by unfair decisions made an individual who shares the same or a different group identity.

Throughout the study, we adopt the concept of *fairness-restorative* dishonest behavior to identify the reaction to unfairness caused by another individual. Individuals might engage in dishonest behavior when they are given the opportunity to increase their material well-being. However, they might be even more willing to act dishonestly when they experience unfair decisions they cannot tolerate.

We investigate how group identity enters the way unfairness is tolerated and, therefore, how it mediates *fairness-restorative* dishonest behavior. We mimic the scenario in which two employees work for the organization, but only one has the power to distribute wages. If the least powerful employee perceives the distribution as unfair, she might feel the need to restore fairness. Since she has no power to restore fairness by directly punishing the counterpart, she might increase her dishonest behavior outside organizational borders.

In our experiment, participants are paired in couples and receive a payoff based on the counterpart's decision in a real-effort dictator game, an ideal setting to impose unfairness on recipients and to induce a shared view of fairness across participants. In particular, prior studies show that dictators allocate unfairly even when they agree on which is the fair allocation (Konow, 2000; Dengler-Roscher *et al.*, 2015) and that the act of exerting effort elicits a shared view of fairness in the proportionality principle (Cappelen *et al.*, 2014). Then, participants are given the opportunity to increase their payoff by engaging in a self-report task which is costly to the experimenter but not to the counterpart. In the baseline condition, participants in the couple are only labeled as recipient and dictator. In the *IN* condition, participants in the couple know that they share the same group identity. In the *OUT* condition, participants in the couple identify with two different groups. To induce group identification, we rely on a modified version of Tajfel *et al.* (1971)'s minimal group paradigm. This way, we investigate recipients' degree of dishonest reactions to dictators' decisions when group identity varies.

Our results show that dishonest behavior is not affected by perceived unfairness when individuals share the same group identity. On the other hand, dishonest behavior is significantly affected by unfair decisions made by individuals with a conflicting group identity. Thus, allowing for the coexistence of diverse group identities within organizations leads individuals to be more sensitive to unfair decisions and to react by increasing their dishonest behavior at cost of society to restore fairness.

Additionally to highlighting a novel mechanism underlying dishonest behavior, our study is of practical relevance to organizations characterized by a decentralized structure. For example, organizations which fail to develop a shared corporate culture across divisions might create the basis for negative externalities on society: employees who cannot tolerate decisions made by a supervisor from a division they find hard to identify with, might be willing to engage in dishonest behavior outside organizational borders to restore fairness. For these organizations, it might be preferable to promote the diffusion of one identity by favoring the interaction between employees from different units and subgroups.

1.2 Relevant Literature and Hypotheses

For being substantially costly to organizations and society, dishonest behavior has received increasing attention by scholars from psychology and economics. The standard economic approach to dishonesty has been shaped by the strict consequentialist logic put forward by the seminal contribution of Becker (1968): people cheat only when the expected benefits of dishonest behavior (e.g., saving money) outweigh its expected costs (e.g., being caught). Experimental evidence shows that the decision to behave dishonestly to maximize earnings is, in fact, influenced by contexts poor of monetary or reputation sanctions (see, among others, Fischbacher & Föllmi-Heusi, 2013; Ploner & Regner, 2013). However, recent research in economics has shown that dishonest behavior is not only driven by expected consequences (Gneezy, 2005; Mazar *et al.*, 2008; Erat & Gneezy, 2012; Abeler *et al.*, 2014), but also by history and context (Gino & Pierce, 2010b; Houser *et al.*, 2012; Shalvi *et al.*, 2015).

Employees are routinely affected by distributive decisions made by other individuals who work for the same organization. They work with the aim of gaining a wage that best reflects their effort. But when a supervisor's choice is perceived as unfair, employees try to restore fairness by indulging in dishonest behavior at cost of others. This is especially true when the employee has no power to react by directly punishing the unfair supervisor. For example, Greenberg (1990) shows that when employees experience unjustified wage cuts, they engage in inventory theft. According to Ambrose *et al.* (2002), when people perceive income distributions as unfair, they engage in sabotage behavior in the attempt to restore equity.

Contextual elements seem to play a fundamental role in shaping fairness perceptions. In particular, they are influenced by the underlying allocation process and the idiosyncratic features of those affected by that allocation (Konow, 2003). Other studies (Gino & Pierce, 2009, 2010a,b; John *et al.*, 2014) report that individuals perceive unfairness in wealth disparities due to different initial endowments and different pay-schemes.

A few studies investigated dishonest behavior as a consequence of experiences of unfairness. In the context of a bargaining game, Ellingsen *et al.* (2009) report that individuals increase their dishonest behavior after experiencing negative actions from their counterpart. In this study, dishonest behavior is costly to the counterpart. Similarly, Alempaki *et al.* (2016) investigate deception as a reciprocity device when individuals experience unkind actions from their counterparts in a dictator game. Houser *et al.* (2012) investigate the decision to cheat after individuals participate in a dictator game. In their experiment, cheating is costly to the experimenter. Our study builds on Houser *et al.* (2012)'s setting. We investigate the decision to engage in dishonest behavior which is costly to the experimenter. Differently from Houser *et al.* (2012), we do not endow participants with windfall money, but we ask them to exert effort as in the case of employees. This way, we mimic the situation in which the least powerful employees are not able to react to unfairness by directly punishing the unfair supervisor, but their dishonest behavior is unlikely to be detected outside the organization while being costly to society.

We add to this stream of research by examining how a salient group identity affects the extent of tolerance of unfair actions. Whether group identity mediates dishonest behavior as a reaction to unfairness has not previously investigated. A recent experimental study suggests that, in the context of communication games, group identity mediates deception (Rong *et al.*, 2016). However, deception is not considered as a device to restore fairness. Employees might be more willing to tolerate unfair decisions when they are made by a supervisor from the same group. This intuition, motivating our research, originates in *Social Identity Theory (SIT)* (Tajfel *et al.*, 1971; Turner, 1985), wherein individuals aim to preserve a positive image of their group members because this is part of their own identity.

Shared group identity has been widely recognized as a means for reducing agency problems and enhancing virtuous behaviors in organizations (Akerlof & Kranton, 2000, 2005, 2008). A bunch of experimental studies show that individuals tend to cooperate more when interacting with others sharing the same group identity (in-group) (Eckel & Grossman, 2005; McLeish & Oxoby, 2011; Weng & Carlsson, 2015). In contrast, when interacting with members of other groups (out-group), individuals display less cooperation (Charness & Jackson, 2007; McLeish & Oxoby, 2007), coordination (Chen & Chen, 2011; Chen *et al.*, 2014) and other-regarding preferences (Chen & Li, 2009). SIT provides a general framework to understand the roots of such inter-group discrimination. When group identity is made salient, the perception of our self-concept changes and also our behavior changes accordingly: we tend to favor the members of our group, while discriminating against those who belong to another group (Balliet *et al.*, 2014).

Studies by Kollock (1998); Goette (2006) and Chen & Li (2009) find that individuals tolerate ingroup unfairness more than outgroup one. In contrast, McLeish & Oxoby (2007, 2011) and Weng & Carlsson (2015) find that ingroup unfairness breeds stronger punishment than outgroup unfairness. Similar to these studies, ours investigates situations in which tolerance of unfair decisions might be mediated by group identity. However, our specific focus is not on reactions against the unfair counterpart, such as second-party punishment, but on a type of reaction that creates negative externalities outside organizational borders: *fairness-restorative* dishonest behavior.

As a measure of fairness of one's behavior, we refer to the proportionality between the amount contributed in the real-effort DG and the amount claimed by the dictator: the closer the amount claimed to the amount contributed, the higher the degree of fairness. Proportionality between inputs and outputs is at the cornerstone of equity theory (Adams, 1965; Homans, 1958; Walster *et al.*, 1973) and of the accountability principle (Konow, 1996). In the following, we refer to this concept of fairness as the *proportionality principle*.

To outline our predictions, we exploit the similarities between the experiment by Houser *et al.* (2012) and our baseline condition in which group identity is absent. In the light of the finding by Houser *et al.*, we predict that individuals are more likely to indulge in dishonest behavior when one's behavior is perceived as unfair. This prediction represents the benchmark against which we assess behavior when group identities are made salient. Thus, we expect one's unfair behavior to determine an increase in another's dishonest behavior. Concerning the conditions in which group identity is made salient, we expect the relationship observed in the baseline condition to be strengthened by conflicting group identities (OUT). Violations of norms from an outgroup member are difficult to be tolerated and justified (Chen & Li, 2009).

Prediction 1. OUT-group

Compared to the Baseline, the stronger the violation of the proportionality principle by an outgroup individual, the higher the likelihood to engage in *fairness-restorative* dishonest behavior.

When individuals share the same group identity (IN), we expect to observe a different pattern. In order to preserve positive beliefs about their group identity and, thus, to avoid cognitive costs associated to self-concept updating, individuals may interpret unfair decisions made by a group member as not unfair. Experiencing unfair decisions made by a group member is likely to prompt self-deception about the real nature of the offer. In particular, individuals interpret unfairness generated within the group through a mentalizing bias (Baumgartner *et al.*, 2013). This leads to the following prediction

Prediction 2. IN-group

Compared to the Baseline, the stronger the violation of the proportionality principle by an ingroup individual, the lower the likelihood to engage in *fairness-restorative* dishonest behavior.

In the next session, we describe the experimental design we adopted to test our hypotheses.

1.3 Method

Our experiment is designed to investigate dishonest behavior as a device to restore fairness when different group identities are salient. At this aim, we rely on the minimal group paradigm (MGP) (Tajfel *et al.*, 1971). We conduct three variants of group identity. In particular, we vary whether no group identity is salient (*BASE*), whether individuals share the same group identity (*IN*), and whether individuals do not share the same group identity (*OUT*).

Participants were recruited to the laboratory in even groups. In all conditions, they were presented with a real-effort task. In the *IN* and *OUT* conditions, participants completed a task aimed to manipulate group identity. Then, participants faced a dictator game aimed at distributing an endowment based on the earnings from the real-effort task. Finally, participants were asked to self-report a number which allowed them to increase their final earnings.

The experiment was programmed and behaviored using z-Tree software (Fischbacher, 2007)[1]. Upon their arrival to the laboratory, participants were randomly allocated to cubicles and asked to privately read the instructions[2]. A member of the staff read aloud the instructions and answered doubts about the experimental procedure. Before starting the experiment, participants had to answer six control questions checking their understanding of the instructions. Participants received on average €9.50 in addition to a show-up fee of €3. Each session lasted on average 1 hour and 30 minutes.

A total of 192 students took part in the experiment. 64 were assigned to the IN condition and 68 to the OUT condition. The remaining 60 participated to the Baseline condition. In all three conditions half participants were randomly assigned to the Dictator role and the other half to the Recipient role.

1.3.1 Stage 1: Real-Effort Task

In the first stage, participants were presented with a task that generates a part of their final earning. The task is a modified version of Gill & Prowse (2012)'s real effort slider task. Participants were asked to position a set of sliders at a correct location on the screen, within 240 seconds. Sliders were presented in blocks of 6, with correct locations randomly defined by the computer. Before starting, all sliders were randomly aligned, to avoid visual learning effects. Participants generated €1 for each block correctly solved.

We adopted this procedure to ensure that participants shared the same perception of fairness. Evidence suggests that the mere act of exerting effort leads individuals to perceive fairness in proportional distributions (Cappelen *et al.*, 2014). While Houser *et al.* (2012) endow dictators with windfall money, we designed this task to prime shared perceptions of fairness in proportional distributions. Throughout the study, we refer to Konow (1996)'s definition of fairness (i.e., *equity* (Adams, 1965; Homans, 1958; Walster *et al.*, 1973)) as the *proportionality principle*: a fair distribution is one proportional to the variables that affect production and that individuals can control (i.e., work effort).

1.3.2 Stage 2: Group Identity

Participants in *IN* and *OUT* conditions were presented with a task based on the MGP. This task enables to make salient a group identity. First, participants were asked to guess a number $\in \{1,2,...,99\}$ randomly drawn by the computer. According to their guess,

[1]Screenshots from the experiment are available in Appendix A.
[2]A translated copy of instructions is available in Appendix A.

they were either allocated to one color group (Red) or to another color group (Yellow). Specifically, those whose guess was closer to the randomly drawn number were assigned to one group and those who were farther to another group. Participants were only told that those in their color group were matched according to the similarity criterion just presented[3]. Second, to strengthen "common fate" feelings - a major constituent of group identity - we asked participants who were assigned the same color to take part in a collective task[4]. Subjects were shown a screen containing a set of unordered pieces of words and were asked to combine them to form a proverb. Participants received an additional €1 if their color group was the fastest in completing the task. To determine which of the two groups was the fastest, individual time records of those in the group were summed up. To test the effectiveness of group identity manipulation, participants were asked to evaluate their perception of similarity with participants affiliated to the same and the other color groups. Answers were reported both on a Likert scale and in a self/other task similar to that adopted by Sani *et al.* (2007).

In the *BASE* condition, participants only played the proverb task and they were told that they had the opportunity to win an additional €1 in the case they were among the fastest half of session participants. All references to group colors were omitted.

1.3.3 Stage 3: Dictator game

In the third stage, each participant was randomly assigned to either the role of dictator or that of recipient and paired with another participant in the other role. Those assigned to the *IN* condition were informed that they belonged to the same group (Red/Red or Yellow/Yellow). Participants assigned to the *OUT* condition were informed that they belonged to different groups (Red/Yellow or Yellow/Red). In these conditions, both players shared common knowledge of group membership. Participants assigned to the *BASE* condition were only informed about their roles. The dictator was asked to allocate between herself and the other the sum that they generated in the *Earnings* stage, which was also of common knowledge by both players.

Our modified dictator game is the ideal setting to induce dictators to make unfair

[3]We did not disclose to participants whether they belonged to the closer or farther group to avoid possible entitlement feelings among those guessing better. Furthermore, we chose a trivial task to avoid potential biases in group composition when the discrimination criterion correlates with unobservable features.

[4] In contexts where group identity is imposed on existing one, as it happens in the laboratory, the salience that commonly categorized individuals are homogeneously treated helps identification (Kramer & Brewer, 1984).

decisions against recipients. Previous studies on allocation choices document that individuals claim more than what they earned when they have a personal stake in the decision outcome (Konow, 2000; Cherry *et al.*, 2002), although they recognize that proportional claims are the fair ones (Dengler-Roscher *et al.*, 2015). Therefore, by claiming non-proportional offers of a co-produced outcome, dictators force recipients into an experience of unfairness.

1.3.4 Stage 4: Questionnaire

The *Questionnaire* stage consists both of non-incentivized self-reported answers and of incentivized answers. For what concerns the former, we asked participants to answer a survey about subjective perceptions of fairness in the allocation task and socio-demographic characteristics. For the latter (*Social Norm Task*), we asked participants to rate in terms of social appropriateness a hypothetical scenario that recalls a dictator game exploiting a task based on Krupka & Weber (2013) [5]. This way, we obtained a measure of the shared perceptions of fairness across participants.

Finally, we provided participants with a self-report task apt to elicit dishonest behavior via untruthful reports (*Dishonesty task*). Previous studies elicited dishonest behavior by looking at self-reported task score (Mazar *et al.*, 2008; Cadsby *et al.*, 2010) or at outcomes of a random event (Bucciol & Piovesan, 2011; Houser *et al.*, 2012; Fischbacher & Föllmi-Heusi, 2013; Ploner & Regner, 2013). We elicited dishonest behavior with a novel method similar to that of Gill *et al.* (2013). They asked subjects to report the last digit of their best friend's number to obtain a more precise distribution of dishonesty degree. We asked to self-report the last digit of the last call they made being aware that they are going to earn €0.50 times the number reported (i.e., maximum earnings are obtained when the last call ends with 9)[6]. While participants were free to report the value without any control from our side, we invited them to check the call list on their mobile phone. This way, we reduced any potential contextual ambiguity that may lead individuals not to perceive that they are lying (Shalvi *et al.*, 2015).

[5] In Krupka & Weber (2013) the group of subjects answering the question was different from the group of subjects that played the dictator game.

[6] The exact payoff rule was presented only in the screen of the computer in concomitance to the task

1.4 Results

We first analyze dictators' allocation choices and provide an assessment of shared perceptions of fairness among participants. Then, we present results of the dishonesty task and a regression analysis inquiring about determinants of dishonest behavior.

1.4.1 Fairness

Figure 1.1 provides a joint representation of claims by the dictators in the *Dictator game* and outcomes in the *Real-Effort* task, in each of the experimental conditions separately.

As the graphs show, dictators' claims largely violate the proportionality principle, with most of the observations lying above the 45° line. This is confirmed also by the intersection between average claims and average contributions, well above the proportionality line in all conditions.

As a measure of the opportunistic stance of dictators, we compute the share of resources of the other appropriated by the dictator. This index of deviation from proportionality is computed as $\phi = \frac{\pi_D - e_D}{e_R}$, where π_D is the actual payoff claimed by the dictator, e_D is the amount earned by the dictator and e_R is the amount earned by the recipient [7].

Table 1.1: Violation of proportionality

	Median	Mean	SD	N
IN	0.464	0.518	0.364	32.000
OUT	0.400	0.388	0.404	33.000
BASE	0.500	0.541	0.356	28.000

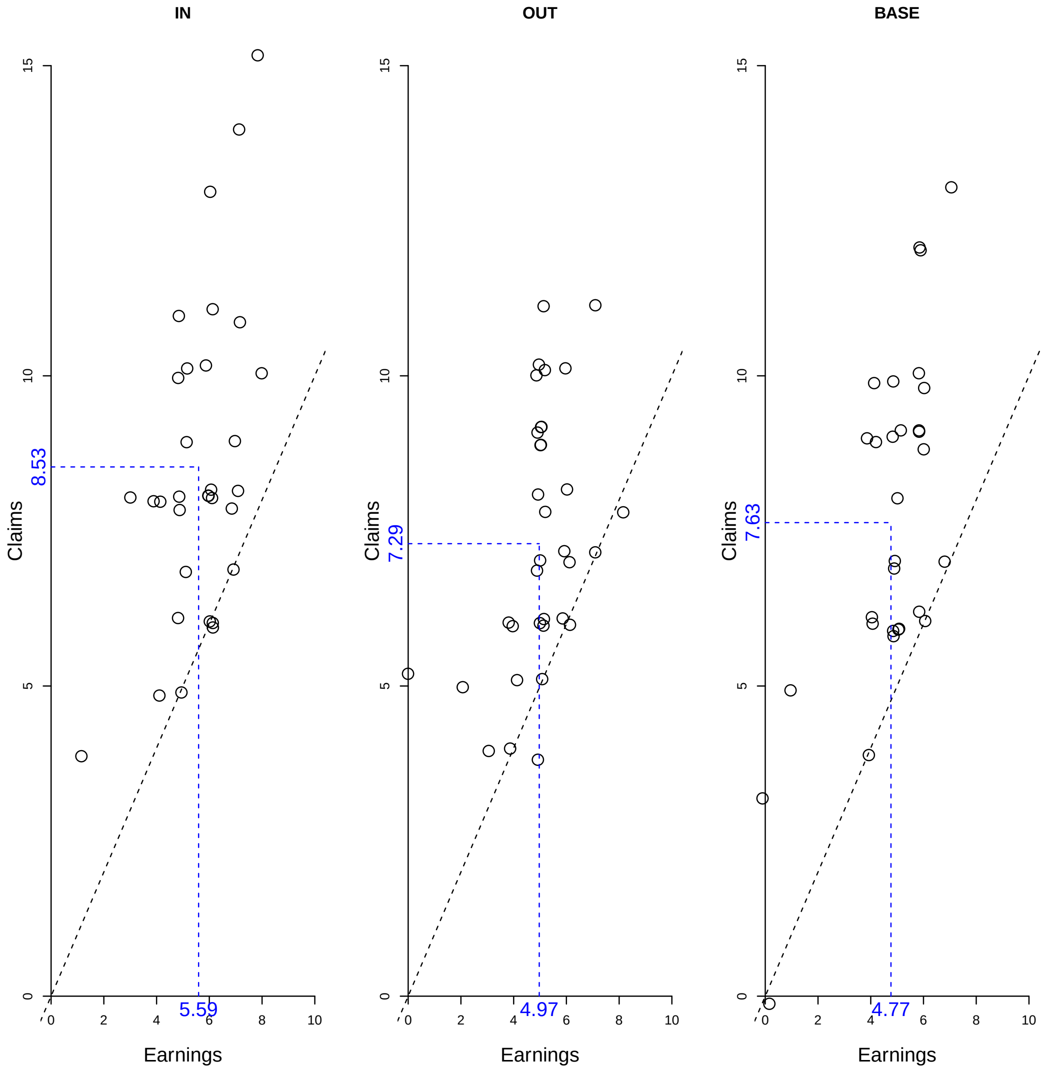
The proportionality index is computed as the amount of other's earnings appropriated relative to other's earnings ($\phi = \frac{\pi_D - e_D}{e_R}$). When $\phi = 0$, the allocation is fully in line with the proportionality principle. When $\phi > 0$, an opportunistic violation of the principle is detected.

As Table 1.1 shows, the largest average (median) violation is observed in condition BASE, followed by IN and OUT. In all conditions the deviations are significantly larger than zero according to a Wilcoxon Signed Rank Test (all p-values < 0.001). A series of Wilcoxon Rank Sum Tests shows that differences across experimental conditions are not statistically significant (all p-values > 0.125).

Dictators largely violate the proportionality principle and appropriate a considerable amount of resources generated by their counterpart. No significant differences in the degree of violation of the proportionality principle are observed across group conditions.

[7] All dictators and recipients exerted a positive amount of effort, i.e. completed at least one page, in the slider task and, therefore, earned a positive amount.

Figure 1.1: Dictators: Contributions and Claims



On the x-axis (Earnings), the amount earned in the slider task. On the y-axis (Claims), the amount claimed in the DG. Points above (below) the 45° line identify claims larger (smaller) than the amount contributed. Average values are reported along the axis. To improve visualization, a small random noise is added to the graph.

This suggests that dictators were not affected by our experimental conditions, in line with previous studies (Konow, 2000). All dictators decided to take more than what they earned, because they were responsible for the allocation decision. By doing so, they forced all recipients into an experience of unfairness.

1.4.2 Perceptions of Fairness

Table 1.2 provides a representation of the shared perception of fairness in the population, as collected in the *Social Norm* task. Participants are presented with a scenario resembling a dictator game and asked to assess the degree of social acceptability of each potential allocation.

Table 1.2: Perceptions of fairness

		Freq (%)			
Allocation	Average	−−	−	+	++
0|6	-0.312	41.1	25.0	23.4	10.4
1|5	**0.420**	**10.4**	**18.8**	**18.2**	**52.6**
2|4	0.250	3.6	17.7	66.1	12.5
3|3	0.299	4.7	22.4	46.4	26.6
4|2	-0.295	22.9	51.0	23.4	2.6
5|1	-0.671	62.0	30.2	4.2	3.6
6|0	-0.861	88.0	6.80	1.6	3.6

The column Allocation reports potential allocations to two individuals, with 1|5 *being the allocation respecting the proportionality principle (bold font). Participants face four assessments for each allocation: "Very unacceptable" (−−), "Quite unacceptable"(−),"Quite acceptable"(+) and "Very acceptable" (++). Frequency of choice for each of the assessment is reported in the table (Freq %). Similarly, to Krupka & Weber (2013), the column "Average" is computed by assigning values -1, -1/3, 1/3, and 1 to the evaluations of acceptability of the allocation in increasing order of acceptability.*

As the table illustrates, the allocation deemed as the most acceptable is the one reflecting proportionality (1|5), with an average of 0.420. All other allocations are judged, on average, unacceptable. This result allows us to rule out the possibility that participants engage in dishonest behavior to restore a different fairness criterion from proportionality, such as inequality.

Table 1.3 reports self-reported measures about the perceived fairness of dictator's choices. Larger values capture a stronger perception of fairness.

As Table 1.3 shows, dictators perceive their choices as fairer than the matched recipient, across all conditions and questions. To test whether perceptions statistically differ, we compute the average at the individual level of the answers to the three questions. When comparing average perceptions of dictators and recipients, a statistically significant difference is observed for condition BASE and IN (Wilcoxon Rank Sum test, both p-values < 0.013), while a marginally significant difference is observed in condition OUT (Wilcoxon Rank Sum test, p-value $= 0.051$). When comparing averages across

Table 1.3: Perception of fairness

Mean(SD)	Dictator	Recipient
	BASE	
Fair	4.133 (2.145)	3.100 (2.369)
Fair | Outcome	3.600 (2.078)	2.633 (2.312)
Fair | Effort	3.800 (1.864)	2.367 (2.042)
	IN	
Fair	4.625 (1.963)	3.125 (2.366)
Fair | Outcome	4.500 (2.125)	2.688 (2.132)
Fair | Effort	4.875 (1.930)	2.375 (1.963)
	OUT	
Fair	3.971 (2.249)	3.353 (2.650)
Fair | Outcome	3.824 (2.355)	2.882 (2.459)
Fair | Effort	3.941 (2.074)	2.676 (2.371)

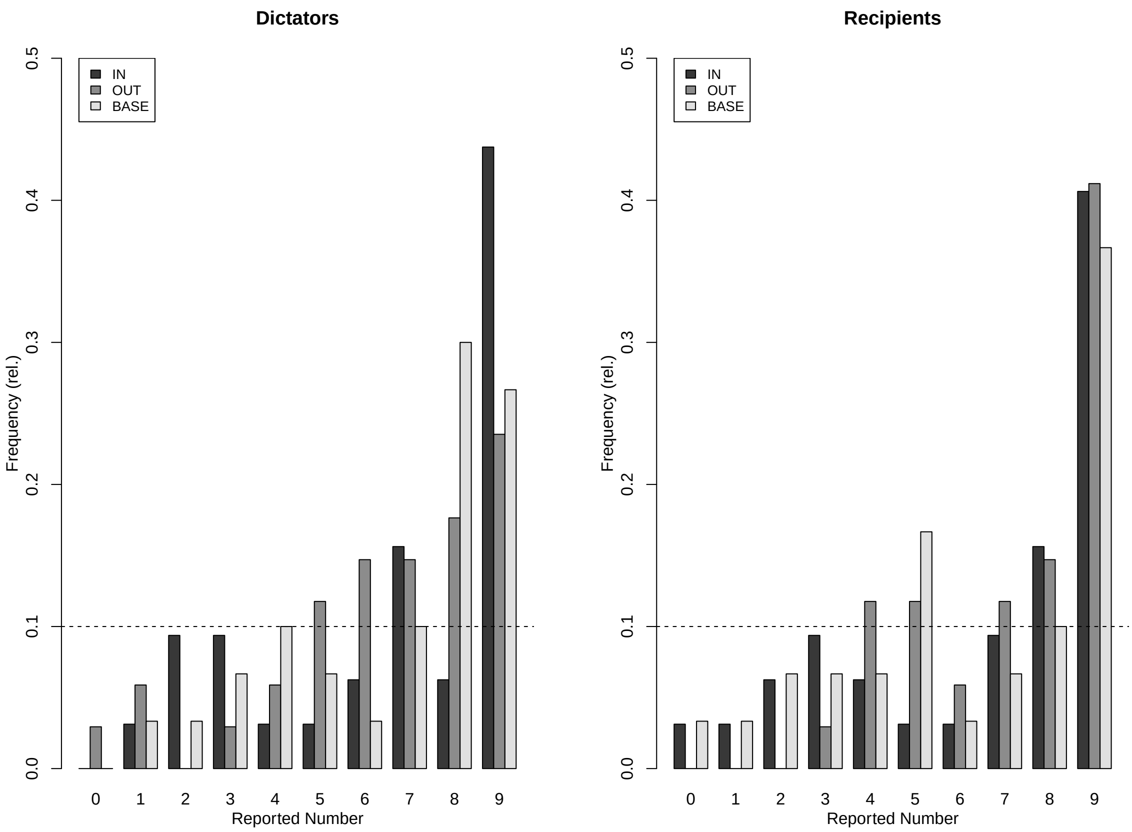
For dictators (recipients), the row labeled Fair *refers to the statement "I feel that I treated the other fairly" ("I feel that the other treated me fairly"). The row* Fair | Outcome *refers to the statement "I feel that I treated the other fairly given the outcome in the slider task" ("I feel that the other treated me fairly given the outcome in the slider task"). The row* Fair | Effort *refers to the statement "I feel that I treated the other fairly given the effort in the slider task" ("I feel that the other treated me fairly given the effort in the slider task"). Answers are collected on a Likert scale 1–7, with 1 meaning "I totally disagree" and 7 meaning "I totally agree".*

conditions given the role, no statistically significant differences are observed (Wilcoxon Rank Sum test, all p-values$\geq$ 0.110).

To gain insights into the source of feelings of fairness, we compute correlations between average feelings of fairness in the statements of Table 1.3 and the proportionality index of Table 1.1 (Spearman's rank correlation). We expect to observe a negative correlation between the index capturing violations of proportionality and perception of fairness. The strongest correlation is observed for the dictators in condition OUT (ρ =-0.830), while the lowest correlation is observed for dictators in condition IN (ρ =-0.418).

When evaluating the fairness content of their actions, dictators seem to adhere to a mentalizing bias aimed at reducing the cognitive dissonance originating in the discrepancy between the—generally acknowledged—fairness norm and the actual—generally selfish—behavior. Accordingly, dictators perceive their actions as fairer than what perceived by their counterpart. Furthermore, dictators in condition IN display a weaker sensitivity to violations of proportionality in terms of fairness perception. Shared group identity would call for fair behavior and when this does not happen, dictators may

Figure 1.2: Self-reported numbers



Self-reported last digit number of the last phone call made by dictators (leftward panel) and recipients (rightward panel), in each of the three experimental conditions separately. The dashed horizontal line provides a reference for the ideal uniform distribution of values. To improve visualization, a small random noise is added to the graph.

reduce their discomfort by re-assessing the fairness of their behavior.

1.4.3 Dishonesty

Figure 1.2 shows a description of behavior in the dishonesty task embedded in the questionnaire. Larger numbers are associated to higher gains.

As the figure shows, the distribution of reported values is negatively skewed, with the mass of the distribution shifted towards higher values and a spike at the highest admitted value. A series of $\tilde{\chi}^2$ tests confirms that reported values are not uniformly distributed (all p-values< 0.038). The tendency to self-report high numbers is testified also by averages of the distributions, with values ranging from 6.233 (Recipients, BASE) to 7.206 (Recipients, OUT). The central tendency of the distributions is significantly larger than the expected average value of 4.5, in all three conditions and for both roles (Wilcoxon Signed Test, all p-values < 0.001).

A comparison of the distributions across experimental conditions for each role shows that there are no significant differences in the central tendencies of the distributions (Wilcoxon Rank Sum Test, all p-values > 0.253). When given the opportunity, participants dishonestly manipulate reported numbers to increase their earnings.

1.4.4 Is dishonesty influenced by unfair decisions?

Individuals might report dishonestly not only because of low monitoring, but also because they do not tolerate the unfairness experienced. In particular, we predict that recipients' dishonest behavior is affected by the unfair dictator's group identity. At this aim, we look at the relationship between reported values and allocations across group conditions and roles.

Figure 1.3 shows that for dictators, a positive correlation between unfairness and reported values is registered in all conditions, though the relation is not statistically significant (Spearman's rank correlation, all p-values≥ 0.113). This suggests that dictators were not affected by the group identity manipulation, but they behaved as unfairly as dishonestly throughout all conditions [8]. For the recipients, a negative and slightly positive correlation is observed in BASE and IN, respectively. However, these correlations are not statistically significant (all p-values ≥ 0.581). In contrast, in condition OUT a positive and marginally significant correlation is observed (p-value=0.080).

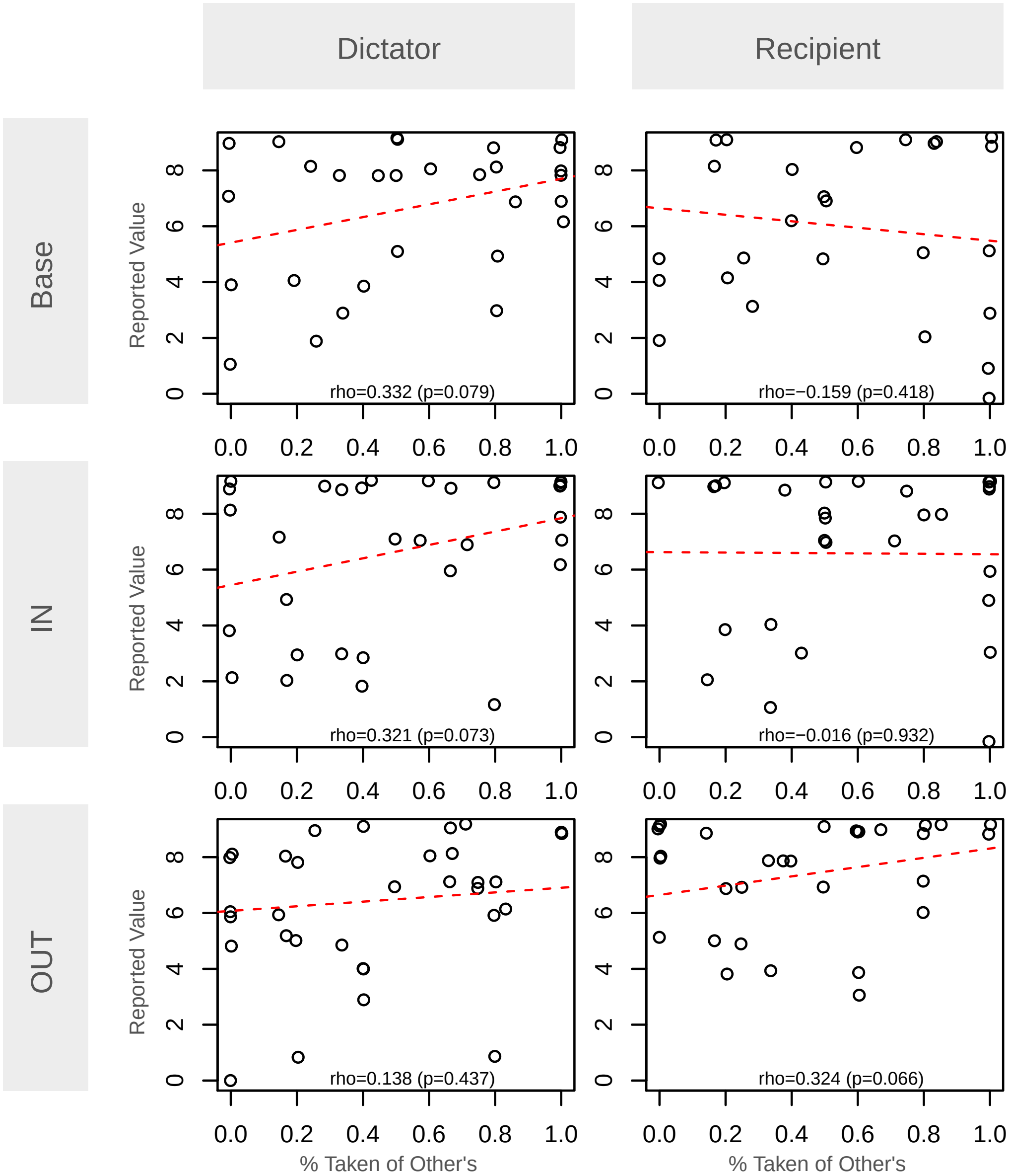
To address the causal relationship between unfairness and dishonest behavior given different types of salient group identity we run a regression model [9]. Table 1.4 reports on the outcomes of a regression estimate about determinants of dishonest behavior of those acting as recipients in the *Dictator game*. The dependent variable $Reported.value$ is given by the integer reported in the dishonesty task described above[10]. When reports are

[8] In Appendix A, a discussion of results of an Ordered Probit Regression model of dictators' dishonest behavior supports this insight.

[9] We relied on the randomness of unfair claims by individuals who have been randomly assigned to the dictator role.

[10] We employed an Ordered Probit model to account for the heterogeneous attitude individuals display when they report increasingly high values. Individuals might care about maintining a positive view of their self-concept (Mazar *et al.*, 2008) or they might suffer from intrinsic costs when they lie (Abeler *et al.*, 2014). Reporting average values might allow to more easily preserve own self-concept or manage lying costs than reporting the highest values. Therefore, the Ordered Probit model allows us to account for

Figure 1.3: Self-reported numbers and unfairness



On the x-axis, the proportionality index $\phi = \frac{\pi_D - e_D}{e_R}$. *On the y-axis the value reported in the Dishonesty task. For dictators, higher values on the horizontal axis imply more unfairness favoring them. Conversely, for recipients higher values on the horizontal axis imply more unfairness damaging them. In each panel, a Spearman's rank correlation* ρ *is presented.*

truthful, the dependent variable and explanatory variables are orthogonal. When this is

conceptual differences between reporting average values, such as 4 and 5 (i.e. mild dishonesty), and high values, such as 8 and 9 (i.e. brazen dishonesty) not the case, we obtain evidence of distorted behavior.

Among explanatory variables, we have dummy variables controlling for group identity conditions: IN is equal to one when the recipient belongs to the same group and zero otherwise; OUT is equal to one when the recipient belongs to the other group and zero otherwise. BASE is the baseline condition.

We enrich the model with fairness-related variables: the proportionality index ϕ ($prop.index$, see Table 1.1) of matched dictator's choices and the extent according to which proportionality is perceived as appropriate in the *Social Norm* task ($prop.norm$, see Table 1.2). The interactions between the fairness-related variables and the group experimental conditions are also added. Finally, we consider a few control variables: $report.time$ (the time in seconds required to report the value); age; $female$, and $civic.score$ (a categorical variable of individuals' participation to collective activities, such as political parties and NGOs).

Table 1.4: Recipients' self-reported values (Ordered Probit regression)

	(1) $rep.value$	(2) $rep.value$	(3) $rep.value$
$prop.index$	0.011(0.249)	-0.57(0.543)	0.648(0.545)
$prop.norm$	-0.187(0.244)	-0.637(0.489)	-0.681(0.493)
IN		-0.365(0.645)	-0.466(0.65)
OUT		-0.781(0.671)	-0.725(0.673)
$prop.index \times IN$		0.427(0.645)	0.559(0.667)
$prop.index \times OUT$		1.486(0.489)*	1.60(0.769)**
$prop.norm \times IN$		0.352(0.628)	0.374(0.639)
$prop.norm \times OUT$		0.837(0.647)	0.689(0.659)
$report.time$			0.00398(0.007)
$female$			-0.2689(0.249)
age			-0.0144(0.045)
$civic.score$			0.013(0.111)
Observations	93	93	93

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

An Ordered Probit regression model is adopted to account for the different attitude individuals display when they report increasingly high values.

As Table 1.4 shows, our explanatory variables cannot systematically explain the value reported. The only exception is observed for the interaction term between $prop.index$ and OUT. Both Model (2) and (3) show that in condition OUT a stronger violation of proportionality by the dictator positively impacts on the reported value. This does not happen in BASE and IN [11].

The regression analysis shows that experiencing higher unfairness induces more dishonest behavior only in condition OUT: when dictators and recipients belong to different groups, dishonest behavior significantly increases to restore fairness.

1.5 Conclusions

In this study, we investigated dishonest behavior as a way to restore fairness when different types of unfair individuals' group identity are salient. We ran a laboratory experiment to mimic the situation in which two employees work for the same organization but only one can control the joint allocation of rewards. In this situation, the employee facing unfairness cannot directly react against the unfair supervisor, but she can react opportunistically outside the organization, without any credible threat of being sanctioned. We find that unfair individuals' group affiliation significantly affects the likelihood of undertaking dishonest behavior to restore fairness, with conflicting group affiliations fostering fairness-restorative dishonest behavior.

To investigate dishonest behavior as a reaction to unfairness we needed to ensure that participants shared the same perception of fairness. Consistent with previous research (Cappelen *et al.*, 2014), we find that the real-effort task was effective to prime perceptions of fairness in the proportionality principle. This was confirmed by responses to the *Social Norm task*.

Crucial for addressing our hypotheses was designing a setting that allowed recipients to experience unfair decisions. Consistent with previous findings (Konow, 2000; Dengler-Roscher *et al.*, 2015), we observed that dictators allocated to themselves more than what they believe is fair. Across all group conditions the index of proportionality revealed that allocations were self-oriented. Therefore, our modified dictator game was an effective setting to impose experiences of unfairness on recipients and to prepare a propellant for *fairness-restorative* dishonest behavior.

Consistent with previous research on dishonest behavior, we find that individuals behave dishonestly to increase their earning when they are given the opportunity (Fischbacher & Föllmi-Heusi, 2013; Ploner & Regner, 2013). Although across all conditions responses to the survey on perceptions of experienced unfairness confirmed that re-

[11] Results are confirmed by a Tobit model to account for potential censoring in the data from the dependent variable - bounded between 0 and 9.

cipients acknowledged that dictators allocated unfairly, dishonest behavior increases only when individuals experience unfair decisions made by an out-group member. Regression results suggest that group identity enters the way unfairness is tolerated and, thus, moderates *fairness-restorative* dishonest behavior. Recipients were more willing to engage in dishonest behavior to restore fairness after experiencing unfair decisions made by an outgroup member. Contrary to our prediction, unfairness was irrelevant to purge dishonest behavior when recipients received unfair allocations from an ingroup member compared to when recipients received unfair allocations from a dictator without a specified group identity.

Why should recipients react to unfairness when they are matched with an outgroup while refraining from the attempt to restore fairness when they are matched with an ingroup? The explanation lies at the roots of SIT. Conditional on which group identity is salient, individuals engage in different mentalizing processes of the norm violation (Baumgartner *et al.*, 2013). Recipients interacting with an unfair dictator with a different group identity have no reasons to put themselves in her shoes. They do not need to preserve their beliefs about their group identity and, thus, mentalize toward the unfair dictator. Instead, they perceive norm violations intolerable and attempt to restore fairness by becoming more inclined to dishonest behavior. In contrast, when a shared group identity is salient, recipients easily mentalize toward unfair dictators to maintain intact their beliefs about the group. A shared group identity would call for fair allocations, and when this is not the case, recipients may avoid potential discomfort by justifying dictator's behavior.

In contrast with Houser *et al.* (2012)'s finding, we did not observe dishonest behavior motivated by unfairness in the Baseline. This might explain why we do not find that ingroup recipients engage in less *fairness-restorative* dishonest behavior than recipients in the Baseline. Our intuition is that the fact that both the recipient and the dictator exerted effort, made salient the idea of *asset legitimacy* (Mittone & Ploner, 2012) and mediated reactions to unfairness. Recipients in Baseline accept unfairness from a dictator because they acknowledge the additional effort exerted by the dictator to decide how to allocate the co-produced endowment.

Our study shows that group identity mediates *fairness-restorative* dishonest behavior and provides us with some insights into the working of organizations. Organizations may try to improve the productivity of employees by introducing competitive payment schemes (i.e., tournament incentives). Schemes of this kind may potentially promote perceptions of unfairness that, in turn, lead individuals to restore fairness through dishonest behavior.

In this study, we addressed the hidden danger of allowing for the coexistence of conflicting group identities: this latter is likely to prompt extreme sensitivity to unfair distributions that translates in determining negative externalities outside organizational borders.

When managers fail to limit local identification in each subunit and division, they might promote employees' dishonest decisions that are costly to society: when employees are imposed unfair decisions by a supervisor from a division with different values and goals, they are likely to perceive unfairness and to correct it by substantially increasing their dishonest behavior outside organizational borders.

CHAPTER 2

Perceptions of others' Risk Attitude and Gender Stereotype

with Caleb Cox - Virginia Commonwealth University
and Morten I. Lau - Durham Business School and Copenhagen Business School

Individual risk attitudes and perceptions of others' risk attitudes might influence many economic outcomes, such as occupational segregation and job promotions. We conduct an incentivized laboratory experiment to assess how individual risk attitudes and perceptions of others' risk attitude are biased by gender stereotype. We structurally estimate coefficients of risk attitudes under the Expected Utility and Rank Dependent models of choice and find no gender difference in the curvature of the utility function. We find evidence of gender stereotype bias: males use gender information for stereotyping others' risk attitude.

Keywords: Risk Preferences; Gender Stereotype; Laboratory Experiment
JEL classification: C91; D81; Z1

2.1 Introduction

Different individual risk attitudes and different perceptions of others' risk attitudes are potential sources of conflict in many social interactions. In particular, if females (males) are wrongly perceived to be more risk averse (taking) than males (females), females or males may be disadvantaged in situations where incorrect perceptions of their risk attitudes have an influence on the decision making process. As an example, employers might base corporate promotions on gender stereotypes if the candidate's gender is known to the decision maker (Wyer Jr & Wyer Jr, 1998). Furthermore, when inaccurate beliefs are internalized, they might channel decisions involving risk, such as occupational choices, and lead to selection into jobs that require less (more) risky decisions (Akerlof & Kranton, 2000).

We conduct a laboratory experiment with students at Durham University and elicit their individual risk attitudes and beliefs about the risk attitudes of other students who participated in the experiment. The experiment consists of two parts. In the first part, we use an existing experimental design and elicit individual risk attitudes over real monetary outcomes. A set of 40 lottery pairs allows us to study several dimensions of individual risk attitudes and compare Expected Utility Theory (EUT) to the alternative Rank-Dependent Utility Theory (RDU). This set of decision tasks is based on the lotteries in Harrison & Swarthout (2012, 2014) and Harrison *et al.* (2015).

Whether females and males differ in their risk attitude is still debated in the literature. Therefore, we disentangle gender difference in risk preferences and address whether the shared belief that females are more risk averse than males is a stereotype. In the second part of the experiment, we use the same set of decision tasks and elicit subjective beliefs with respect to risk attitudes of other participants in the experiment, where the subjects are informed about the gender of the other (target) person. This allows us to investigate whether individuals use gender as an information device to form beliefs about others' risk attitude. Particularly, perceptions of others' risk attitudes may be biased by observable characteristics, such as gender (Ames, 2004). Therefore, by providing information on the gender of the other person we investigate stereotype bias.

Our study adds to the literature on perceptions of others' risk attitude and gender difference in individual risk preferences, first, by eliciting risk behavior and beliefs in fully incentivized tasks and, second, by estimating risk attitude coefficients underlying non-EUT latent models of choice. While we do not find gender difference in risk preferences, we find that individuals use gender to predict others' risk attitude. In particular, we find that perceptions of others' risk attitude are biased by the inaccurate stereotype that females are more risk averse than males.

2.2 Existing literature

Surprisingly little work has been done on the association between individual risk attitudes and stereotyped perceptions of others' risk attitudes in controlled laboratory settings. Hsee & Weber (1997, 1998, 1999) were the first to address this issue. They present subjects with seven hypothetical decision tasks in which they are asked to select a sure amount or a lottery with a 50-50 probability distribution over two outcomes. Subjects are then asked to predict the choices made by others in similar decision tasks, where the reference to others differs across the three studies, from "somebody somewhere in the US" to "an average student on campus" to "the individual sitting next to them." One concern in this study is that the decision tasks are not incentive compatible [1]. The subjects are not paid according to their decisions, and both observed risk attitudes and subjective beliefs may be influenced by hypothetical bias [2].

This early work provided inspiration for later studies with incentive compatible decision tasks, and two elicitation methods in particular have been used to elicit individual risk aversion and stereotyped beliefs: in the Multiple Price List (MPL) design the subject is presented with an ordered array of binary lottery choices, whereas in the Ordered Lottery Selection (OLS) design the subject chooses one lottery from an ordered set of several lotteries.

In the MPL, popularized by Holt & Laury (2002), subjects are presented with an array of binary choice tasks between two lotteries, where each lottery typically has two outcomes. The subjects are asked to select one of the two lotteries in each decision task and one decision task is randomly selected for payment. This method has been widely

[1] Subjects are paid $50 in the third experiment by Hsee & Weber (1997) if all predictions are accurate.

[2] Other studies have used hypothetical decisions tasks to elicit individual risk attitudes and stereotyped beliefs with respect to the risk attitudes of others.

Siegrist *et al.* (2002) use the same design as the first and second experiment in Hsee & Weber (1997), but instead of referring to "most other students in the US" they also look at gender differences and refer to "most other female students in the US" and "most other male students in the US".

Daruvala (2007) employs a modification of the Becker *et al.* (1964) procedure (BDM). Although the elicitation method in theory is an incentive compatible measure, there is some concern about the cognitive efforts that are required to understand the decision tasks, even if they are hypothetical (Holt & Laury, 2002; Plott & Zeiler, 2005). Daruvala (2007) conducts a laboratory experiment in which the participants are asked to state the certainty equivalent of a lottery with two outcomes, and then asked to predict the certainty equivalent of the same lottery for each participant in the room. While no gender difference is found in this study, there is some evidence of gender stereotype bias in predictions of others' risk attitude.

used to elicit individual risk attitudes since the reward structure is transparent and simple, and one can easily derive measures of absolute and relative risk aversion from observed choices. Despite the simplicity of the elicitation method there is some concern about possible framing effects in this design, although it is possible to identify and correct for possible framing effects (Andersen *et al.*, 2006).

The MPL was adopted by Roth & Voskort (2014) who conducted an artefactual field experiment with senior and junior financial advisors and students. First, they present subjects with the MPL task. Then, they ask subjects to predict the choices made in the same MPL task by some profiles of respondents to a web-based survey [3]. They can base their predictions on information about the profile which varied according to two conditions. Subjects assigned to the "Rank" condition are, first, asked to assign a preference ranking to each of the profile's attributes and, then, to predict the risky choices of four profiles. Subjects assigned to the "Pay" condition are, first, asked to pay for disclosing each profile's attributes and, then, to predict the risky choices of four profiles. The prediction task in both conditions is incentivized [4]. The authors find that on average professionals are more risk seeking in their decisions than when making predictions on the profile. The authors report evidence of stereotype bias. Subjects use available information about the profile to form their beliefs. Particularly, they find that predicted number of safe choices increases when the profile is female [5].

[3]Web-survey respondents are recruited via E-mail and are asked to circulate the survey. They are presented the same MPL task used in the experiment. However, they are not paid according to their choices. The authors mention that participants who complete the web-based survey are given the opportunity to win €50. From the web-survey the authors create eight profiles, each characterized by a different combination of seven attributes, such as sex, age, marital status, level of education, number of children and income.

[4]In the "Rank" condition, for each profile the computer assigns a random number to the preference ranking and determines the number of attributes to disclose to subjects. In the "Pay" condition, for each attribute and for each profile, subjects choose which characteristics are uncovered by paying according to a scoring rule (e.g., the first attribute costs €0.01 and all attributes cost €0.99). In both conditions, if the prediction for a profile is correct the subject receives €0.50.

[5]In addition to the MPL, Roth & Voskort (2014) adopt another incentivized mechanism to elicit risk preferences and beliefs about risk attitude of others. They ask subjects to hypothetically invest up to €100.000 into a lottery that doubles or halves the amount invested with 50/50 probability. To provide incentives, the high stake amounts are divided by 4 for the actual payment. Then, they ask subjects to predict the amount invested by each profile. As for the MPL, web-survey respondents are presented with the same investment question. However, they are not paid according to their choices. According to the "Rank" and "Pay" conditions, subjects can base their predictions on available attributes of each target profile. If the prediction for a profile is correct, subjects received €0.50.

To define prediction errors, they use responses to the investment question by respondents to another survey (the German Socioeconomic Panel) which contains 20,750 subjects. However, the authors do not mention whether in the Panel the investment question is incentivized. First, they compute representative counterparties for the combination of the seven attributes of each profile. Then, they define prediction errors as the squared difference between predicted amount not invested and the mean amount not invested

The second method adopted to derive stereotyped beliefs over others' risk attitudes is the OLS developed by Binswanger (1980, 1981). In particular, Eckel & Grossman (2002, 2008) use an OLS-based design in which subjects are presented with a choice of five lotteries and asked to pick one. Each lottery has two possible outcomes, each with a 50-50 chance. While the instrument is easy to understand, it restricts the probabilities to $\frac{1}{2}$. This way, it does not allow to make inferences about non-EUT models of choice behavior, such as probability weighting. To investigate predictions of risk attitude of others, they conduct an experiment in which they ask 261 students to predict each other's choices. Subjects are distributed across loss and gain treatments and are seated at tables with unobstructed views of all session participants. Subjects are asked to select one lottery from a set of five gambles. Then, they are asked to predict how each participant in the session would choose in the same task. First, they report gender difference in risky choices: on average males choose riskier alternatives than females in all treatments. Then, they report gender difference in predicting risk attitudes of others: males are predicted by both males and females to be less risk averse than females. Finally, they do not find gender difference in the accuracy of predictions made by females and males.

Other studies took inspiration from Eckel & Grossman (2002, 2008)'s design to investigate stereotyped predictions of others' risk attitude. Ball *et al.* (2010) recruit 182 students to investigate how predictions of risk attitudes of others are influenced by targets' physical appearance. Subjects are first asked to select a gamble from a set of six alternative gambles and, then, to predict the gamble chosen by each session participant. Each subject is also asked to use a dynamometer to collect a measure of individual physical prowess. They find that predictions reflect gender stereotype. Particularly, they find that observable characteristics associated to gender stereotyping, such as attractiveness for females and strength for males, affect predictions of females' and males' risky choices.

The study by Grossman (2013) is also an adaptation of Eckel & Grossman (2002, 2008)' s design. They recruit 90 students to predict each session participant's risky choice. First, they ask to select a gamble from a set of six alternative gambles. Then they ask to predict the risky choice of each subject in the session twice. Subjects are randomly assigned to two conditions. Those assigned to the "Visual/Info" condition

by the representative profile. By running a regression model with the prediction error as dependent variable and available attributes as independent variables, the authors claim that prediction accuracy increases when more information is available. A major concern associated to this conclusion is the bias derived from the Panel responses, since these have not been incentivized. Experimental subjects' prediction error is computed with Panel responses as representative counterpart for each target profile, therefore prediction accuracy is likely to embed the bias from the non-incentivized Panel responses.

are asked to predict the gamble chosen by each session participant the first time by having unobstructed view of each participant; next, they are told they can revise the first predictions after being provided each target's responses to a question about social risk and financial statement. Those assigned to the "Info/Visual" condition are asked to predict each sessions' gamble choice only being provided each target's responses; then, they are told they can revise their prediction by directly observing the target. While they do not find that choices differ by treatment, they find evidence of gender stereotype in predictions of others' risk attitude when subjects are provided with the visual clues. Particularly, predictions in the "Info/Visual" condition are significantly revised as more risk taking for males and more risk averse for females.

Previous studies inferred gender difference in risk preferences and stereotyping in predictions of others' risk attitude from observed choices. However, none directly estimated the risk attitude coefficients underlying a latent model of choice in a prediction task. Chakravarty *et al.* (2011) structurally estimate risk attitude, however, they investigate how individual risk preferences affect risky choices made on behalf of others, and not how the formers affect predictions of risk attitude of others. Subjects are presented with two MPLs in which they have to make choices having consequences on themselves and on behalf on others. They find a widespread higher risk aversion when choices are made on behalf of others. Only as control for the assessment of decisions made on behalf of others, they ask participants to predict the average risky choice in the room both when payoff have consequence on the decision maker and when the choice is made on behalf of others. They report that average risk attitude on one's money does not differ from beliefs about average risk attitudes on one's own money. The closest study to ours is that of Conte *et al.* (2016). They adopted a maximum simulated likelihood approach to estimate risk preferences under Expected Utility Theory and the alternative Rank-Dependent model. However, they investigate how risk preferences are affected by emotions, and not how gender affects both risk preferences and perceptions of others' risk attitude. Subjects are presented with 100 pairwise choices based on Hey (2001). Then, subjects are randomly assigned to one of five treatments in which different types of emotions are induced. They find females to be more risk averse than males and that different emotions drive gender difference in risk preferences, i.e. sadness for males and joviality for females.

The evidence from experimental designs is not clear-cut with respect to gender difference in risk preferences and to the role of gender in perceptions of others' risk attitude. Moreover, conclusions about prediction accuracy also depend on the type of risk measure considered. Second, experimental studies addressing the issue of predictions of others' risk attitude have not considered the direct estimation of a structural model of a latent choice process that best describes observed choices. To the best of our knowledge, this is the first study addressing predictions of others' risk attitude by structurally specifying utility and probability weighting function underlying EUT and non-EUT latent choice processes. We employ the Random Lottery Pair Design to have multiple observations for the same subject. This provides a richer dataset to characterize individual risk attitude and compare it with predicted ones for both sexes after controlling for unobserved heterogeneity [6].

2.3 Experimental Design

The experiment was conducted with 115 students at Durham University in spring 2015. There are two sets of decision tasks. The first set is used to elicit individual risk attitudes, and the second set is used to elicit subjective beliefs about the risk attitudes of a randomly selected subject in the experiment when only the gender of that person is known. These two sets of decision tasks are similar and the subjects are paid according to their choices in each set. The design allows us to test for gender difference in risk preferences and stereotype bias in predictions of others' risk attitude.

2.3.1 Eliciting individual risk attitude

To elicit individual risk attitudes, we employed a design based on the Random Lottery Pair design developed by Hey & Orme (1994), albeit with different decision tasks. The subjects were presented with 40 decision tasks in which they were asked to make a choice between two lotteries, where 25 of those decision tasks included compound lotteries. The employed payment protocol was the Random Lottery Incentive Mechanism: one decision task was chosen at random and the preferred lottery was played out for payment in cash at the end of the session.

The 40 decision tasks are based on those in Harrison & Swarthout (2012, 2014) and Harrison *et al.* (2015). Each lottery has between one and three monetary prizes from a set of seven possible outcomes of £0, £1, £2, £3.5, £4, £7 and £14. The probability of each outcome in the lotteries is $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ or 1. The 40 decision tasks are divided between

[6] Our study relates to that by Conte *et al.* (2016) in the structural estimation of a latent model of choice. However, our procedure for estimating risk attitude is the Maximum Likelihood Estimation approach and not the Maximum Simulated Likelihood approach.

15 simple-compound lottery pairs (S-C), 15 simple-actuarially equivalent lottery pairs (S-AE), and 10 actuarially equivalent-compound lottery pairs (AE-C) [7]. With this design we obtain multiple observations for each subject and we can estimate a latent structural model of choice that describes observed choices [8]. We identify individual risk attitudes using the econometric specification described in Harrison & Rutstrom (2008).

Studies investigating predictions of risk attitude of others have not considered the estimation of a model of the underlying (latent) choice process that best describes risky choices. In our study, we consider both Expected Utility Theory (Von Neumann & Morgenstern, 1944) and the Rank-Dependent model (Quiggin, 1982). RDU captures the idea that subjective perceptions of probabilities may be different from actual probability distribution of random outcomes, even when the actual probability distribution is known to the individual (Kahneman & Tversky, 1979).

Gender difference in behavior under risk might depend not only on valuations of outcomes but also on perceptions of probabilities. In particular, gender difference in emotions - higher fear for females and higher anger for males - might explain gender difference in probability optimism and pessimism (Croson & Gneezy, 2009). A study related to this issue is Fehr-Duda *et al.* (2006): in estimating parameters of the choice model of Prospect Theory by eliciting certainty equivalents, they found evidence of gender difference in how probabilities are assessed: females are more susceptible to probability weighting than males.

In our study, we investigate the existence of gender differences both in the degree of utility curvature and the degree of probability pessimism and optimism, and we characterize behavior in both risky choices and predictions of others' risky choices under EUT and RDU.

2.3.2 Eliciting beliefs about risk attitude of others

In the second set of decision tasks, subjects were asked to predict the choices of a randomly chosen subject who participated in another session of the same experiment. Particularly, they were asked to predict the choices made by one among ten students from the same university - five males and five females - who participated in the first session. Each target was assigned a random number, even for female and odd for male. In order to determine the prediction target, participants were asked to roll a ten-sided

[7] The prospect parameters are in Appendix B

[8] This design allows us also to collect observations useful for a follow-up aimed at investigating gender differences in violations of the Reduction of Compound Lotteries (ROCL) axiom.

die. For instance, if a subject rolled the number 4, she knew that the target subject was a female student whose risky choices were observed in a previous session. Each of the ten target subjects' answer sheets were kept in separate envelopes with the target subject's number visible on the envelope. Subjects were not given any information about the target other than the target's gender. Thus, gender was the only signal to form beliefs about the target's risky choices.

In order to elicit beliefs about risk attitudes of others, subjects were told that, at the end of the task, one choice would be randomly selected. If their prediction was correct, they would be paid £5, otherwise nothing. Participants could verify whether or not they matched the target subject's response by looking at the answer sheet contained in the envelope given before starting the prediction task. The envelope was kept closed until the end of the task. To control for potential order effects, in six of the 14 sessions we reversed the order of the risk and the prediction tasks. Our instructions described the typical scenario each subject would face to make risky choices [9].

Predictions of others' risk attitude can be biased by available information of a target. In particular, individuals' behaviour can be driven by the salience of peculiar features which activate the heuristic thinking of stereotypes (Wyer Jr & Wyer Jr, 1998). This is the case of gender that is associated to devaluing beliefs (Goffman, 1963), such that females (males) are more risk averse (taking) than males (females). Therefore, if individuals use gender as a shortcut to assess others' risk attitude, we expect to see predictions made by both males and females to be biased by gender stereotype (Ames, 2004).

Prediction

If individuals use gender to form predictions of others' risk attitude, females (males) will be predicted to be more risk averse (taking) than males (females).

Among the 115 participants (Female: 67%, sd: 0.472; Age: 22.14, sd: 2.13) 10 were recruited for the first session needed for creating the control group [10]. All participants received a £5 show up fee. At the end of the experiment, subjects completed a survey to collect demographic characteristics [11]. There were 14 experimental sessions in total and each lasted approximately 1 hour. The average payment was £11.

[9] Instructions for both tasks are in Appendix B

[10] Subjects were recruited using ORSEE (Greiner, 2015).

[11] The survey is available in Appendix B.

2.4 Statistical Model

To estimate risk attitudes two methods can be adopted. One involves the calculation of bonds implied by observed choices and exploits utility functions with one parameter to estimate. The other directly estimates a structural model of a latent process of choice with a maximum likelihood estimation of the parameters of risk attitudes specific to the choice model considered (Harrison & Rutstrom, 2008). We adopt this latter approach to characterize risk attitudes under models of choices alternative to the Expected Utility Theory. We initially estimate risk attitudes under EUT because it is parsimonious. Then, we characterize risk attitudes under RDU to capture probability pessimism and optimism.

2.4.1 Expected Utility Theory

We assume a Constant Relative Risk Aversion (CRRA) utility function, defined as

$$U(y) = \frac{y^{(1-r)}}{(1-r)}$$

where r is the parameter to be estimated, and y is income. We estimate the utility function by using maximum likelihood and a latent Expected Utility Theory (EUT) structural model of choice. Let there be K possible outcomes in a lottery - in our case $K \leq 3$. Under EUT the probabilities for each outcome k, p_k represent the objective probabilities induced by the experimenter. The expected utility is simply the probability weighted utility of each outcome in a lottery:

$$EU_i = \sum_{K=1,K} (p_k \cdot u_k)$$

Participants choose between two lotteries (left and right). Therefore, the latent index for each lottery pair from which we estimate r is:

$$\nabla EU = EU_R - EU_L$$

This index is based on latent preferences and is linked to the observed choices using a standard cumulative normal distribution function $\Phi(\nabla EU)$. The probit function transforms any argument between $\pm\infty$ into 0 and 1.

The probability that lottery R is chosen is:

$$\text{Prob(choose lottery R)} = \Phi(\nabla EU)$$

The likelihood of the observed responses, assuming that individuals choose according to EUT and CRRA utility, depends on the estimates of r given the above statistical specification and the observed choices. The log-likelihood is

$$lnL(r;y,X) = \sum_i [(ln\Phi(\nabla EU) \cdot I(y_i = 1)) + ln(1 - \Phi(\nabla EU)) \cdot I(y_i = -1))]$$

where $I(\cdot)$ is the indicator function, $y_i = 1(-1)$ is the choice of the right (left) lottery in task i, and X is a vector of observable characteristics, such as gender and treatments. The parameter r is modelled as a linear function of the covariates in X. In the model, we include the possibility that subjects make the behavioral error of not choosing a lottery when its EU is larger than the EU of the other lottery. In particular, we include the error specification proposed by Fechner and popularized by Hey & Orme (1994). The latent index becomes

$$\text{Prob(choose lottery R)} = \Phi(\frac{\nabla EU}{\mu})$$

where $\mu > 0$ is a noise parameter that reflects the fact that subjects state their preferences with some error. The likelihood specification maximizes $lnL(r,\mu;y,X)$ and we estimate r and μ given the observations on y and X.

2.4.2 Rank Dependent Utility Theory

One alternative to EUT is to allow risk preferences to depend on the rank of the final outcome by using weighted probabilities as decision weights when evaluating lotteries. As before, we assume a Constant Relative Risk Aversion (CRRA) utility function. Under RDU the objective probabilities for each outcome k, p_k are replaced by subjective probabilities w_k. The expected utility is the probability weighted utility of each outcome in each lottery is

$$EU_i = \sum_{K=1,K} (w_k \cdot u_k)$$

where

$$w_i = \omega(p_i + ... + p_n) - \omega(p_{i+1} + ... + p_n)$$

for i =1,..., n-1, and $\omega(p_i)$ for i=n. The subscript indicates that outcomes are ranked from worst to best, and $\omega(p)$ is a probability weighting function. We assume the weighting function proposed by Tversky & Kahneman (1992). This function has weights

$$\omega(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$$

for

$$0 < p < 1$$

that gives the weighting function an "inverse S-shape". The weighting function is concave for small probabilities - implying overweighting of small probabilities - and convex for large probabilities - implying underweighting of large probabilities. The parameter γ informs about the curvature of the weighting function. In reviewing a substantial amount of evidence, Gonzalez & Wu (1999) show that $0 < \gamma < 1$: the smaller is γ, the more curved is the weighting function (the more probability weighting deviates from linear weighting), the more responsive to changes in extreme probabilities is the subject, the lower the crossover point. When γ is equal to 1, the model reduces to EUT.

The likelihood of the observed responses, assuming individuals choose according to RDU and displaying a CRRA, depends on the estimates of r given the above statistical specification and the observed choices. The log-likelihood is

$$lnL(r,\omega,\gamma;y,X) = \sum_i [(ln\Phi(\nabla RDU) \cdot I(y_i = 1)) + ln(1 - \Phi(\nabla RDU)) \cdot I(y_i = -1))]$$

where $I(\cdot)$ is the indicator function, $y_i = 1(-1)$ is the choice of the right (left) lottery in task i, and X is a vector of observable characteristics. As before, we include the error specification proposed by Fechner and popularized by Hey & Orme (1994). The latent index becomes

$$\text{Prob(choose lottery R)} = \Phi(\frac{\nabla RDU}{\mu})$$

The likelihood specification maximizes, then, $lnL(r,\omega,\gamma,\mu;y,X)$ to estimate r,ω,γ and μ given the observations on y and X.

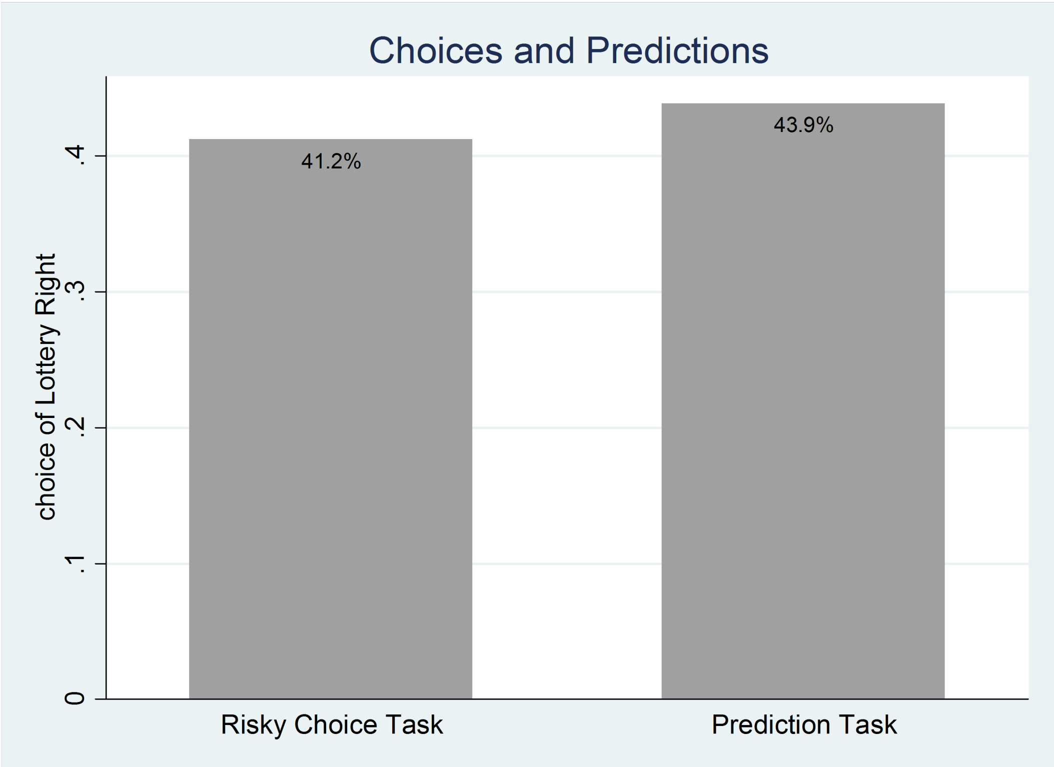
2.5 Results

2.5.1 Aggregate Results

Figure 2.1 depicts the distribution of individual choices and predictions.

Subjects chose lottery Right 41.2% (.017) times over the 40 decision tasks. Predictions of risk attitudes of others were measured by asking subjects to predict how another student (female or male) chose in the individual task. Over the 40 prediction tasks subjects predicted the target to choose lottery Right 43.9% (.016) times. A means test (t= -1.1469, p-value= 0.2528) accepts the null hypothesis of no difference between individual

Figure 2.1: Individual choices and Predictions



risky choices and predictions of Lottery Right. Descriptive analysis suggests no difference between aggregate observed choices and predictions.

When we look at gender difference in individual choices, we find that females (39.4%(.022)) and males (45.1%(.024)) chose, on average, the same percentages of Lottery Right (t= 1.5530, p-value= 0.1235). Similarly, we find that females (43.2%(.020)) and males (45.4% (.024)) made similar predictions about the target's choice of lottery Right (t= 0.6351, p-value= 0.5267).

Descriptive analysis suggests no gender difference between observed choices and predictions.

2.5.2 Structural coefficients

Table 2.1 displays the maximum likelihood estimates of the EUT model applied to participants' behavior. The estimates are a linear function of the CRRA coefficient *r* for the choices in the decision tasks. The dependent variable is the binary choice between lottery Left and Right, and a standard probit function links the difference in expected utility to the observed choices.

Results show that subjects are on average risk averse (the CRRA coefficient $r > 0$) and that gender has no effect on risk aversion. The estimate for the dummy picking out the risk aversion responses by females shows that the RRA is not significantly higher than that of males. Similarly, we find that the RRA does not change when making decisions in the risky choices task. On the other hand, we find that the estimate for the dummy picking out the measurement error μ associated to females is 0.614 higher than that associated to males.

Overall, when we assume subjects to choose according to EUT, we find that the curvature of utility function is the same for females and males, but that the measurement error is higher for females than for males.

Table 2.1: Statistical Model of Risk Aversion under EUT

VARIABLES	r	μ
Female	0.0815	0.614**
	(0.141)	(0.240)
Risk Choice Task	0.0726	0.062
	(0.073)	(0.147)
Constant	0.294**	1.318***
	(0.092)	(0.148)
Observations	8,400	8,400

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

Maximum-likelihood estimates of CRRA utility function under EUT.
N= 8400 binary choices by 105 subjects in the risky choice and prediction tasks
Estimates corrected for clustering on the individual.

When we apply the RDU model to individuals' behavior, we find evidence of probability weighting (Table 2.2). The estimate γ is 0.820, and thus we reject the hypothesis that subjects value probabilities linearly (i.e. that the estimate collapses to EUT, $\gamma = 1$). When we look at the estimate of γ for female subjects, we find that the weighting function for females is the same as that of males. Similarly, the γ associated to choices in the Risky Choice Task is not different from that associated to predictions.

When allowing for probability weighting, we do not find gender difference in risk attitudes. Although the estimate for the dummy picking out the risk aversion responses by females shows that the RRA is 0.3 higher than that of males, this is not significant. Similarly, the RRA does not change across the risky choices and the prediction tasks.

Similarly to the coefficient estimated under EUT, we find that the estimate for the dummy picking out the measurement error μ associated to females is higher than that associated to males.

We can conclude that rank-dependent transformation of the probability weights into decision weights better describes individuals' behavior in risky choices than when allowing for probability to be linear, since we find evidence of probability weighting. Similarly to the estimates assuming EUT, we find gender difference neither in the curvature of the utility function, nor in the probability weighting. On the other hand, we do find that females display a higher measurement error than males both under EUT and RDU models of choice.

These results are in contrast with the maximum likelihood estimates of Conte *et al.* (2016) and Fehr-Duda *et al.* (2006), but are similar to those in the high-payoff treatments of Holt & Laury (2002).

Table 2.2: Statistical Model of Risk Aversion under RDU

VARIABLES	r	γ	μ
Female	0.303	-0.403	0.425*
	(0.274)	(0.353)	(0.236)
Risk Choice Task	0.20	-0.286	-0.0614
	(0.172)	(0.206)	(0.179)
Constant	0.421**	0.820***	1.141***
	(0.1666)	(0.195)	(0.185)
Observations	8,400	8,400	8,400

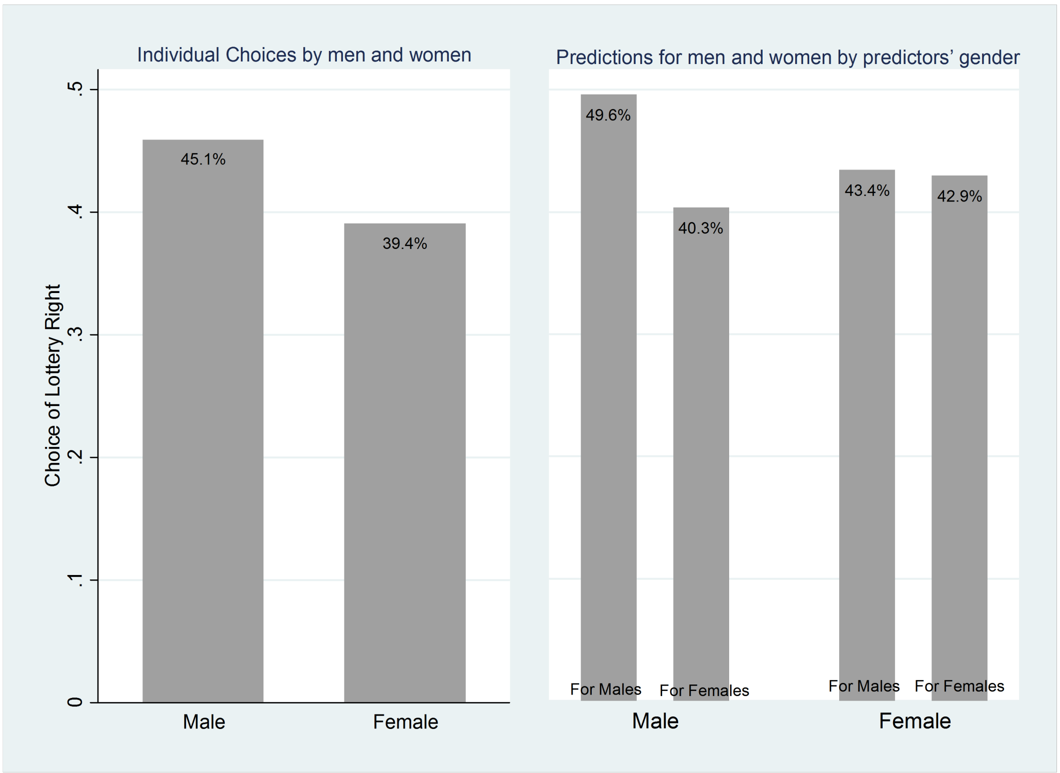
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Maximum-likelihood estimates of CRRA utility function under RDU.
N= 8400 binary choices by 105 subjects in the risky choice and prediction tasks
Estimates corrected for clustering on the individual.

2.5.3 Analysis of Prediction Bias

We start the analysis of stereotype bias in predictions by describing the relationship between predictor and target gender. Figure 2.2 reports the predictions for lottery Right by the gender of predictor and target, and the aggregate observed choices by males and females. We observe that females predict their target to choose the same amount of Lotteries Right for males (43.4%(.027) and females (42.9%(.031)) (t-test, t= 0.1173, p-value= 0.9069). On the other hand, we find that males predict females (40.3%(.037)) to choose significantly less Lotteries Right than males (49.6%(.031)) (t-test, t= 1.9258, p-value= 0.0633).

Figure 2.2: Individual choices and Predictions by predictors' gender



The latter result suggests that predictions by males are biased by gender stereotype. To corroborate this insight, we analyse the difference between predicted and actual choice by running an OLS model. Our index of *Prediction Bias* (PC - AC) is computed as the difference between *Predicted Choice* by the predictor and *Actual Choice*. This latter is computed as mean percentage of lottery Right chosen in each of the 40 tasks by the gender population to which the predictor's target belong.

When the index is negative ($PC < AC$), the predictor overpredicts the risk aversion of the gender population to which the target belong. On the other hand, when the index is positive ($PC > AC$), the predictor overpredicts the risk tolerance of the gender population to which the target belong.

Table 2.3 shows the results of the OLS model producing two sets of regression coefficients, one for female predictors and one for male predictors [12].

[12]We also estimated a Tobit model (see Appendix B) to account for the censoring of the index from -1 to

Model a tests whether information that the target is *Female* results in a systematic bias. We find that female predictors are not biased by gender information, as the coefficient is not significant. On the other hand, we find that male predictors use gender information inaccurately, as they predict female targets to be more risk averse than they actually are.

In Model b we also test for another fundamental factor that might bias perceptions of others' risk attitude. In particular, we test for *false consensus bias*, that is whether own behavior is projected in perceptions of others' behavior (Ross *et al.*, 1977). We, therefore, include the predictor's *OwnChoice* in the same task. If own risk preferences bias predictions of others' risk attitude, then more risk taking individuals will perceive the target as more risk taking than she actually is. We find positive significant coefficients for both female and male predictors. Therefore, unconditionally on gender, we find predictors' beliefs to be regressive on predictors' own choices.

Table 2.3: Prediction Bias

VARIABLES	$(Female_a)$ PC-AC	$(Female_b)$ PC-AC	$(Male_a)$ PC-AC	$(Male_b)$ PC-AC
$FemaleTarget$	0.00338	0.00318	-0.0843*	-0.0777*
	(0.0408)	(0.0364)	(0.0477)	(0.0434)
$PredictorChoice$		0.263***		0.289***
		(0.0364)		(0.0510)
$Constant$	0.0174	-0.0864***	0.0790**	-0.0541
	(0.0270)	(0.0248)	(0.0305)	(0.0340)
Observations	2,880	2,880	1,320	1,320
R-squared	0.000	0.078	0.009	0.112

Estimates corrected for clustering on the individual
*** p<0.01, ** p<0.05, * p<0.1
Robust standard errors in parentheses

1. The model produces similar results.

2.6 Conclusions

This study shows that individuals use gender as informative device to make predictions of others' risk attitude. In a controlled laboratory environment, we presented subjects with two incentive compatible risk aversion and prediction tasks in which subjects were only informed about the target's gender. While previous studies inferred risk attitude and predicted risk attitude of others from observed choices, ours is the first attempt to directly estimate the risk attitude coefficients underlying a latent model of choice that best describes observed behavior.

RDU model of choice explains subjects' behavior better than EUT, as we find evidence of probability weighting. However, we find gender difference neither in the structural estimates of the risk attitude coefficients underlying EUT nor in those underlying RDU.

We do find gender difference in measurement error: females state their choices with a higher error than males. Therefore, in our study, beliefs that females are more risk averse than males results to be a stereotype.

Our study shows that predictions of others' risk attitude are inaccurate and reflect gender stereotype. In particular, we find that males inaccurately overpredict females' risk aversion.

CHAPTER 3

HONESTY UNDER THREAT

with Francesca Gino - Harvard Business School
and Marco Piovesan - University of Copenhagen

Team members may contribute to enforce a culture of corruption within organizations and institutions when these are not optimally designed. In three experiments, we predict and find that individuals' perception of dishonesty in suboptimal group settings, like that of Weisel & Shalvi (2015), can be shaped by the surrounding context. We find that honesty in groups is negatively affected when punishment among peers and by an uninvolved party is allowed. On the other hand, we find that introducing a conflict of interest in suboptimal group settings slows down the diffusion of collaborative dishonesty.

Keywords: Group Dishonesty; Punishment; Conflict of Interest; Loss Aversion
JEL classification: C92; D03

3.1 Introduction

Dishonesty in the workplace is a widespread and multifaceted phenomenon: from embellishing the truth on work performance, to stealing office supplies; from promised bonuses and promotions that never arrive, to co-workers hiding valuable information. In the last decade, economists and psychologists have shown that dishonesty is malleable and dynamic (Monin & Jordan, 2009; Falk & Tirole, 2016). Contrary to what the traditional approach to criminal behavior (Becker, 1968) suggests, dishonest decisions are influenced by a much broader set of factors than just the costs and benefits associated with the act. Individuals care about their self-image and act dishonestly to increase their earnings up to the point in which they can mitigate the threat to the moral self (Mazar *et al.*, 2008; Bénabou & Tirole, 2011). To do so, they engage in self-serving justification (Shalvi et al., 2011), especially when norms are ambiguous (Shalvi *et al.*, 2015) and the dishonest action benefits others (Wiltermuth, 2011). Individuals refrain from full dishonesty also because they suffer from intrinsic costs of lying (Ellingsen & Johannesson, 2004; Gneezy, 2005; Abeler *et al.*, 2014) and guilt aversion when they view honesty as the norm to follow (Charness & Dufwenberg, 2006). However, the extent to which they perceive honesty as a value can be shaped by the context. People are less lying averse when financial incentives are negatively framed, because loss aversion changes the perception about dishonesty (Grolleau *et al.*, 2016). Similarly, institutions lead individuals to view dishonesty as normal, as may happen within banks (Cohn *et al.*, 2014) and societies with high corruption levels (Gächter & Schulz, 2016).

While most research has been conducted on individuals' decision to act dishonestly and its underlying mechanisms, less we know about "collaborative dishonesty", that is when the decision to act dishonestly involves more than one agent. This combined dishonest action of two or more individuals is at the roots of extremely costly phenomena for the whole society, such as corruption, bribery activities, financial and sport scandals. To mitigate the propagation of negative externalities, it is fundamental to understand whether the mechanisms behind individual dishonest behavior can also explain dishonesty in group settings, or whether alternative mechanisms are better predictors of cooperative dishonesty.

Sutter (2009) suggests that dishonest behavior in groups is driven by lower level of trust and higher strategic behavior than those of individuals. In particular, in the setting of a sender-receiver game, he finds that team senders engage in deception through sophisticated truth-telling more than individual senders [1].

Conrads *et al.* (2013) show that diffusion of responsibility increases dishonest behavior. Adding a team compensation scheme [2] to the die-rolling game (Fischbacher & Föllmi-Heusi, 2013), they find that dishonesty is higher under the team incentives than under the individual piece-rate scheme. Gino *et al.* (2013) further disentangle the effect of "altruism" (Wiltermuth, 2011) on dishonest behavior by using problem-solving tasks in which benefits are split with either another or two other students. They show that the presence of a beneficiary promotes more dishonest behavior and allows to preserve the moral self more than when dishonesty benefits the self only, suggesting that both social concerns and self-justification drive dishonest behavior.

Weisel & Shalvi (2015) study dishonest collaboration with a "dyadic die-rolling" paradigm in which individuals are forced by the institution to corrupt in order to increase their earning. By manipulating the incentive structure such that outcomes are either aligned, fixed, or higher for one of two players [3] , they show that the "moral" act of cooperating on equal terms (i.e. outcomes are aligned) outweighs the costs associated to lying, thus, promoting dishonesty in group settings. Recently, Kocher *et al.* (2016) have showed that allowing individuals in the group to communicate promotes dishonesty as in the payoff commonality framework of Weisel & Shalvi (2015). By adding the possibility to exchange information in a chat before making the decision in the die-rolling task (Fischbacher & Föllmi-Heusi, 2013), they show that communication shapes the perception of which norm to follow within group settings. Taken together, these findings suggest that different contextual features might increase dishonesty in group settings. In particular, when cooperation (equally) benefits other individuals and communication

[1] Senders are informed about the payoffs of two options: option A favors receivers, while option B favors senders. Then, senders choose the message to disclose to receivers: i) "option A will earn the receiver more money than option B", ii) "option B will earn the receiver more money than option A". Receivers choose which of the two options is implemented for payment. In this setting, the distribution of deception is defined in four types. "Benevolent truthtellers" send the message about the option that benefits receivers (A) and expect receivers to implement it (A). "Sophisticated truth-tellers" send the message about the option that benefits receivers (A) and expect receivers not to trust them, thus to implement the other option (B). "Benevolent liars" do not suffer from lying costs but care about receivers: they send the message about the option that benefits them more than receivers (B), but expect receivers not to trust them, thus to implement the other option (A). "Liars" send the message about the option that benefits them more than receivers (B) and expect the receivers to implement it (B).

[2]The sum of the numbers reported by two randomly matched individuals is split equally between the two individuals.

[3]*Aligned outcomes*: if they get a double players get the same payoff equal to the number reported. *A(B) high(low)*: A(B) gets a fixed high(low) payoff if they get a double independently on the number reported, while B(A) gets the value of the double. *A(B) fixed*: A(B) gets a fixed payoff regardless of whether they get a double or not.

is allowed, individuals change their perception of honesty as norm to follow and engage in higher dishonesty.

By shedding light on the effect of three contextual changes on perceptions of collaborative dishonesty, our study contributes to previous research showing that when individuals are given the incentive to dishonestly cooperate to equally share benefits they cheat more. We conduct three laboratory experiments at the Laboratory of Experimental Economics of the University of Copenhagen using the "dyadic die-rolling paradigm" of Weisel & Shalvi (2015). In the first experiment, we study the effect that second and third-party punishment has on preventing or promoting dishonesty in groups. On the one hand, the threat of punishment may limit the willingness to violate moral rules: each agent may act honestly out of the fear that the other player can punish her if she believes that she was dishonest. On the other hand, the threat of punishment can induce honest people to act dishonestly: punishment can represent a tool to create a social norm of dishonesty and/or make "rationalization" easier for those people that are looking for an excuse for their dishonest intentions. In the second experiment, we study the effect of introducing a conflict of interest in collaborating dishonestly. Introducing unequal rewards might make it more difficult to coordinate on a tacit agreement, since players might be concerned about different motives, such as inequality aversion, than profit-maximization only. Finally, the third experiment extends the findings from Grolleau *et al.* (2016) and Schindler & Pfattheicher (2016) on the detrimental effect that loss aversion might have on perceptions of dishonesty in group settings.

We find that allowing for punishment mechanisms in suboptimal group settings has a detrimental effect on normative perceptions of corruption. We also find that adding a conflict of interest in group settings that give the incentive to corrupt slows down the diffusion of collaborative dishonesty. Finally, we show that the frame of group settings is not crucial for promoting collaborative dishonesty when these are not designed optimally.

3.2 Method

A total of 268 students from the University of Copenhagen took part in the study (female: 52.7%, sd. 0.50; age: 26.02, sd. 0.21). 50 were assigned to the Baseline condition, 46 to the Second Party Punishment Condition, 78 to the Third Party Punishment, 46 to the Conflict of Interest condition, and 48 to the Loss Frame condition. Participants were recruited through ORSEE (Greiner, 2004) to the Laboratory of Experimental Economics of the same University in even groups for the conditions Baseline, Second Party Punishment, Conflict of Interest and Loss Frame. For the condition Third Party Punishment they were recruited in groups of three.

The experiment was programmed and conducted using z-Tree software (Fischbacher, 2007). Upon their arrival to the laboratory, participants were randomly allocated to cubicles and asked to privately read the instructions. Before starting the experiment, they were asked to complete control questions verifying their understanding of the experiment. Participants received on average DKK 112 in addition to a show-up fee of DKK 50. Each session lasted on average 1 hour and 30 minutes.

In the first stage of all experiments, participants are presented with the experiment-specific variants of the "dyadic die-rolling paradigm" from Weisel & Shalvi (2015). Participants are randomly paired and assigned to roles of Player A or B. First, Player A rolls the die and reports the outcome. After that, Player B is informed about player A's report, she rolls the die and reports the outcome. This task is repeated for 20 periods, keeping the role and the pair fixed for all the periods. If Player A and B report the same number (i.e., a "double") each participant earns an amount of points equal to the number reported multiplied by twenty (for instance 120 points for a "double six"); zero points otherwise. Die rolls are truly private and misreporting is possible. The expected probability of rolling a double in a single trial is 1/6 (16.7%). Any significant statistical deviation can be interpreted as an evidence of cheating. Each player is informed about her report, the counterpart's report and the payoff associated from the previous periods. The experimenter pays participants for one of the 20 periods, randomly selected at the end of the experiment. Each point corresponds to 1 DKK. In our study, we modify Weisel & Shalvi (2015)'s paradigm by giving participants a fixed endowment of 46 points at the beginning of each period independently on the outcome of the die-rolling task.

After completing the dyadic die-rolling task, participants are presented with the Bomb Risk Elicitation Task (Crosetto & Filippin, 2013) and the Personality Traits questionnaire (Ashton & Lee, 2009). The BRET is a choice-based elicitation method that asks subjects to decide the point at which they prefer to stop collecting 100 available boxes, one of which contains a bomb. Subjects receive an earning in DKK which is equal to the number of boxes collected, however if one of those contains the bomb they receive nothing. The cumulative probability of collecting the bomb increases linearly from 0.01 to 1. This setting allows to derive a measure for risk preferences from the position in which the subject stopped, i.e. the less boxes participants collect, the more risk averse they are. The Personality Traits questionnaire provides us with six measures of personality traits: Honesty-Humility, Extraversion, Agreeableness, Conscientiousness and Openness to Experience. Participants are asked to state how much they agree on a scale from 1 to 5 to 60 non-incentivized statements. Finally, they are asked to complete a demographic survey.

3.3 Experiment 1

In this experiment, we study the effect of providing a punishment mechanism in settings that give individuals no other option to increase their benefits than to corrupt. This experiment contends that second party and third party punishment mechanisms exacerbate the negative side of contexts that embed normative conflicts.

Previous studies on settings characterized by normative conflicts (Nikiforakis *et al.*, 2012; Xiao & Kunreuther, 2016) show that peer punishment is less effective in enforcing cooperation in public good settings. Ours examines cooperation in a setting that requires to violate honesty norm and that, in turn, embeds two plausible rules about how one should behave: cooperate against honesty or acting honestly while hurting the group.

There are two arguments that provide us with two alternative predictions. The first relates to intrinsic preference for honesty (Abeler *et al.*, 2014): some may believe that the group should increase welfare only through honest procedures, because they have a strong preference for honesty. Thus, we predict that individuals with preference for honesty sanction to enforce the norm of honesty. If they are the majority, there will be a negative effect on collaborative dishonesty.

The second relates to self-serving justification mechanisms and social concerns (Shalvi *et al.*, 2011; Gino *et al.*, 2013), and suboptimal institutions (Gächter & Schulz, 2016): suboptimal settings that provide individuals with the prospect of benefiting others against ethical values, on the one hand, allow to preserve the moral self and, on the other, shape over time individual ethical values. Thus, we predict individuals to punish to enforce dark cooperation. This will have a positive effect on collaborative dishonesty.

If only involved parties were given the opportunity to punish, we would not be able to isolate different forces that shape norm enforcement, as sanctions can be motivated by non-normative motives such as retaliation. On the other hand, giving the option to sanction to a third party that is not affected by the norm violation "reveals the truly normative standards of behavior" (Fehr & Fischbacher (2004): p.65). Therefore, third-party punishment enables us to shed light on the normative conflict characterizing our setting.

As before, there are two arguments that provide us with two predictions. Third-parties might sanction when players violate honesty, because they have internalized a norm of honesty (Abeler *et al.*, 2014). Thus, we predict third-parties with internalized norms of honesty to sanction violations of honesty. At the same time, institutions may shape individual's honesty (Gächter & Schulz, 2016). Dishonest cooperation might appear justifiable also to the eye of an uninvolved party if the context does not give other option to increase benefits than through unethical procedures. We predict suboptimal context to shape third parties' perceptions of ethical values and to trigger sanctions of violations of dishonest cooperation.

To address these predictions we implement a between-subject design with two treatments (*Second Party Punishment* and *Third Party Punishment*). In our Baseline Condition [4] (Baseline, 2 sessions, 50 subjects) participants are endowed with a fixed amount of 46 points. In our Condition (Second Party Punishment, 2 sessions, 46 subjects) each player -after having observed the counterpart's report- can "spend" 6 of these 46 points to "punish" the other player [5].

In our Third Party Condition (Third Party Punishment, 3 sessions, 78 subjects) we add a third player with a punishment option to our Baseline condition. Only Player C - after having observed what Player A and Player B report - can "spend" 6 points of her endowment to "punish" the other players. Player C does not participate in the "dyadic die-rolling paradigm" and can earn only her fixed endowment. Therefore, we give Player C an additional endowment of 12 points to avoid punishment actions motivated by perceived unfairness. The additional endowment is equal to the expected value of winning an amount of points for Player A and Player B from participating in the "dyadic die-rolling paradigm" [6].

3.3.1 Results

3.3.1.1 Collaborative and Individual Dishonesty

Figure 3.1 provides a joint representation of individual and collaborative dishonesty. In our setting, participants earn money only if they report a "double". The number of doubles are represented by the dots in the diagonal, that is the joint report of Player A and Player B in a single period for the two treatments (Panel a, *Second Party Punishment*, 23 pairs x 20 periods = 460 dots; Panel b, *Third Party Punishment*, 26 pairs x 20 periods = 520 dots; Panel c, *Baseline*, 25 pairs x 20 periods = 500 dots). Figure 3.1 shows that

[4] The Baseline Condition is the same in all the three experiments.

[5] Every point allocated for punishment decreases the payoff of the other player of 2 points.

[6] $[(1/6 \cdot 20 + 1/6 \cdot 40 + 1/60 \cdot 60 + 1/6 \cdot 80 + 1/6 \cdot 100 + 1/6 \cdot 120)/6] = 11.67 \sim 12$ points

Players dishonestly collaborated. In fact, the average percentages of reported doubles in *Baseline* (Wilcoxon signed-rank test: z= 4.361, p-value=0.0000), *Second Party Punishment* (Wilcoxon signed-rank test: z= 3.991, p-value=0.0001) and the *Third Party Punishment* (Wilcoxon signed-rank test: z= 4.311, p-value= 0.0000) are significantly higher than the average percentage assuming honesty (16.7%).

However, the average percentage of doubles in *Second Party Punishment* (69,6%, sd:0.323; Wilcoxon rank-sum test: z=-1.000, p-value=0.3174) and in *Third Party Punishment* (64,6%, sd:0.316; Wilcoxon rank-sum test: z= -0.409, p-value=0.6829) are not different from that in *Baseline* (56,6%, sd: 0.343).

The side histograms in Figure 3.1 show the distribution of numbers reported for each player. The not-uniform distributions suggest that both players reported the number dishonestly. In fact, both players A and B report a number significantly higher than the expected 3.5 assuming honesty (Wilcoxon signed-rank test: A in *Baseline*: z= 4.349, p-value 0.000; A in *Second Party Punishment*: z = 4.110, p-value 0.000; A in *Third Party Punishment*: z=3.876, p-value 0.0001; B in *Baseline*: z= 4.120, p-value 0.000; B in *Second Party Punishment*: z= 4.046, p-value 0.0001; B in *Third Party Punishment*: z= 3.956, p-value=0.0001).

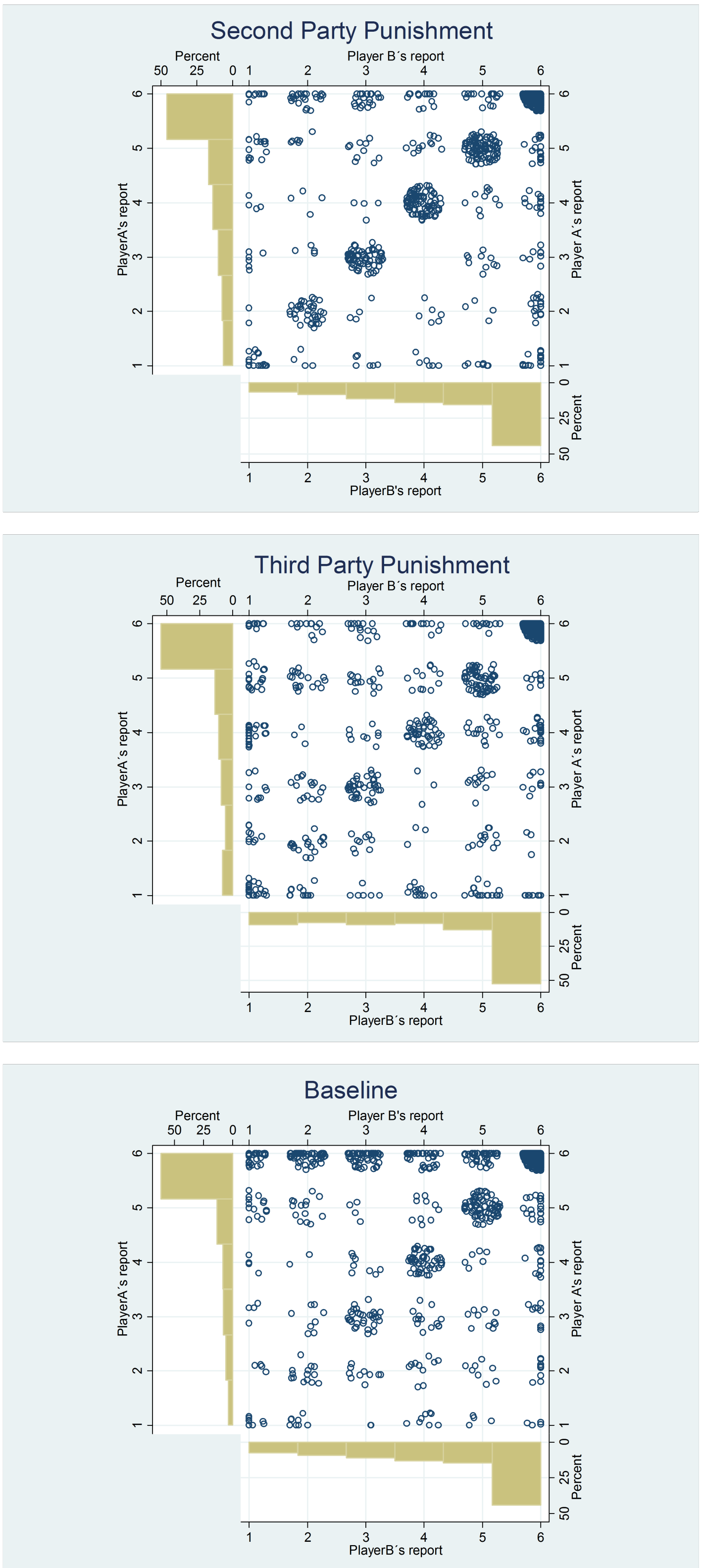
Player A in *Second Party Punishment* (4.65: sd: 0.90, Wilcoxon rank-sum test: z= 1.394, p-value= 0.1633) and in *Third Party Punishment* (4.86: sd: 1.12, Wilcoxon rank-sum test: z= 0.574, p-value= 0.5658) report a number that does not differ from that reported by Player A in *Baseline* (5.064: sd: 0.932).

Similarly, Player B in *Second Party Punishment* (4.56: sd: 0.84, Wilcoxon rank-sum test: z= -0.620, p-value= 0.5352) and in *Third Party Punishment* (4.65: sd: 1.07, Wilcoxon rank-sum test: z= -0.539, p-value= 0.5897) report a number that does not differ from that reported by Player B in *Baseline* (4.49: sd: 0.98).

3.3.1.2 Analysis of Types

Fig 3.2 and Fig 3.3 provide a joint representation of all decisions made by each group in *Baseline*, *Second Party Punishment* and *Third Party Punishment*. They show that individual behaviour is heterogeneous across groups. In particular, three main types of behaviour can be observed: i) brazen (individuals who reported 6 across all the 20 periods); ii) dishonest, i.e. individuals who reported on average a number larger than the honest benchmark but less than the maximum (6) across the 20 periods; iii) honest, i.e. individuals who reported on average truthfully (3.5) across the 20 periods.

Figure 3.1: Second Party and Third Punishment - Baseline



The proportion of Brazen Players across conditions is homogeneous. In particular, the proportion of Brazen Players A in *Second Party Punishment* (17%(0.39)) and *Third Party Punishment* (31%(0.47)) are the same as the proportion of Brazen Player A in *Baseline* (32%(0.48), Wilcoxon Rank Sum test 2P: z = 1.155, p-value = 0.2479; 3P: z = 0.094, p-value 0.9253). Similarly, the proportion of Brazen Players B in *Second Party Punishment* (9%(0.29)) and *Third Party Punishment* (23%(0.43)) is not different from that in Baseline (16%(0.37), Wilcoxon Rank Sum test 2P: test z = 0.756, p-value = 0.4494; 3P: z = -0.630 p-value = 0.5286).

As for Brazen Players, the proportion of Dishonest Players is homogeneous across conditions. The proportion of dishonest Players A in *Second Party Punishment* (74%(0.45)) and *Third Party Punishment* (46%(0.51)) is not different from that in Baseline (64%(0.49), Wilcoxon Rank Sum test 2P: z = -0.732, p-value= 0.4639; 3P: z = 1.268, p-value= 0.2049). Similarly, the proportion of Dishonest Player B in *Second Party Punishment* (82%(0.39)) and *Third Party Punishment* (65%(0.49)) is not different from that in *Baseline* (72%(0.46), Wilcoxon Rank Sum test 2P: test z = -0.864, p-value = 0.3873; 3P: z = 0.504, p-value = 0.6143).

Interestingly, while the proportion of Honest Player A in *Second Party Punishment* (9%(0.29)) does not differ from that in *Baseline* (4%(0.2), Wilcoxon Rank Sum test 2P: z = -0.664, p-value = 0.5065), we find that the proportion of Honest Player A in *Third Party Punishment* (23%(0.43), Wilcoxon Rank Sum test 3P: z = -1.960, p-value = 0.0500) is significantly higher than that in *Baseline*. This suggests that introducing a monitoring mechanism, such as a third party, elicits more honest types than in *Baseline*: Player A fears to be punished by the third party if she does not report honestly. On the contrary, the proportion of Honest Player B in *Second Party Punishment* (9%(0.29)) and *Third Party Punishment* (12%(0.33)) does not differ from that in *Baseline* (12%(0.33), Wilcoxon Rank Sum test 2P: z = 0.370 p-value= 0.7110; 3P: z = 0.051 p-value = 0.9596).

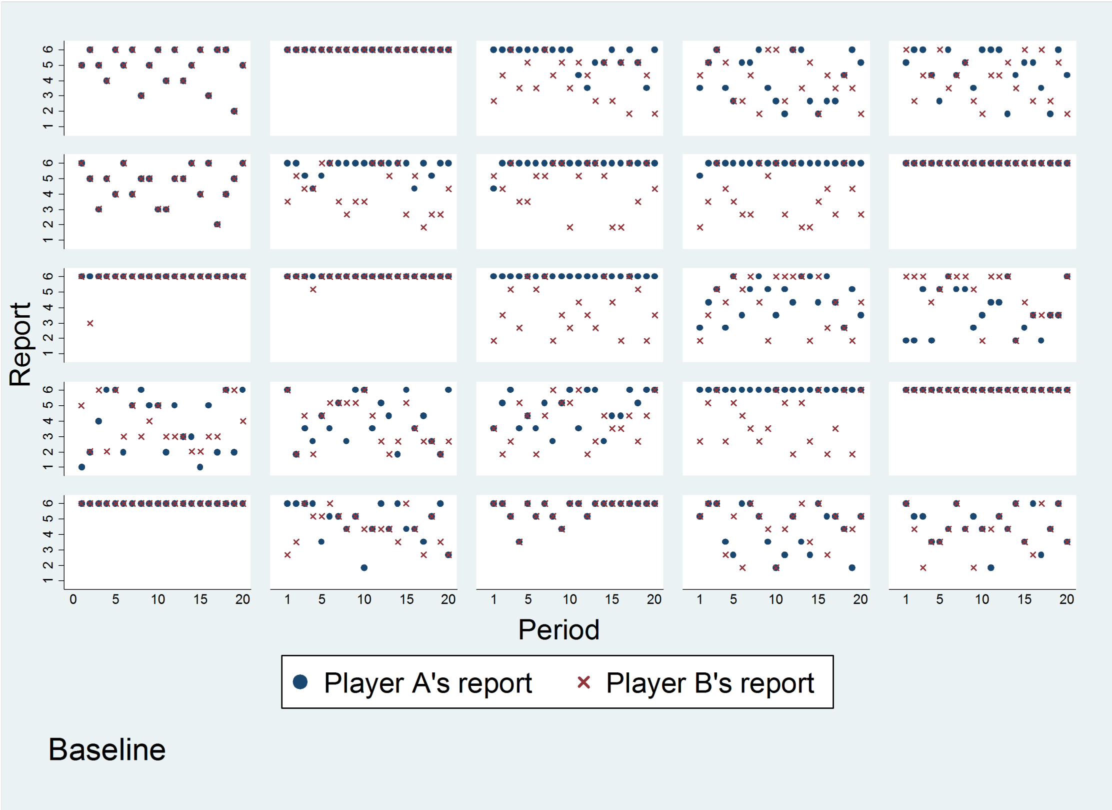
3.3.1.3 Evolution of Collaborative Dishonesty

Figure 3.4 shows the effect of *Second Party Punishment* and *Third Party Punishment* on collaborative dishonesty over the 20 periods of the game. It shows that in *Second Party Punishment* and in *Third Party Punishment* collaborative dishonesty is higher than in Baseline after the second half of the game.

We complement this insight with a random individual effect[7] probit regression

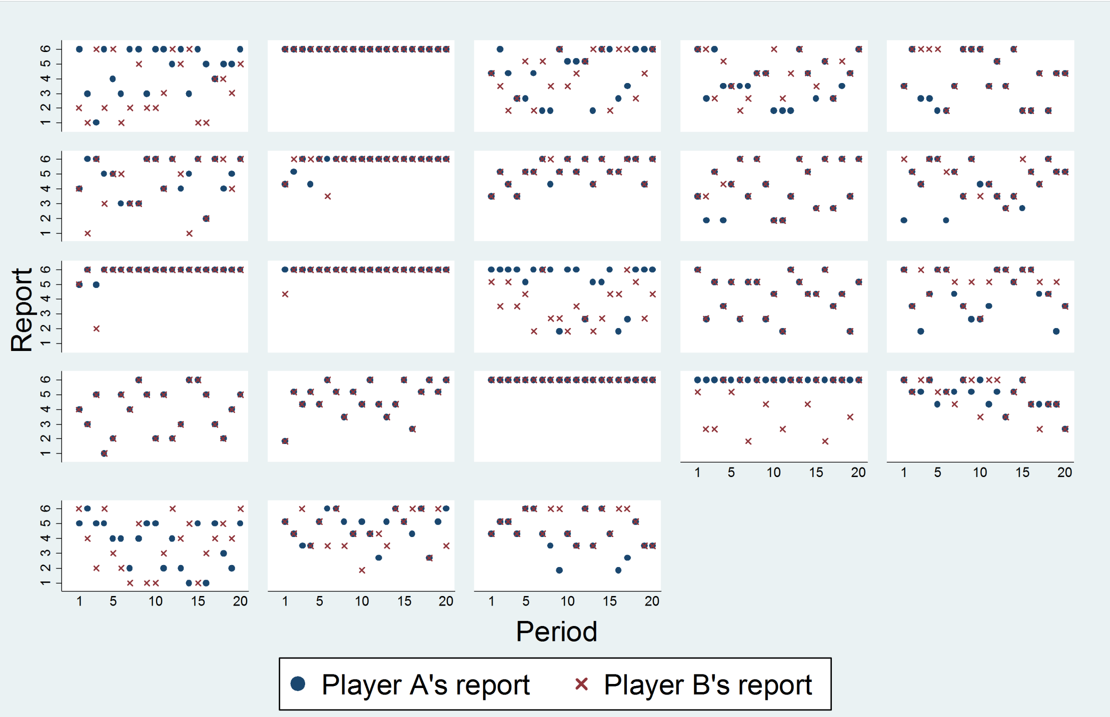
[7]Throughout the study we implemented random effects models to capture between-subject differences and between-subject heterogeneity.

Figure 3.2: Decisions by Groups Across Periods: Baseline

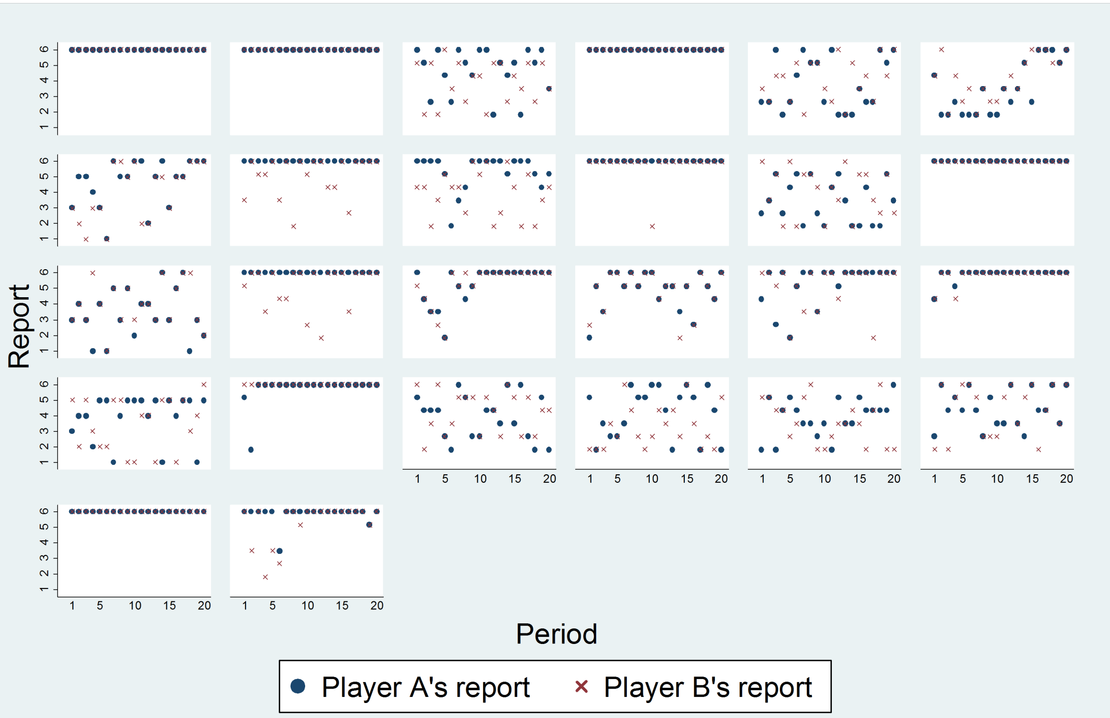


analysis of Player B's decision to report a double in *Second Party Punishment* and *Third Party Punishment* to isolate the time effect (Table 3.1). We include a dummy for a block of ten periods (*11-20*), using the first 10 periods as baseline. We add the dummies for the treatments *Second Party Punishment* (*SP*) and *Third Party Punishment* (*TP*). Finally, we add the interactions between *Second Party Punishment* and *11-20*, and *Third Party Punishment* and *11-20* to capture the combined effect. Results show that collaborative dishonesty is likely to increase over time compared to Baseline, as the coefficients associated to 11-20 in *Second Party Punishment* and *Third Party Punishment* are significantly positive. Wald test for the null hypothesis that the difference between the coefficients *Second Party Punishment x 11-20* and *Third Party Punishment x 11-20* is zero indicates that the coefficients have the same effect on collaborative dishonesty (χ^2_1 = 0.18, $p > \chi^2_1$ =0.6727).

Figure 3.3: Decisions by Groups Across Periods: Second Party Punishment and Third Party Punishment

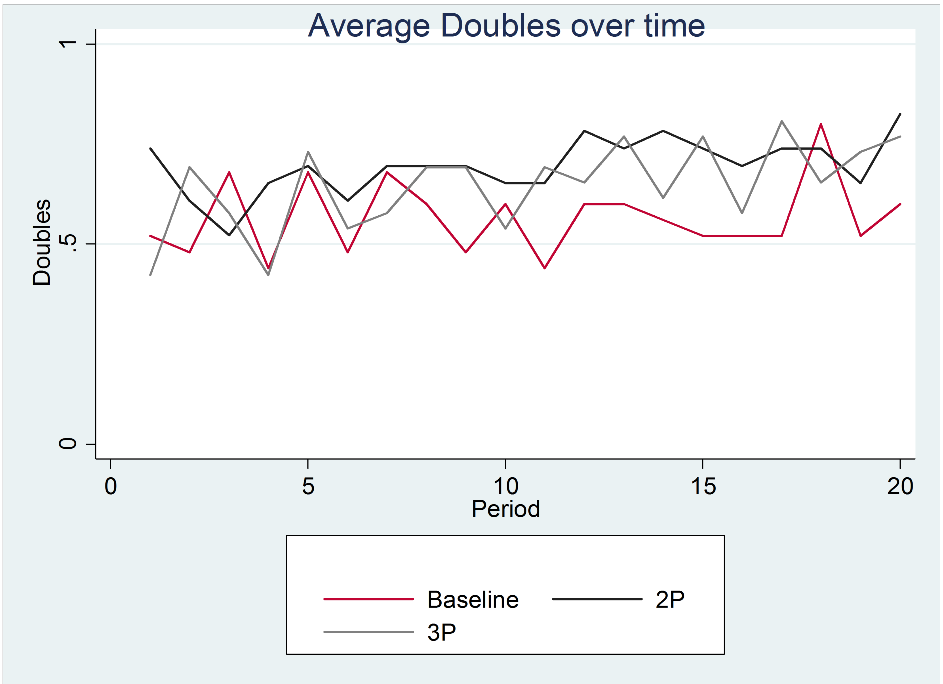


Second Party Punishment



Third Party Punishment

Figure 3.4: Punishment: Doubles over period



3.3.1.4 Analysis of punishment

In *Second Party Punishment* only 9,46% (sd:0.25) of players decided to punish the other player and when they used punishment they used it moderately at the same extent: on average Player A received 3.15 points and B received 2.48 points (Wilcoxon signed-rank test: z -0.179, p-value= 0.8578).

In *Second Party Punishment* players take part in the die-rolling paradigm, and their punishment decisions might be predicted by their preference for dishonesty. As a result, we conduct two separate random individual effect ordered probit regression analyses with the decision to punish a positive amount or not as a dependent variable by Player A and Player B in *Second Party Punishment*. As independent variables, we include the player's decision and the counterpart's decision to report brazenly or not. We define the variable Brazen A as a dummy variable taking the value 1 if Player A reports 6, zero otherwise. The variable Brazen B is a dummy variable taking the value 1 if Player B reports a double, zero otherwise. Table 3.2 shows that the more Player B is brazen, i.e. reports a double, the less she punishes the counterpart. This suggests that the more Player B gest along with Player A by accepting her action, the less likely is her willingness to punish her counterpart. Hence, reporting a double is a signal of

Table 3.1: Random individual effect probit regression analysis of Double decisions: Second- Party and Third-Party Punishment

VARIABLES	(1) Double	(2) Double
SP	0.299	0.113
	(0.411)	(0.425)
TP	0.149	-0.0781
	(0.396)	(0.409)
11-20	0.307***	0.0202
	(0.0840)	(0.141)
11-20 x SP		0.396*
		(0.213)
11-20 x TP		0.486**
		(0.199)
Constant	0.369	0.509*
	(0.289)	(0.296)
Observations	1,480	1,480
Number of id	74	74

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

collaboration and dilutes peer-punishment.

In *Third Party Punishment*, the percentage of uninvolved parties (23,852% sd: 0.24) that decided to punish the other players is higher than players who punished in *Second Party Punishment* (Wilcoxon rank-sum test: z= -4.413, p-value= 0.0000). However, as for *Second Party Punishment*, third parties sanctioned Player A and B moderately and at the same extent (A: 2.20 points; B 2.01 points; Wilcoxon signed-rank test: z 0.402, p-value: 0.6876).

To address the normative motivations behind punishment decisions in each treatment, we separate the case when players systematically sanctioned Player B for violating honesty norm - hence when she reported a double - and for acting against dark cooperation - hence when she did not report a double.

Within *Second Party Punishment*, we observe that the percentage of Player A that punished Player B for violating honesty ($3.9\% sd: 0.13$) is significantly higher than the percentage of Player A that punished for violating dark cooperation ($1.7\% sd: 0.08$) (Wilcoxon signed-rank test: z= 2.232, p-value: 0.0256). Similarly, within *Third Party Punishment*, we observe that the percentage of third parties that punished Player B for violating honesty ($15.8\% sd: 0.19$) is significantly higher than the percentage of third parties that punished for violating dark cooperation ($8\% sd: 0.13$) (Wilcoxon signed-rank test: z= 1.972, p-value: 0.0487).

This suggests that those who decided to consistently punish Player B were motivated by normative concerns in favour of honesty. However, the actual use of punishment for consistently violating a norm was too scarce to enforce honesty in this suboptimal group setting.

In Second Party Punishment, the almost null use of punishment for norm violations suggests that individuals perceive a normative conflict, but the fear of retaliation provides them with a tool to accept dishonest cooperation as a norm to earn some benefits. In Third Party Punishment, the scarce use of punishment for norm violations suggests that also uninvolved parties are uncertain about which is the right rule to enforce. As a result, third party becomes a non-credible monitoring mechanism to enforce honesty. Building on the insight from the analysis of types that the presence of a third party elicits more honest Players A than in Baseline, we can conclude that if third parties were more credible, they would be able to crowd-in more honest types.

Overall, these results suggest that this suboptimal group setting shapes the normative perception of dark cooperation at cost of honesty.

Table 3.2: Random individual effect probit regression analysis of Punishment decisions

VARIABLES	Player A Pun Dec	Player B Pun Dec
Brazen A	0.177	-0.0417
	(0.483)	(0.307)
Brazen B	0.970	-0.693*
	(0.887)	(0.372)
Constant	-5.834***	-3.215***
	(1.054)	(0.566)
Observations	460	460
Number of id	23	23

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.3.2 Discussion

Results of Experiment 1 show that introducing the norm enforcement mechanism of second-party and third-party punishment exacerbates the negative effects of suboptimal group settings (i.e. collaborative dishonesty), that is settings that give individuals no other option to increase their benefits than to corrupt. This is in line with recent evidence (Meier *et al.*, 2016) showing that norm enforcement mechanisms augment the negative consequences of suboptimal institutions, such as those characterized by a culture of Mafia. Our results are also in line with evidence that suboptimal institutions shapes individual ethical values (Gächter & Schulz, 2016), since both involved parties and third-parties barely use punishment to enforce a rule.

The same tool, punishment, that can be used to foster group cooperation to contribute to public goods (Fehr & Fischbacher, 2004), can produce more cooperation against ethical values.

3.4 Experiment 2

So far, we reported that suboptimal group settings promote collaborative dishonesty when individuals are forced to coordinate to equally share benefits. One explanation behind the diffusion of collaborative dishonesty is the ability of groups to make better decisions than individuals. Groups are more successful in strategic interactions (Kocher & Sutter, 2005; Sutter, 2009) and learn faster how to coordinate on efficient actions, because they are more concerned about profit-maximization than individuals (Feri *et al.*, 2009).

In our Baseline, efficiency is increasing as the group lies down the most profitable outcome. As a result, groups converge fast to reporting the maximum value, i.e. the most efficient action [8], because it guarantees the maximum payoff to all group members.

In Experiment 2 (Conflict of Interest, 2 sessions, 46 subjects), we add a mechanism that might deter diffusion of collaborative dishonesty enhanced by those profit-maximization and efficiency concerns which are proper of groups. In particular, we introduce an asymmetric payoff scheme to create a conflict of interest in reporting a double (Table 3.3).

As for the Baseline, players have the incentive to report a double to earn an additional amount of money to the initial endowment. However, in this setting, any combined action is an efficient equilibrium and yields different levels of inequality between one's and her counterpart's payoff: Player A has an incentive to report small values, while Player B is better off if Player A reports high values.

[8] Efficiency is here defined as the sum of payoffs (Engelmann & Strobel, 2004)

Table 3.3: Payoff scheme in Conflict of Interest

Player A's outcome	Player B's outcome	Player A's Earning	Player B's Earning
1	1	120	20
2	2	100	40
3	3	80	60
4	4	60	80
5	5	40	100
6	6	20	120

If individuals in the group are purely concerned about maximizing the group payoff, they will coordinate on the combined action yielding the highest payoff to Player A, who is in the position of first mover, and the lowest but still positive payoff to Player B. However, individuals might make dishonest decisions not only to maximize profit, but also because they are motivated by inequality aversion (Abeler *et al.*, 2016). Based on this insight, we predict that if individuals are inequality averse, they will converge to the combined dishonest action that minimizes the difference between their own and the counterpart's payoff. We also predict that if individuals are inequality averse, they will converge slower to collaborative dishonesty, because they are more concerned about how their stakes compare to their counterpart's ones than about profit maximization. Thus, conflict of interest might dilute the diffusion of collaborative dishonesty, by awakening social preferences at the expenses of the profit-maximization mindset enhanced by group settings.

3.4.1 Results

3.4.1.1 Collaborative and Individual Dishonesty

Figure 3.5 provides a joint representation of individual and collaborative dishonesty. As before, we look at the average percentages of doubles to investigate collaborative dishonesty, since participants earn money only if they report a "double". The number of doubles are represented by the dots in the diagonal, that is the joint report of Player A and Player B in a single period in *Conflict of Interest* (Panel a, *Conflict of Interest*, 23 pairs x 20 periods = 460 dots)), in the simulation assuming honesty (Panel b, *Decisions assuming honesty*, 23 pairs x 20 periods = 460 dots) and in the Baseline (Panel c). Differently from *Decisions assuming honesty*, the number of doubles (the number of dots in the diagonal) in *Conflict of Interest* are not random, but concentrated around the values 3 and 4. This can be better observed in Figure 3.6 in which doubles are concentrated around the value 3.5.

Figure 3.5 shows that Players dishonestly collaborated. In fact, the average percentages of reported doubles in *Conflict of Interest* (50.2% sd:0.26) is significantly higher than the average percentage assuming honesty (16.7%)(Wilcoxon signed-rank test: z= 4.110, p-value=0.0000). However, it is not significantly lower than that in *Baseline* (56.6% sd: 0.343)(Wilcoxon rank-sum test: z= 0.560, p-value= 0.5753).

The side histograms in Figure 3.5 show the distribution of numbers reported for each player in *Conflict of Interest* (Panel a), in the simulation of *Decision assuming honesty* (Panel b), and in *Baseline* (Panel c). The distribution of numbers reported for each player in *Conflict of Interest* is the same as if they reported honestly. In fact, while in *Baseline* the honesty benchmark against which compares Players' reported values is 3.5 ([1 + 2 + 3 + 4 + 5 + 6]/6), in *Conflict of Interest* 3.5 represents an indicator of collaborative dishonesty.

On average, we find that Player A (3.08, sd:0.38) in *Conflict of Interest* reports a number significantly smaller than 3.5 (Wilcoxon signed-rank test: A : z= -3.791, p-value 0.0002). On the other hand, Player B (3.56, sd:0.32) reports a value equal to the collaborative dishonesty indicator 3.5 (B in *Conflict of Interest*: z= 1.127, p-value 0.2596). Compared with the *Baseline*, both Player A and Player B report a number significantly smaller than Player A and Player B in *Baseline* (A, Wilcoxon rank-sum test: z= 5.835, p-value=0.0000; and B, Wilcoxon rank-sum test: z= 3.711, p-value=0.0002).

3.4.1.2 Analysis of Types

Results presented so far suggest that individuals in group settings are not always profit-maximizers, but also inequity averse when presented with unequal rewards: in *Conflict of Interest* individuals in the group converge to the combined action that minimizes difference between players' payoff.

To corroborate this insight, we compare the percentage of profit-maximizers in *Baseline* and *Conflict of Interest*. Fig 3.7 provides a joint representation of all decisions made by each group in *Baseline* and *Conflict of Interest*. It shows that that the percentage of Player A (20%(0.09)) and Player B (16.9%(0.08)) in *Conflict of Interest* who reported the value that maximizes their payoff, is significantly lower than the percentage of Player A (61%(0.34)) and Player B (44%(0.34)) in *Baseline* who reported 6 (Wilcoxon signed-rank test: A z= 4.630, p-value 0.0000; B z = 3.372, p-value=0.0007).

We, then, conduct a within-treatment analysis of the percentages of profit-maximizer and inequality averse players, these latter defined as individuals who get disutility when

Figure 3.5: Conflict of Interest and Decision Assuming Honesty - Baseline

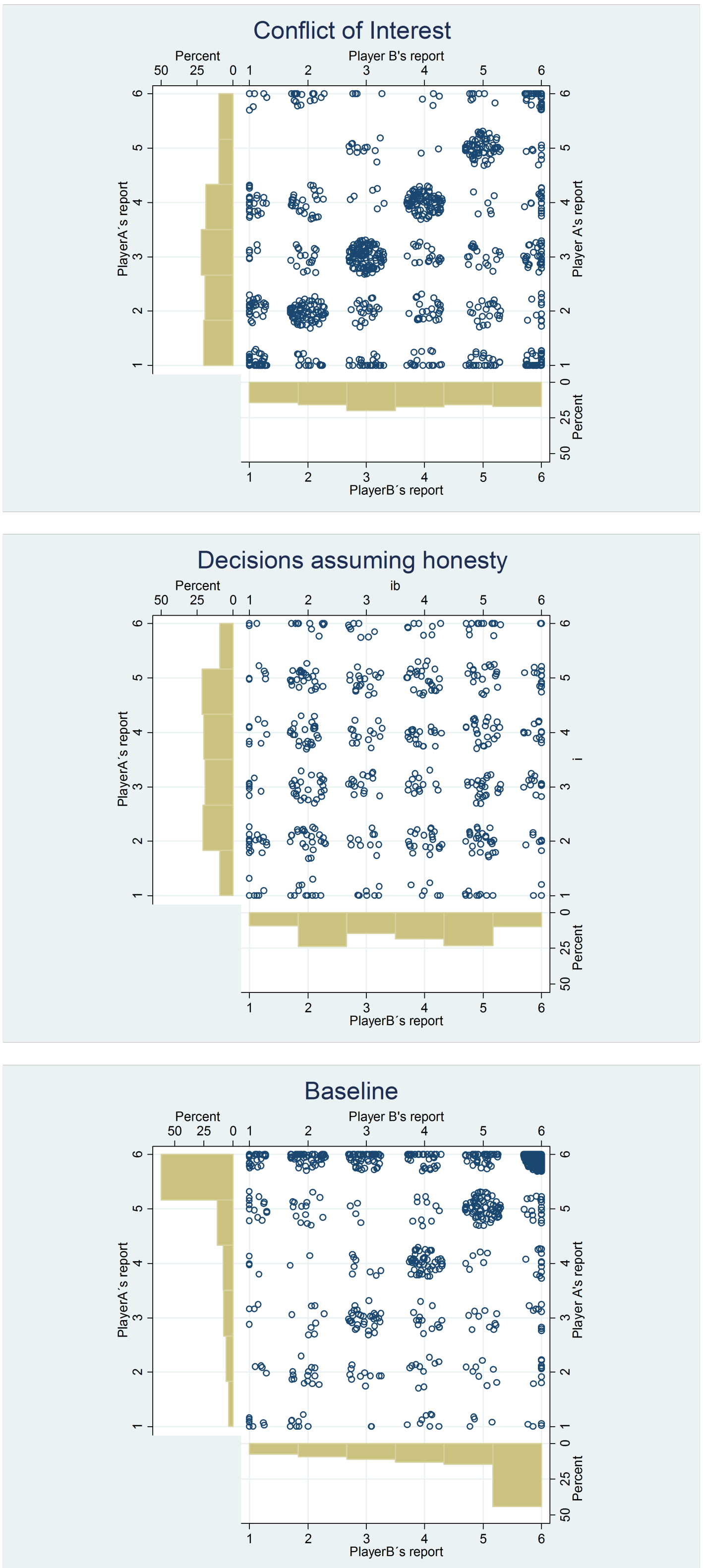
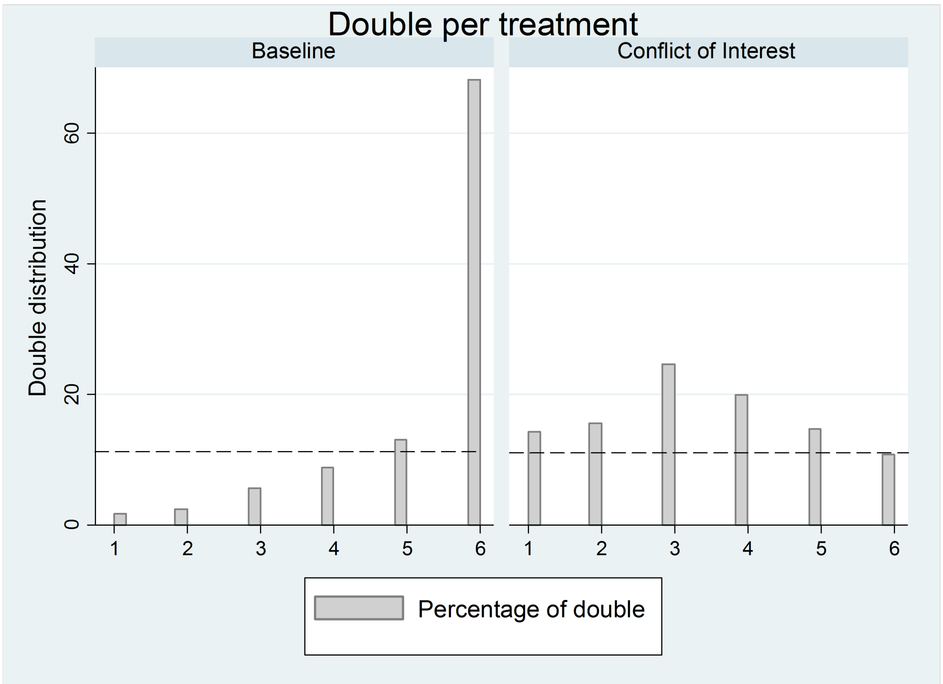


Figure 3.6: Distribution of Doubles in Baseline and Conflict of Interest



their payoff deviates from the average payoff (Bolton & Ockenfels, 2000). Building on this definition, our measure of inequality aversion is the percentage of Players who reported either 3 or 4.

Fig 3.7 (panel B) shows that both the percentages of inequality averse Players A (41%(0.147)) and Players B (37.9%(0.08)) are significantly higher than the percentage of profit-maximizer Players A and Players B (Wilcoxon signed-rank test; A: z = 18.571, p-value = 0.0000; B: z = 18.579, p-value = 0.0000).

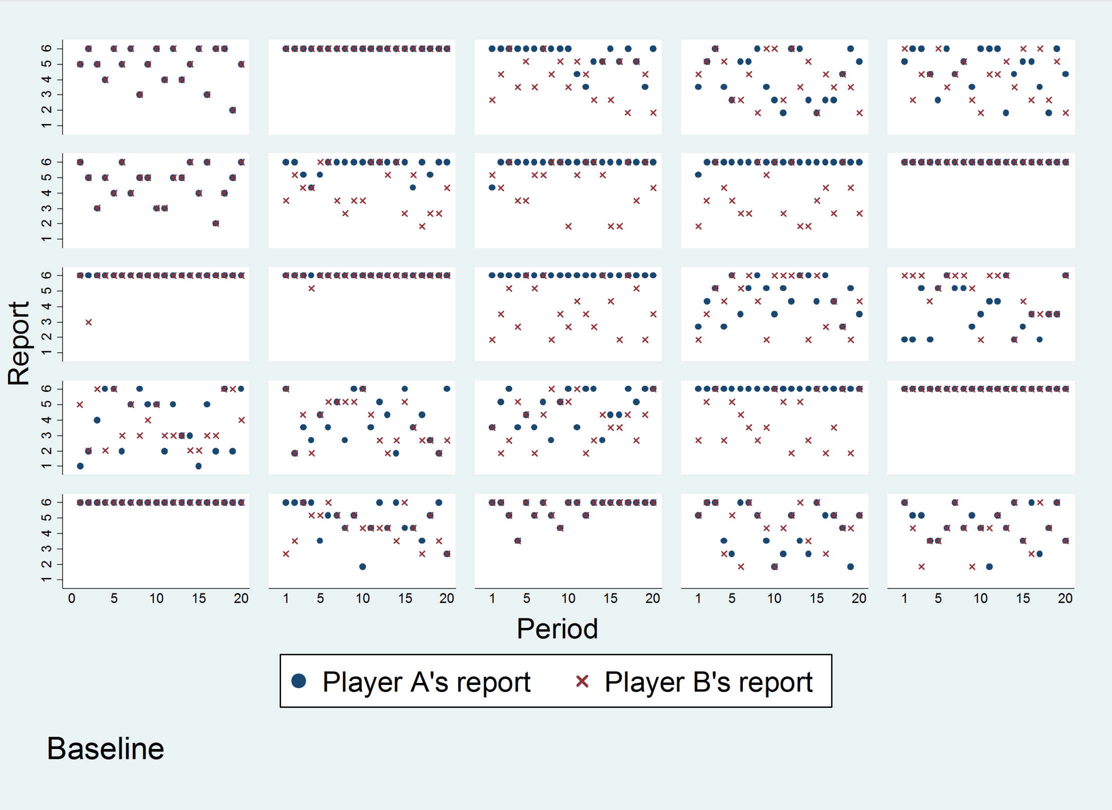
These results suggest that Conflict of Interest awakens social preferences at the expenses of the profit-maximization mindset which is proper of groups.

3.4.1.3 Evolution of Collaborative Dishonesty

Figure 3.8 shows the effect of *Conflict of Interest* on collaborative dishonesty over the 20 periods of the game. It shows that in *Conflict of Interest* collaborative dishonesty is substantially lower than in Baseline since the initial periods. However, doubles reported in *Conflict of Interest* seem to be similar to those reported in *Baseline* in the second half of the game.

We investigate this insight with a random individual effect probit regression analysis

Figure 3.7: Decisions by Groups Across Periods: Baseline and Conflict of Interest



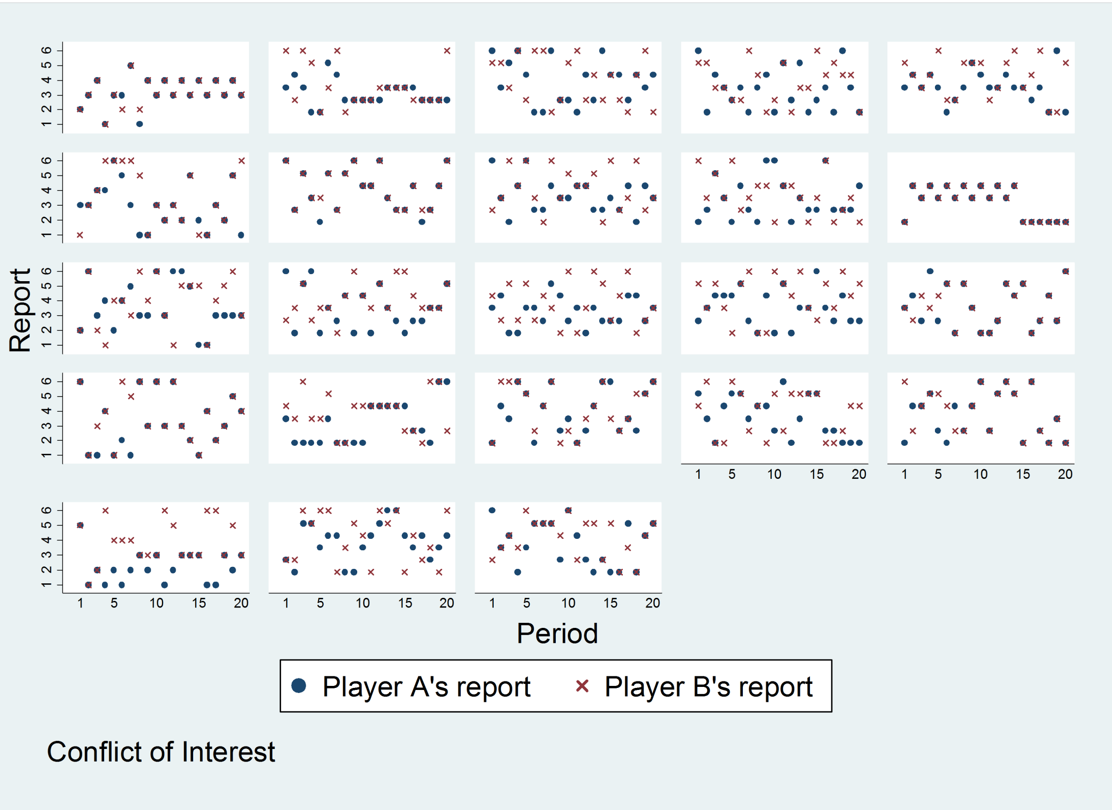


Figure 3.8: Conflict of Interest: Doubles over period

of Player B's decision to report a double to isolate the time effect (Table 3.4). We include a dummy for a block of ten periods (*11-20*), using the first 10 periods as baseline. We add the dummy for the treatment *Conflict of Interest* (*CI*). Finally, we add the interaction between *Conflict of Interest* and *11-20* to capture the combined effect. Results show that collaborative dishonesty is likely to increase in the last periods, but that the effect of *Conflict of Interest* is negative enough to make collaborative dishonesty less likely over time.

3.4.2 Discussion

Results of Experiment 2 indicate that introducing a conflict of interest in group settings mitigates the diffusion of negative externalities (i.e. diffusion of corruption) associated to teams with a joint stake in dishonest decisions. Adding a mechanism that enhances social comparison within the group awakens individuals' concerns for inequality aversion and decreases those for profit-maximization which are proper of groups. This way, individuals converge more slowly to a combined dishonest action.

While in the baseline individuals in groups coordinate fast on efficient outcomes because they are concerned about profit-maximization (Feri *et al.*, 2009), when faced

Table 3.4: Random individual effect probit regression analysis of Double decisions: Conflict of Interest

VARIABLES	(1) Double	(2) Double
CI	-0.395	-0.574*
	(0.331)	(0.344)
11-20	0.215**	0.0192
	(0.0948)	(0.139)
11-20 x CI		0.368*
		(0.190)
Constant	0.341	0.437*
	(0.239)	(0.245)
Observations	960	960
Number of Individuals	48	48

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

with unequal rewards they become more concerned about how their individual stake differs from that of the counterpart, rather than about profit-maximization. As a result, they take more time to converge to a combined action.

Our results show that introducing conflict of interest in a setting in which individuals have a joint stake in corruption creates noise within the group by making individuals more focused about reducing inequality and less concerned about increasing earnings through collaborative dishonesty. Results from Experiment 2 also suggests that, while lowering the diffusion of collaborative dishonesty, conflict of interest makes individual dishonesty hard to be detected. Conflict of interest gives individuals the incentive to report as if they were honest, with the joint stake to report the same outcome. In this setting, individual behavior is not a good predictor for dishonesty, while is so collaborative dishonesty.

3.5 Experiment 3

Our third experiment examines whether loss aversion further promotes dishonesty in group settings. Building on Grolleau *et al.* (2016) and Schindler & Pfattheicher (2016) findings on the detrimental effect that loss aversion has on individual honesty, we test how it interacts with group settings that promote collaborative dishonesty. Therefore, in Experiment 3 we add to the "dyadic die-rolling" paradigm the threat of loss aversion. The

Loss Frame Condition (Loss Frame, 2 sessions, 48 subjects) is implemented by providing participants at the beginning of each of period the maximum payoff attainable from reporting a double in the Baseline condition (i.e. reporting a double of 6 gives 120). In the instructions, participants were told that they could lose all points if they do not report a double (excluding the Baseline endowment of 46), while they could lose a certain amount of points if they report a double as shown in the column Loss Frame in Table 3.5 [9].

Table 3.5: Baseline and Loss frame payoff schedule and final profit

Player A's outcome	Player B's outcome	Baseline	Baseline Profit	Loss Frame	Loss Frame Profit
1	1	20	46+20=66	-100	166-100=66
2	2	40	46+40=86	-80	166-80=86
3	3	60	46+60=106	-60	166-60=106
4	4	80	46+80=126	-40	166-40=126
5	5	100	46+100=146	-20	166-20=146
6	6	120	46+120=166	0	166-0=166

3.5.1 Results

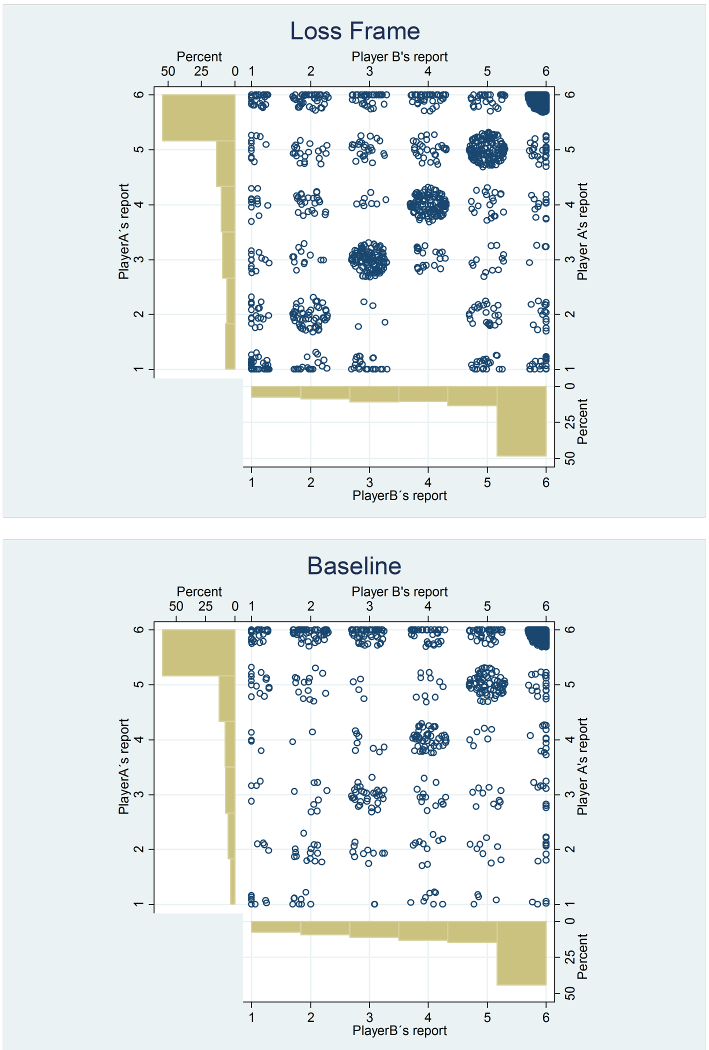
3.5.1.1 Collaborative and Individual Dishonesty

Figure 3.9 provides a joint representation of individual and collaborative dishonesty. As before, to investigate collaborative dishonesty, we look at the amount of doubles. The number of doubles are represented by the dots in the diagonal, that is the joint report of Player A and Player B in a single period for the *Loss Frame* (Loss Frame, 24 pairs x 20 periods = 480 dots) and *Baseline* conditions. Figure 3.9 shows that Players dishonestly collaborated. In fact, the average percentages of reported doubles in *Loss Frame* is significantly higher than the average percentage assuming honesty (16.7%) (Wilcoxon signed-rank test: z= 4.296, p-value=0.0000). However, the average percentage of doubles in *Loss Frame* (69.8%, sd:0.28) is not different from that in *Baseline* (56,6%, sd: 0.343) (Wilcoxon rank-sum test: z= -1.203, p-value= 0.2288).

The side histograms in Figure 3.9 show the distribution of numbers reported for each player. The not-uniform distributions suggest that both players reported the number dishonestly. As for the *Baseline*, both players A and B in *Loss Frame* report a number significantly higher than the expected 3.5 assuming honesty (Wilcoxon signed-rank test: A: z= 4.202, p-value 0.000; B: z= 3.758, p-value=0.0002).

[9] $e_{loss} = e_{gain} + \pi_{gain,max} = 46 + 120 = 166$

Figure 3.9: Loss Frame - Baseline



Player A in *Loss Frame* (4.79: sd: 0.90) report a number that does not differ from that reported by Player A in *Baseline* (5.064: sd: 0.932) (Wilcoxon rank-sum test: z= 1.296, p-value= 0.1949).

Similarly, Player B in *Loss Frame* (4.59: sd: 0.94) report a number that does not differ from that reported by Player B in *Baseline* (4.49: sd: 0.98) (Wilcoxon rank-sum test: z= -0.260, p-value=0.7947).

3.5.1.2 Analysis of Types

Fig 3.10 provides a joint representation of all decisions made by each group in *Baseline* and *Loss Frame*. It shows that individual behaviour across groups can be categorized into three main types of behaviour can be observed: i) brazen (individuals who reported 6 across all the 20 periods); ii) dishonest, i.e. individuals who reported on average a number larger than the honest benchmark but less than the maximum (6) across the 20 periods; iii) honest, i.e. individuals who reported on average truthfully (3.5) across the 20 periods.

The proportion of Honest Players across conditions is homogeneous. The proportion of Honest Players A in *Loss Frame* (8%(0.28)) is the same as the proportion of Honest Player A in *Baseline* (4%(0.2), Wilcoxon Rank Sum test: z = -0.626 p-value = 0.5313). Similarly, the proportion of Honest Players B in *Loss Frame* (21%(0.41)) is not different from that in *Baseline* (9%(0.29), Wilcoxon Rank Sum test: z = -0.828 p-value = 0.4078).

As for Honest Players, the proportion of Dishonest Players is homogeneous across conditions. The proportion of Dishonest Players A in *Loss Frame* (84%(0.38)) is not different from that in *Baseline* (64%(0.49), Wilcoxon Rank Sum test: z = -1.517 p-value = 0.1294). Similarly, the proportion of Dishonest Player B in *Loss Frame* (75%(0.44)) is not different from that in *Baseline* (72%(0.46), Wilcoxon Rank Sum test: z = -0.235 p-value = 0.8139).

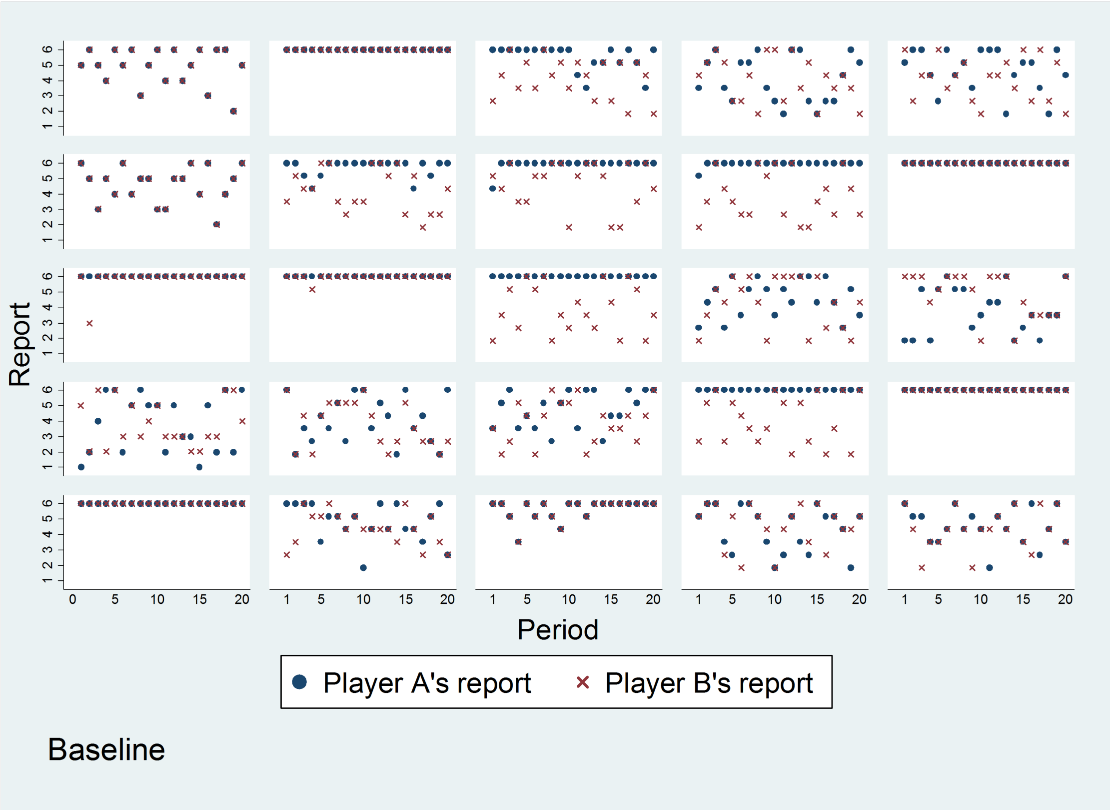
Interestingly, while the proportion of Brazen Player B in *Loss Frame* (4%(0.20)) does not differ from that in *Baseline* (16%(0.37), Wilcoxon Rank Sum test: z = 1.354 p-value = 0.1758), we find that the proportion of Brazen Player A in *Loss Frame* (8%(0.28)) is significantly higher than that in *Baseline* (32%(0.48), Wilcoxon Rank Sum test: z = 2.034 p-value = 0.0420). This suggests that loss aversion makes it more difficult for Player A to push Player B towards brazen corruption.

3.5.1.3 Evolution of Collaborative Dishonesty

Figure 3.11 shows the effect of *Loss Frame* on collaborative dishonesty over the 20 periods of the game. It shows that in *Conflict of Interest* collaborative dishonesty is higher than in Baseline especially in the second half of the game.

We isolate the time effect on collaborative dishonesty by running a random individual effect probit regression analysis of Player B's decision to report a double (Table 3.6). We include a dummy for a block of ten periods (*11-20*), using the first 10 periods as baseline. We add the dummy for the treatment *Loss Frame* (*LOSS*). Finally, we add the

Figure 3.10: Decisions by Groups Across Periods: Baseline and Loss Frame



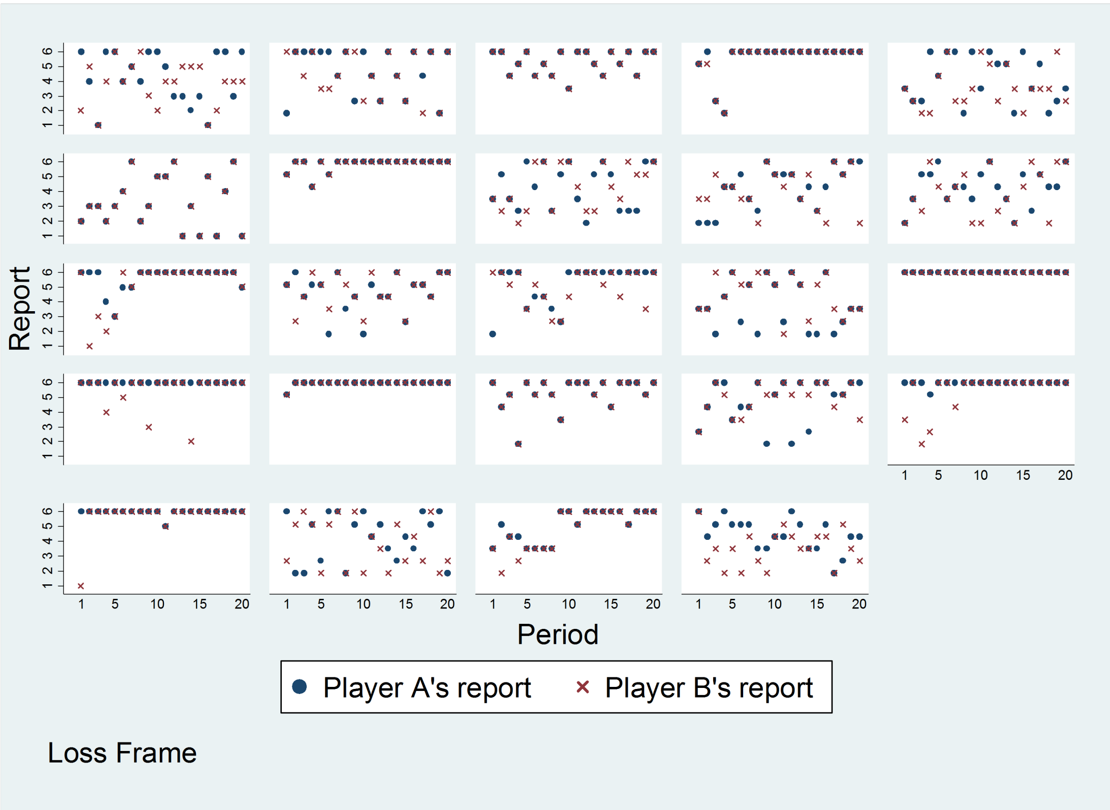


Figure 3.11: Loss Frame: Doubles over time

interaction between *Loss Frame* and *11-20* to capture the combined effect. Results show that collaborative dishonesty in *Loss Frame* is not likely to change over the last periods compared to Baseline.

This non-finding suggests that loss aversion does not amplify the negative sides of a suboptimal environment that forces individuals to corrupt to earn (avoid) some benefits (losses).

3.5.2 Discussion

Results of Experiment 3 show that loss aversion does not amplify the negative sides of suboptimal group settings. Differently from previous research on the interaction between loss aversion and individual dishonest behavior (Kocher *et al.*, 2016; Schindler & Pfattheicher, 2016), our results suggest that loss aversion does not interact with suboptimal group settings, as the mere exposure to them is sufficient to change perceptions of ethical values.

Table 3.6: Random individual effect probit regression analysis of Double decisions: Loss Frame

VARIABLES	(1) Double	(2) Double
LOSS	0.375	0.251
	(0.393)	(0.404)
11-20	0.149	0.0201
	(0.101)	(0.140)
11-20 x LOSS		0.266
		(0.202)
Constant	0.432	0.495*
	(0.283)	(0.287)
Observations	980	980
Number of Individuals	49	49

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

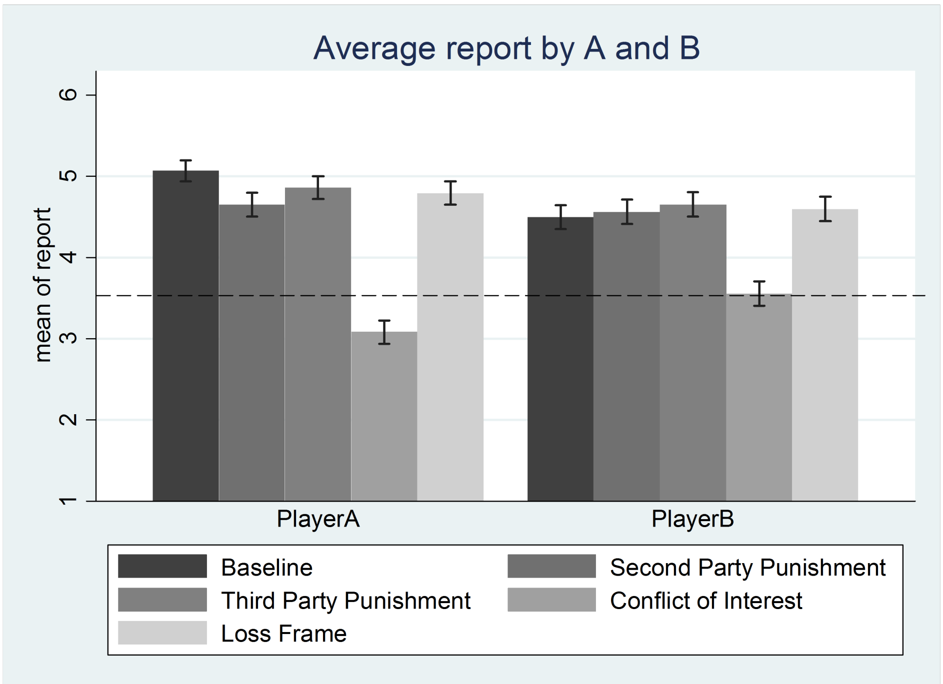
3.6 Personality Traits and Risk Preferences

In all experiments, we asked participants to complete two additional tasks: the questionnaire based on the Hexaco model of personality traits (Ashton & Lee, 2009) and the Bomb Risk Elicitation task (BRET) developed by Crosetto & Filippin (2013).

Among the six measures of personality traits, the Honesty-Humility measure is especially crucial for our study since it represents "the tendency to be fair and genuine in dealing with others, in the sense of cooperating with others even when one might exploit them without suffering retaliation" (Ashton & Lee (2007): p. 156). Hilbig *et al.* (2012) show that cooperation (i.e. contributions to public goods) depends on high dispositional levels of Honesty-Humility. However, in our setting cooperation is against ethical norms and far from being fair. We investigate whether high dispositional levels of Honesty-Humility decrease the likelihood to act dishonestly.

According to traditional economic approach to dishonesty (Becker, 1968), individual attitudes toward risk play a crucial role in the assessment of expected costs and, thus, on the subsequent choice to act dishonestly. On the other hand, extensive experimental evidence suggests that individual decision to act dishonestly is influenced by a much broader set of factors than just the costs and benefits associated with the act. In particular, in a setting characterized by a normative conflict, individual risk preference might not

Figure 3.12: Average report by A and B by treatment



The dashed line is the expected number reported assuming honesty (3.5)

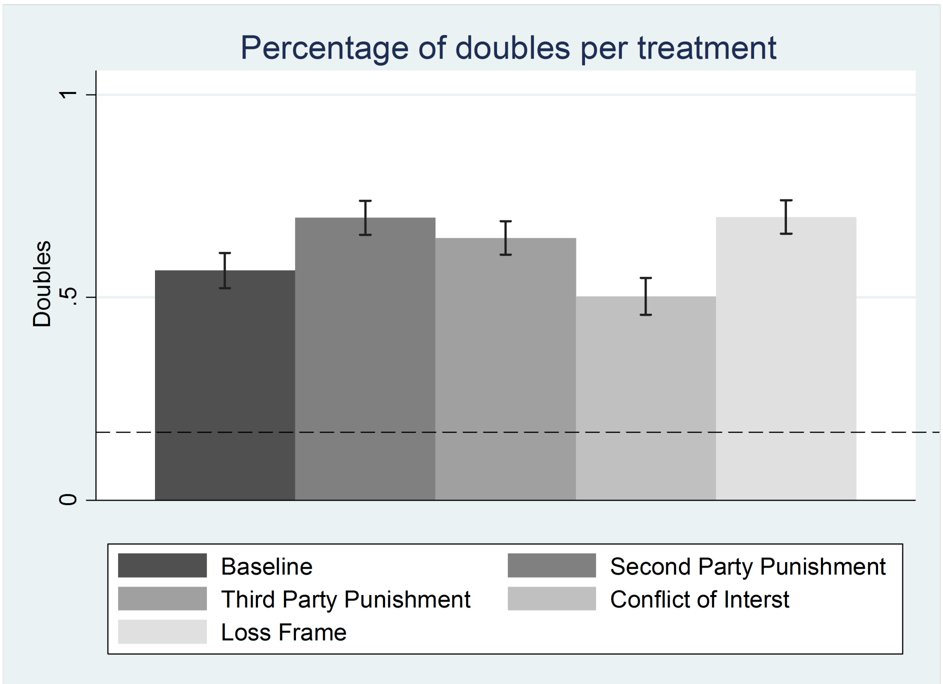
play the substantial role that is played by other-regarding and moral concerns. We derive an intuitive measure of risk aversion to investigate whether risk preferences have an effect on the likelihood to make dishonest decisions in a context that triggers moral and other-regarding concerns.

These two additional measures allow us to further disentangle across treatments the determinants behind dishonest reporting by each player (Figure 3.13), and dishonest cooperation by Player B (Figure 3.12).

First, we run a Random Effect Tobit model [10] (Table 3.7) with reported number by Player A (Model a) and Player B (Model b) as dependent variable. Among the explanatory variables we include the dummies variables identifying the experimental conditions: SP is equal to 1 if the Player is in the Second Party Punishment condition, TP is equal to 1 if the Player is in the Third Party Punishment condition, CI is equal to 1 if the Player is in the Conflict of Interest condition, LOSS is equal to 1 if the Player is in the Loss Frame condition. The model is enriched with the personality measure for Honesty-Humility and risk aversion (Risky choice).

[10]We run a Tobit model to account for potential censoring in the data from the dependent variable - bounded between 1 and 6.

Figure 3.13: Percentage of doubles per treatment



The dashed line is the expected percentage of doubles assuming honesty (16.7%)

As Table 3.7 shows, the dummy associated to Player A's decision to report a higher number in Second Party Punishment is negative and significant: the threat of being sanctioned by Player B for bullying her to act brazenly against honesty decreases the likelihood to report a higher numbers. This confirms the presence of a normative conflict characterizing our suboptimal setting. Consistently with the incentive scheme introduced in *Conflict of Interest* to report a number around the optimal value of 3.5, the dummy associated to both Player A's and Player B's decision to report higher number in Conflict of Interest is negative and significant. Finally, individuals with high dispositional Honesty-Humility are less willing to report high numbers.

Second, we run an Individual Random Effect Probit model (Table 3.8) with the decision to report a double by Player B as dependent variable. As before, the independent variables are the dummies identifying the experimental conditions: SP is equal to 1 if the Player is in the Second Party Punishment condition, TP is equal to 1 if the Player is in the Third Party Punishment condition, CI is equal to 1 if the Player is in the Conflict of Interest condition, LOSS is equal to 1 if the Player is in the Loss Frame condition. The model is enriched with the personality measure for Honesty-Humility and risk aversion (Risky choice).

Table 3.7: Individual Random Effect Tobit model: Player A and Player B Report

VARIABLES	(a) $Report_A$	(a) $Report_A$	(b) $Report_B$	(b) $Report_B$
SP	-1.452*	-1.568*	-0.0975	-0.177
	(0.816)	(0.809)	(0.691)	(0.684)
TP	-0.451	-0.441	0.514	0.652
	(0.800)	(0.806)	(0.676)	(0.673)
CI	-4.569***	-4.652***	-2.196***	-2.104***
	(0.811)	(0.804)	(0.686)	(0.679)
LOSS	-1.178	-1.155	-0.0262	0.0228
	(0.806)	(0.811)	(0.683)	(0.684)
Honesty Humility		-0.575*		-0.576*
		(0.366)		(0.343)
Risky Choice		0.0112		-0.00687
		(0.0146)		(0.0128)
Constant	7.477***	8.938***	5.789***	8.041***
	(0.580)	(1.638)	(0.482)	(1.416)
Observations	2,420	2,420	2,420	2,420
Number of Individuals	121	121	121	121

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As Table 3.8 shows, individuals with high dispositional Honesty-Humility are less willing to dishonestly collaborate. Moreover, it shows that *Second Party Punishment*, *Third Party Punishment*, *Conflict of Interest* and *Loss Frame* have no effect on the probability to report a double. In order for *Second Party Punishment*, *Third Party Punishment* and *Conflict of Interest* to have an effect on collaborative dishonesty, we have to control for the time effect, as we reported in the result sessions of each experiment. This is consistent with the insight that suboptimal settings shape normative perceptions of corruption over time.

3.7 General Discussion and Conclusions

In many real-life situations, it is the combined of actions of many individuals belonging to a group rather than the isolated decision of an individual that produce undesirable outcomes. Individuals belonging to teams, families, administrative divisions or organizations may contribute to spread group-specific values and create suboptimal self-enforcing

Table 3.8: Random Effect Individual Probit model: Double decision

VARIABLES	(1) Double	(2) Double
SP	0.311	0.239
	(0.358)	(0.353)
TP	0.161	0.253
	(0.345)	(0.341)
CI	-0.413	-0.370
	(0.352)	(0.347)
LOSS	0.375	0.466
	(0.354)	(0.353)
Honesty Humility		-0.365**
		(0.178)
Risky Choice		0.00242
		(0.00650)
Constant	0.468*	1.576**
	(0.249)	(0.727)
Observations	2,420	2,420
Number of Individuals	121	121

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

institutions that have a negative influence on the members of a society. In our study, we use three laboratory experiments to investigate the diffusion of dishonest actions in group settings when three mechanisms are introduced. Our study focuses on suboptimal group settings, such as the one depicted by Weisel & Shalvi (2015): individuals are denied a fair access to opportunities, but are given the incentive to dishonestly collaborate to earn some benefits. We complement the argument by Gächter & Schulz (2016) that honesty is impaired by the surrounding context. We show that honesty in groups is negatively affected by the context when this allows for informal norm enforcement mechanisms, while it is not impaired when the joint stake in the dishonest decision is added a conflict of interest.

First, our findings add to the literature on suboptimal institutions (Herrmann *et al.*, 2008; Meier *et al.*, 2016) by showing that group settings that give the incentive to corrupt distort the typical efficacy of punishment mechanisms to enforce positive norms. Allowing for norm-enforcement mechanisms, such as second-party and third-party punishment, produces anti-social norms, i.e. collaborative dishonesty. Punishment makes individuals in the group cooperate more, however the type of cooperation involved is against ethical norms. Dark cooperation outweighs the moral costs associated to dishonesty and changes negative perceptions of dishonesty.

Second, our study complements evidence on the superior strategic mindset of groups to that of individuals (Kocher & Sutter, 2005; Sutter, 2009; Feri *et al.*, 2009). Adding unequal rewards in suboptimal group settings lowers the salience of the joint stake in dishonest collaboration while increasing that of individual stakes in comparison with those of others. This way, individuals in the group become more concerned about inequality than profit-maximization and, as a result, converge slower to dishonest coordination.

Finally, our results contribute to the literature studying the underlying mechanisms of dishonest behaviour extended to group settings. Contrary to studies on the relationship between loss aversion and individual dishonesty (Grolleau *et al.*, 2016; Schindler & Pfattheicher, 2016), we do not find that loss aversion has an effect on collaborative dishonesty when the group setting is suboptimally designed.

We also show that differences in ethical behaviour can be explained by personality traits, such as the Honesty-Humility factor of the model of personality structure (Hilbig & Zettler, 2015), also in suboptimal group settings.

Our study is of relevance for institutions and organizations that promote aligned incentives in group settings. These can be threatened by a culture of corruption when punishment among peers or non-credible monitoring are allowed. On the other hand, it shows that nesting inequality in the aligned outcomes scheme might represent a tool to weaken the diffusion of collaborative dishonesty, as team members become less group-focused and ready to identify corruption as the best strategy for the group.

Conclusions

This Doctoral Thesis exploits the laboratory experimental approach to disentangle the influence of suboptimal social contexts on individual decisions. Laboratory experimental method provides a clean tool to disentangle the role that several mechanisms play in complex scenarios, like suboptimal social environments: it allows to derive causal inference on how social patterns shape individuals' ultimate decisions. Laboratory experiments allow to collect rigorous evidence on the underlying mechanisms that drive behavior in situations that cannot be easily observed in the field. This is of particular relevance for research aimed at disentangling social determinants of preferences. It is very difficult to isolate the role that social mental models have on behavior in the field: throughout their life, individuals are exposed to several environments that may contribute to develop complex responses to different scenarios.

The three essays of this Doctoral Thesis manipulate social context with the aim to isolate social determinants on behavior.

The First Chapter focused on experiences of unfairness in a social context in which group identity is salient, inspired by the evidence that episodes of unfairness promote unethical decisions. We manipulated group identity to induce a mental model such that individuals perceive individuals sharing the same group identity differently from those with a conflicting group identity. This treatment was necessary to investigate whether individuals tolerate unfairness by an ingroup differently from that made by an outgroup. The results support the predictions that when social context makes a particular group identity salient, individuals change their perceptions of unfairness and, in turn, react accordingly in the form of unethical behavior. The design focused on the peril of allowing for the coexistence of different cultures across organizational divisions and subunits. The findings are of practical relevance for organizations which have to deliver a wage that does not reflect the employee's effort. By failing to build a shared corporate culture, they might foster employees' fairness-restorative dishonest behavior which is costly to society.

The Second Chapter investigated the role of a cultural mental model, like that of gender stereotype, in decisions that require predicting others' risk attitudes. We first assessed whether the belief that women are more risk averse than men is accurate by structurally estimating risk preferences under two alternative models of choice. Second, we assessed the extent according to which individuals exploit gender information to predict others' risk attitude. Our findings suggest that predictions of others' risk attitude are biased by gender stereotype: they emphasize the need to avoid providing gender information to individuals that make decisions in which one's risk attitude has to be taken into account.

The Third Chapter focused on a social context that induces individuals to gain benefits other than through collaborative dishonesty. This context was the setting of Weisel & Shalvi (2015) and mimics real world situations, such as organizations that equally reward employees working in teams only if the team performs well. This essay analyzed the impact that punishment, conflict of interest and loss aversion have on perceptions of corruption. Our results suggest that introducing a norm enforcement mechanism, such as second party and third party punishment, augments the negative externalities associated to suboptimal social contexts: this enforces the diffusion of collaborative dishonesty. Organizations that allow for peer punishment and non-credible monitoring are at risk of facing a dramatic diffusion of cases of collaborative dishonesty. Second, it shows that introducing a conflict of interest in the aligned incentive scheme lowers the diffusion of corruption. Organizations that introduces inequality in the team reward scheme are less likely to face the diffusion of collaborative dishonesty. Finally, our results suggest that loss aversion does not promote collaborative dishonesty, as the mere exposure to suboptimal group settings is sufficient to change perceptions of ethical values.

Limitations and further research

One of the most popular critique to laboratory experimental findings is the extent to which they can be generalized to the real world, as they lack external validity (Levitt & List, 2007). However, the aim of laboratory experiments is not to provide generalizability, but a benchmark to policy makers and organizations before they implement costly interventions. Laboratory experiments allow to isolate the effect of a desired contextual change on behavior in a fully controlled environment. Moreover, they allow to derive rigorous evidence on behaviors that correlate well with behaviors outside the lab (Gächter & Schulz, 2016).

However, as for any laboratory evidence, the findings of this research would highly benefit from complementing evidence from field experiments. The first Chapter suffers from manipulating group identity in the laboratory. Building a real group identity in the field would probably create a real change in the self-concept that, in turn, cues the mental model of differently interpreting interactions with individuals belonging to the same or to a different group.

The Second Chapter restricts the investigation of predictions of others' risk attitude to female and male targets to address gender stereotype bias. However, false consensus bias is also crucial in predictions of others' behavior. A follow-up aimed at better isolating false consensus bias from gender stereotype might include a treatment in which no information about the target is provided.

Results from Experiment 3 in the Third Chapter might be corroborated by running a control condition aimed at comparing behavior in individual settings and in group settings. This would allow to better grasp framing effect on dishonest behavior.

APPENDIX A

APPENDIX A

A.1 Dictator's Dishonest Behavior

As for Recipients, to address the causal relationship between unfairness and dictators' dishonest behavior given different types of salient group identity we run a regression model (Table A.1). The dependent variable $Reported.value$ is given by the integer reported in the dishonesty task. Among explanatory variables, we have dummy variables controlling for group identity conditions: IN is equal to one when the dictator belongs to the same group and zero otherwise; OUT is equal to one when the dictator belongs to the other group and zero otherwise. BASE is the baseline condition.

We enrich the model with fairness-related variables: the proportionality index ϕ ($prop.index$, see Table 1.1) of dictator's choices and the extent according to which proportionality is perceived as appropriate in the *Social Norm* task ($prop.norm$, see Table 1.2). The interactions between the fairness-related variables and the group experimental conditions are also added. Finally, we consider a few control variables: $report.time$ (the time in seconds required to report the value); age; $female$, and $civic.score$ (a categorical variable of individuals' participation to collective activities, such as political parties and NGOs).

As Table A.1 shows, we find that group identity does not affect dictators' decision to report dishonestly. In fact, the treatment dummies are not statistically significant. Similarly, we find that higher violations of proportionality predicts higher willingness to report dishonestly, but unconditionally on the group identity conditions.

Table A.1: Dictators' self-reported values (Ordered Probit regression)

	(1)	(2)	(3)
	rep.value	*rep.value*	*rep.value*
prop.index	0.665^{**}(0.299)	0.814(0.573)	1.039(0.582)*
prop.norm	-0.218(0.253)	-0.505(0.427)	-0.580(0.437)
IN		-0.0829(0.782)	-0.293(0.800)
OUT		-0.147(0.628)	-0.0833(0.642)
$prop.index \times IN$		0.150(0.803)	0.166(0.818)
$prop.index \times OUT$		-0.428(0.742)	-0.835(0.777)
$prop.norm \times IN$		0.290(0.692)	0.288(0.711)
$prop.norm \times OUT$		0.433(0.592)	0.226(0.608)
report.time			0.00956(0.00557)*
female			-0.430(0.239)*
age			-0.0952(0.0485)**
civic.score			-0.180(0.102)
Observations	95	95	95

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

An Ordered Probit regression model is adopted to account for the different attitude individuals display when they report increasingly high values.

In fact, mirroring Konow (2000), we found that, unconditionally on the experimental conditions, on average, dictators allocate to themselves more than what they earned because they have a personal stake in the allocation choice outcome (Table 1.1). The evidence that higher violations of proprionality principle lead to higher reported values suggests that dictators infer their own type and values from their past choices (Bénabou & Tirole, 2011): the more they have been selfish in the allocation task, the more they are selfish in the dishonesty task. This insight is supported by the evidence that dictators taking more time to report the value, are more likely to report higher values: to infer who they are, dictators require more time.

Finally, we find that dictators' dishonest behaviour correlates with observable characteristics. Women are less likely to engage in dishonest behaviour than men (Dreber & Johannesson, 2008; Houser *et al.*, 2012; Bucciol *et al.*, 2013). The likelihood to engage in dishonest behaviour decreases with age since adults find it easier to self-restrain (Glätzle-Rützler & Lergetporer, 2015).

A.2 Robustness Check

Table A.2: Recipients' self-reported values (Tobit regression)

	(1)	(2)	(3)
	rep.value	*rep.value*	*rep.value*
prop.index	0.048(0.998)	-1.92(2.066)	-2.232(2.05)
prop.norm	-0.661(0.981)	-2.394(1.88)	-2.562(1.87)
IN		-1.27(2.473)	-1.663(2.461)
OUT		-3.023(2.581)	-2.775(2.553)
$prop.index \times IN$		1.329(2.466)	1.872(2.481)
$prop.index \times OUT$		5.686(2.884)*	6.095(2.913)**
$prop.norm \times IN$		1.512(2.411)	1.607(2.422)
$prop.norm \times OUT$		3.212(2.489)	2.622(2.499)
report.time			0.0141(0.0249)
female			-1.12(0.941)
age			-0.0487(0.172)
civic.score			0.0498(0.421)
Constant	8.204(0.967)***	9.669(2.072)***	10.96(4.519)**
sigma			
Constant	4.051(0.434)***	3.862(0.412)***	3.812(0.406)***
Observations	93	93	93

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Tobit regression model is adopted to account for bounded support of choices (ul=9, ll=0). Standard errors in parentheses. Significance symbols: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.3 Screenshots from the experiment

Figure A.1: Real Effort Task

Figure A.2: Group Identity: Guess Task.

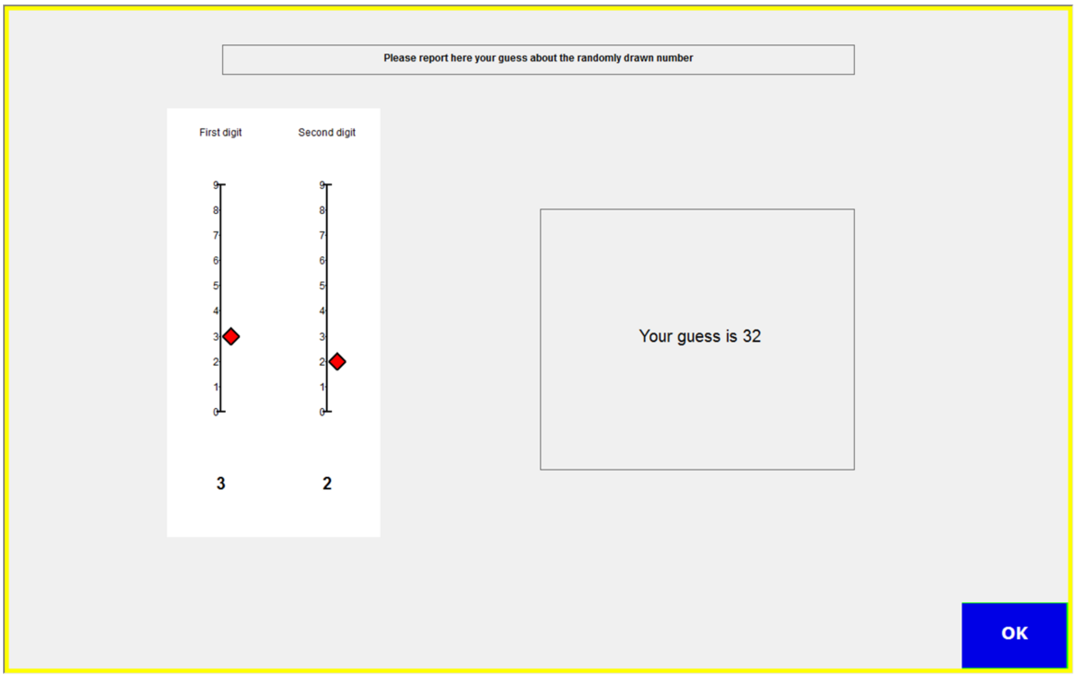


Figure A.3: Group Identity: Group Assignment.

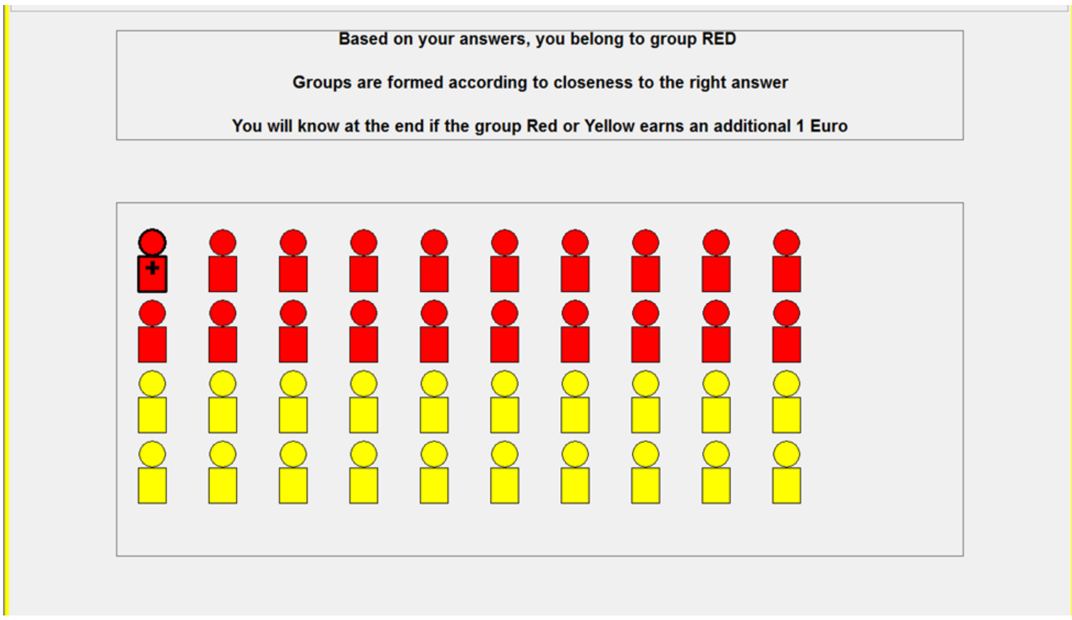


Figure A.4: Group Identity: Proverb task.

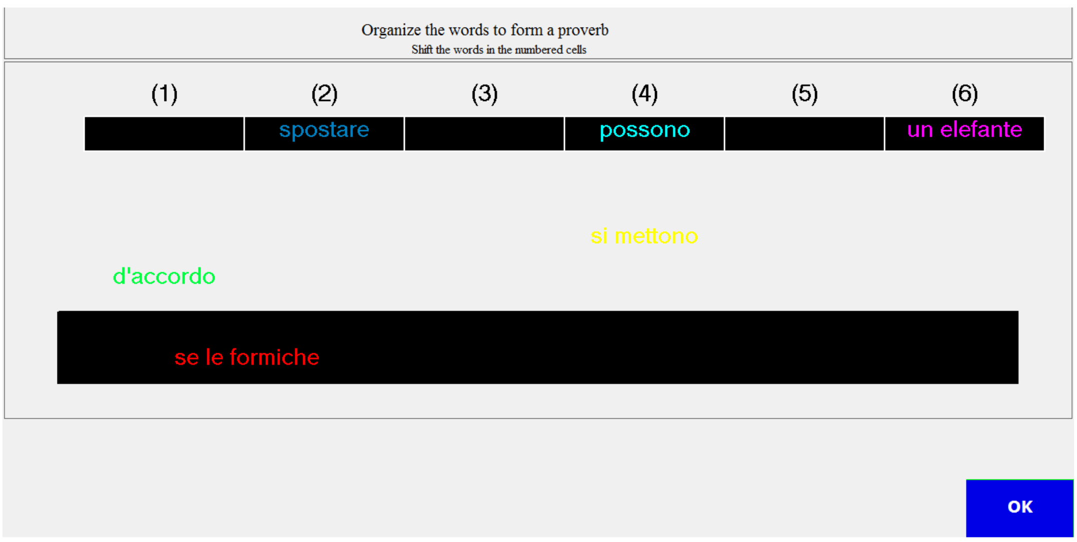


Figure A.5: Dictator Game.

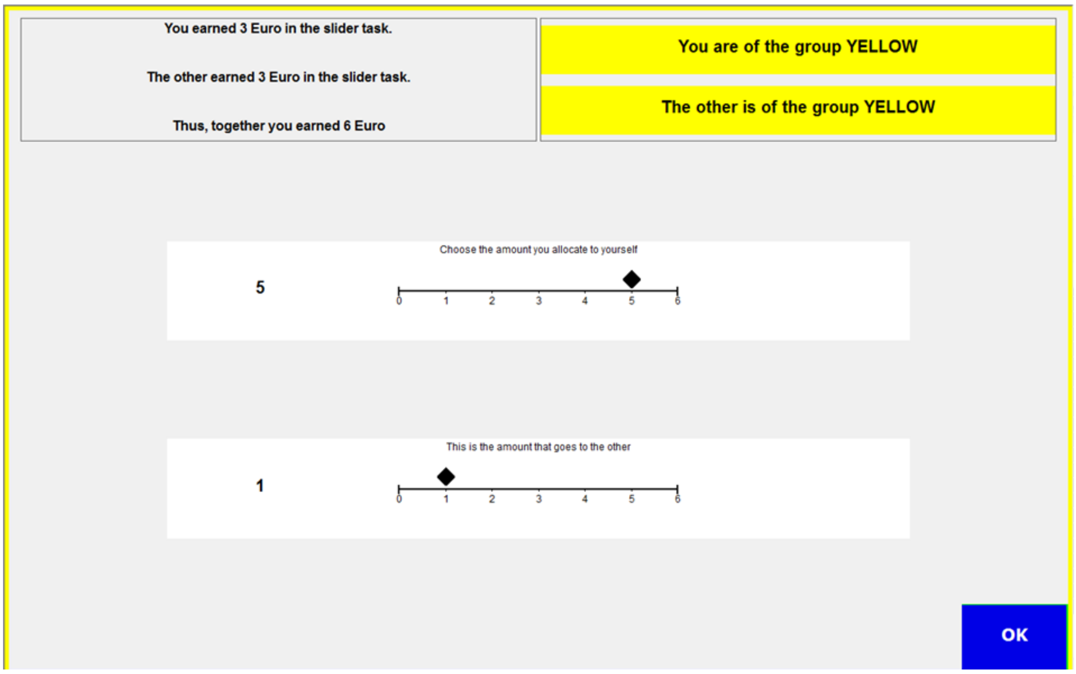
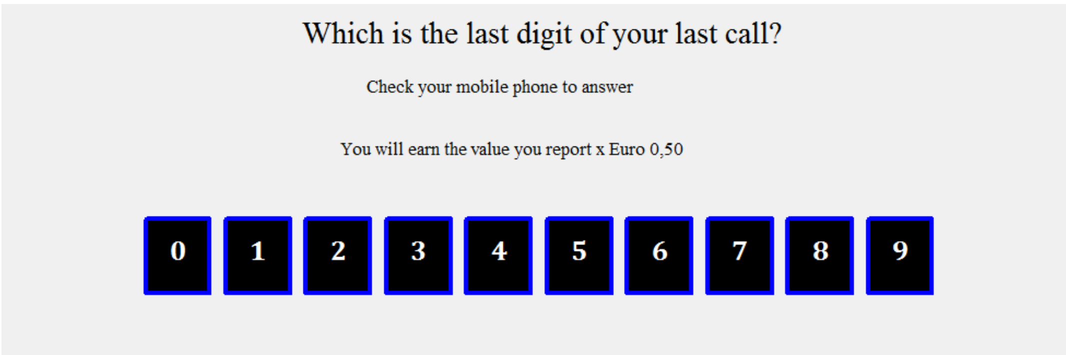


Figure A.6: Dishonesty task.



A.4 Instructions (translated)

You are now taking part in an economic experiment which has been financed by various foundations for research purposes. Please read carefully the instructions that we have distributed to you. During the experiment you will have the opportunity to earn a sum of money that will depend on your actions, your decisions, the other participants' decisions and some random factors. You will receive this sum of money at the end of the experiment. You will earn anyway €3 for showing up to the experiment.

It is prohibited to communicate with the other participants during the experiment. If you violate this rule, we shall have to exclude you from the experiment and from all payments. Should you have any questions please raise your hand: a lab assistant will come to your place.

During the experiment your earning will be calculated in tokens. At the end of the experiment the total amount of tokens you have earned will be converted into real money at the following rate:

1 token = 1 €

You are free to leave the experiment if you want to, however you will not receive any sum of money.

During the experiment you will have the opportunity of making choices that will influence both your earning and that of other participants. The choices made by each subject will be totally anonymous.
Anonymity will be maintained both during and after the experiment: all the money you will earn will be privately paid in another room when the experiment will be over.

General overview

Please read carefully the description of the situation you are about to face. You and other fifteen people are participating in the experimental session. The experiment is made of four stages. All tasks will be computerized. After completing each stage, the next stage instructions will appear on the screen of your computer. As is the case in all economics experiments, we will always provide you true information that never deceives you in any way.

Slider Task

In the first stage you will have the opportunity to earn a portion of your final earning. After reading the instructions, a white page with six sliders will appear on the screen of your computer. Each slider is initially positioned at 0 and can be moved as far as 100. You can use your mouse or touchscreen to move each slider. Your goal is to position the slider at the value shown on its right. Once you have positioned the slider at the goal value, the value shown on its right will turn from red to green. The current slider position is on the left of the slider. You can readjust the position of each slider as many times as you want. After adjusting the six sliders in each page, a new page with six sliders will appear on the screen of your computer. The total number of pages you will complete within 300 seconds will be the first part of your earning. The second stage instructions will appear on the screen of your computer once the 300 seconds will be over.

Guessing task

After completing the first stage, you will be asked to answer a simple question that will appear on the screen of your computer. You can use your mouse or touchscreen to answer to this question. You will be asked to choose a random number between 1 and 99. To communicate the number you choose you will have to position two sliders: the position of the first slider will be the tens of your number, while the position of the second slider will be the units of your number. Depending on your answers, you will be assigned to the Red group or the Yellow group. The division in groups will take place according to a similarity/distance criterion with a number randomly chosen by the computer. Participants who will choose the closest numbers to the one randomly chosen by the computer will be assigned to one colour, while participants who will choose the farther numbers to the one randomly chosen by the computer will be assigned to the other colour. The colors will be randomly assigned by the computer to the criterion of similarity and distance.
After communicating the number you have chosen, you will be shown if you have been assigned to the Red or Yellow group on the screen of your computer.

Proverb task

At this stage you will be asked to complete a task together with participants assigned to the same group colour as yours. Particularly, you and the other mates will be shown a series of words and letters. You and your mates will be asked to organize the words and letters to form a proverb. There is no time constraint. However, only the fastest group at completing the proverb will enable all group members to earn an additional amount of money at the end of the experiment.

Feedback

After completing the proverb bask together with your mates, you will be shown on the screen of your computer the total number of pages you have completed in the slider task. Before starting the next stage, the computer will match you with another participant. You will be informed about the group affiliation and your role. Your partner and your role will be randomly chosen by the computer.

Allocation task

In this stage of the experiment you will be asked to complete a task with the partner you have been previously informed of.

First, you and your partner will be shown the sum of the partner's and your earnings from the slider task. Depending on the role you have been randomly assigned, you will be shown the details on your computer screen. If you have been assigned to the role of dictator, you will be asked to decide how to divide the sum of the earnings between you and your partner. Your partner will be shown the amount you will offer at the end of the task. If you have been assigned to the role of recipient, you will have to wait your partner's offer.

Questionnaire

After completing the allocation task, you will be asked to answer to a short questionnaire. You will have the opportunity of earning an additional amount of money for your time. After answering all questions, you will receive a final feedback about the additional earnings from completing all tasks. At this stage, you will have to wait for a lab assistant who will call your seat number for being paid in the other room.
We would also be grateful if you did not discuss the experiment with the other participants outside the laboratory.

A.5 On-screen Questionnaire

Allocation task

Recipient Please indicate how much you agree with the following statements (1-not at all 7-very much):

- The amount offered by the dictator is fair ☐
- The amount offered by the dictator is fair given the dictator's and my results from the slider task ☐
- The amount offered by the dictator is fair given the effort the dictator and I have exerted in the slider task ☐
- If I had been assigned to the role of dictator, I would have offered the same amount the dictator has offered to me ☐
- If you do not agree with the previous statements, please indicate the amount you would have offered if you had been assigned to the role of dictator ☐

Dictator Please indicate how much you agree with the following statements (1-not at all 7-very much):

- The amount I offered is fair ☐
- The amount I offered is fair given my and the recipient's results from the slider task ☐
- The amount I offered is fair given the effort the recipient and I have exerted in the slider task ☐
- I would have offered a different amount if the initial sum to divide had been different ☐
- The offer I made to the recipient was based on the information about the effort that the recipient and I have exerted in the slider task ☐
- The offer I made to the recipient was based on the information about the recipient's group affiliation ☐

All
Individual A finds €10 in the street. He decides to share it with a pedestrian. The table below presents a list of the possible choices available to Individual A. How much socially appropriate do you believe is each option?

Individual A	Very socially inappropriate	Socially inappropriate	Socially appropriate	Very socially appropriate
Offers 0, keeps 10				
Offers 2, keeps 8				
Offers 4, keeps 6				
Offers 5, keeps 5				
Offers 6, keeps 4				
Offers 8, keeps 2				
Offers 10, keeps 0				

All
Individual A finds €10 in the street. He meets a friend and decides to share it. The table below presents a list of the possible choices available to Individual A. How much socially appropriate do you believe is each option?

Individual A	Very socially inappropriate	Socially inappropriate	Socially appropriate	Very socially appropriate
Offers 0, keeps 10				
Offers 2, keeps 8				
Offers 4, keeps 6				
Offers 5, keeps 5				
Offers 6, keeps 4				
Offers 8, keeps 2				
Offers 10, keeps 0				

Group

1. Please indicate how much you agree with the following statements (1-not at all 7-very much):
   - I feel similar to the member of my color group □
   - I identify myself with the members of my color group □
   - I might behave differently with the members of the other color group □
2. Slide the circle "I" towards the circle "Other" to describe how connected you feel to the group you have been assigned.

How do you feel

Recipient Please indicate how much you agree with the following statements (1-not at all 7-very much):

- I feel I have been treated fairly by the dictator □
- I feel I have been treated fairly by the dictator given my result from the slider task □

- I feel I have been treated fairly by the dictator given the effort I exerted in the slider task ☐
- I was disappointed by the dictator's behavior given that the dictator is a member of my group color ☐
- I was angry because of the dictator's behavior given that the dictator is a member of my group color☐

Dictator Please indicate how much you agree with the following statements (1-not at all 7-very much):

- I feel I have treated fairly the recipient ☐
- I feel I was selfish at offering to the recipient the amount I chose ☐
- I felt unpleasant at offering to the recipient the amount I chose ☐
- I felt unpleasant at offering the amount I chose given that the recipient is a member of my group ☐

About you

- Department ☐
- Gender ☐
- Age ☐
- Only child ☐
- How big is the place you spent most of your life? ☐ (1:up to 2000 inhabitants abitanti; 2:2000-10000; 3:10000-100000; 4: more than 100000)
- How often do you attend religious events? ☐ (0:Never; 1: sometimes; 2:more than once in a week)
- Do you participate in one of the following organizations as member? (0:Never; 1: sometimes; 2:more than once in a week)
  - Sport club ☐
  - Choir, orchestra ☐

- Political party ☐
- NGOs ☐
- Other organizations ☐

As you spent some time answering this questionnaire, you have now the opportunity to earn an additional sum of money. Wait for the instructions on the screen of your computer.

APPENDIX

B

APPENDIX B

B.1 Robustness Check

Table B.1: Prediction Bias: Tobit Model

VARIABLES	$(Female_a)$ PC-AC	$(Female_b)$ PC-AC	$(Male_a)$ PC-AC	$(Male_b)$ PC-AC
$FemaleTarget$	0.00338	0.00318	-0.0843*	-0.0777*
	(0.0408)	(0.0364)	(0.0477)	(0.0434)
$PredictorChoice$		0.263***		0.289***
		(0.0364)		(0.0509)
$Constant$	0.0174	-0.0864**	0.03046**	-0.0541
	(0.0248)	(0.0790)	(0.00456)	(0.0340)
Observations	2,880	2,880	2,880	2,880

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

B.2 Prospect Parameters

Table B.2: Left Prospect Parameters

(S-C)	Lp1	Lp2	Lp3	La1	La2	La3	EV(L)
1	0.5	0.5	0	0	2	0	1
2	1	0	0	2	0	0	2
3	1	0	0	2	0	0	2
4	0.25	0.75	0	0	2	0	1.5
5	0.25	0.75	0	0	2	0	1.5
6	1	0	0	2	0	0	2
7	1	0	0	4	0	0	4
8	0.75	0.25	0	4	7	0	4.75
9	0.5	0.5	0	4	14	0	9
10	0.5	0.5	0	0	14	0	7
11	1	0	0	4	0	0	4
12	0.75	0.25	0	4	14	0	6.5
13	1	0	0	7	0	0	7
14	0.75	0.25	0	7	14	0	8.75
15	0.75	0.25	0	7	14	0	8.75
(S-AE)	Lp1	Lp2	Lp3	La1	La2	La3	EV(L)
1	0.5	0.5	0	0	2	0	1
2	1	0	0	2	0	0	2
3	1	0	0	2	0	0	2
4	0.75	0.25	0	2	0	0	1.5
5	0.75	0.25	0	2	0	0	1.5
6	1	0	0	2	0	0	2
7	1	0	0	4	0	0	4
8	0.25	0.75	0	7	4	0	4.75
9	0.5	0.5	0	4	14	0	9
10	0.5	0.5	0	0	14	0	7
11	1	0	0	4	0	0	4
12	0.25	0.75	0	14	4	0	6.5
13	1	0	0	7	0	0	7
14	0.25	0.75	0	14	7	0	8.75
15	0.25	0.75	0	14	7	0	8.75
(AE-C)	Lp1	Lp2	Lp3	La1	La2	La3	EV(L)
1	0.25	0.75	0	4	0	0	1
2	0.25	0.25	0.5	2	7	0	2.25
3	0.75	0.25	0	0	7	0	1.75
4	0.5	0.5	0	0	14	0	7
5	0.25	0.25	0.5	0	4	7	4.5
6	0.25	0.75	0	0	7	0	5.25
7	0.25	0.75	0	0	14	0	10.5
8	0.25	0.25	0.5	0	14	4	5.5
0	0.25	0.75	0	0	14	0	10.5
10	0.5	0.5	0	0	14	0	7

Table B.3: Right Prospect Parameters

(S-C)	Rp1	Rp2	Rp3	Ra1	Ra2	Ra3	Compound	EV(R)
1	0.5	0.5	0	0	2	0	if Ra2 (-2 or 2)	1
2	0.5	0.5	0	0	2	0	if Ra2 (-2 or 2)	1
3	0.5	0.5	0	1	3.5	0	if Ra1 (1 or -1)	2.25
4	0.5	0.5	0	0	3.5	0	if Ra2 (3.5 or -3.5)	1.75
5	1	0	0	7	0	0	if Ra1 (7 or -7)	7
6	1	0	0	7	0	0	if Ra1 (7 or -7)	7
7	0.5	0.5	0	2	7	0	if Ra1 (2 or -2)	4.5
8	0.5	0.5	0	3.5	7	0	if Ra1 (3.5 or -3.5)	5.25
9	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
10	0.5	0.5	0	7	4	0	if Ra1 (7 or -7)	5.5
11	0.5	0.5	0	7	4	0	if Ra1 (7 or -7)	5.5
12	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
13	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
14	1	0	0	7	0	0	if Ra1 (7 or -7)	7
15	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
(S-AE)	Rp1	Rp2	Rp3	Ra1	Ra2	Ra3	Compound	EV(R)
1	0.25	0.75	0	4	0	0	-	1
2	0.25	0.75	0	4	0	0	-	1
3	0.25	0.25	0.5	2	7	0	-	2.25
4	0.25	0.75	0	7	0	0	-	1.75
5	0.5	0.5	0	0	14	0	-	7
6	0.5	0.5	0	0	14	0	-	7
7	0.25	0.25	0.5	0	4	7	-	4.5
8	0.25	0.75	0	0	7	0	-	5.25
9	0.25	0.75	0	0	14	0	-	10.5
10	0.25	0.25	0.5	14	4	0	-	5.5
11	0.25	0.25	0.5	0	14	4	-	5.5
12	0.25	0.75	0	0	14	0	-	10.5
13	0.25	0.75	0	0	14	0	-	10.5
14	0.5	0.5	0	0	14	0	-	7
15	0.25	0.75	0	0	14	0	-	10.5
(AE-C)	Rp1	Rp2	Rp3	Ra1	Ra2	Ra3	Compound	EV(R)
1	0.5	0.5	0	0	2	0	if Ra2 (2 or -2)	1
2	0.5	0.5	0	1	3.5	0	if Ra2 (3.5 or -3.5)	2.25
3	0.5	0.5	0	0	3.5	0	if Ra1 (3.5 or -3.5)	1.75
4	1	0	0	7	0	0	if Ra2 (3.5 or -3.5)	7
5	0.5	0.5	0	2	7	0	if Ra1 (2 or -2)	4.5
6	0.5	0.5	0	7	3.5	0	if Ra2 (3.5 or -3.5)	5.25
7	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
8	0.5	0.5	0	7	4	0	if Ra1 (7 or -7)	5.5
0	0.5	0.5	0	7	14	0	if Ra1 (7 or -7)	10.5
10	1	0	0	7	0	0	if Ra1 (7 or -7)	7

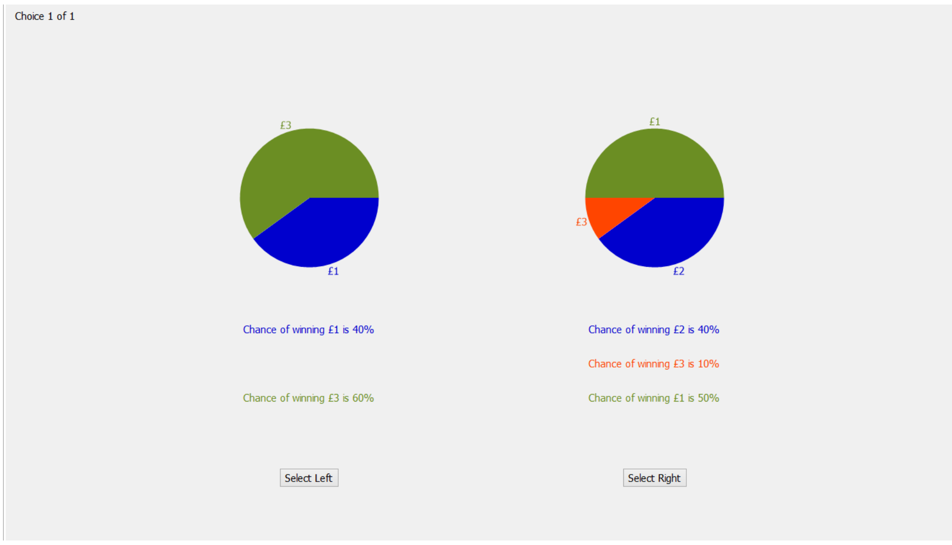
B.3 Instructions

Choices over Prospects

This is a task where you will choose between prospects with varying prizes and chances of winning. You will be presented with 40 pairs of prospects. For each pair of prospects, you should choose the prospect you prefer to play. You will actually get the chance to play one of the prospects you choose, and you will be paid according to the outcome of that prospect, so you should think carefully about which prospect you prefer.

Image B.1 is an example of what the computer display of such a pair of prospects might look like.

Figure B.1: Example of on-screen prospects



The outcome of the prospects will be determined by the draw of a random number between 1 and 100. Each number between, and including, 1 and 100 is equally likely to occur. In fact, you will be able to draw the number yourself using two 10-sided dice.

In the above example the left prospect pays one pound (£1) if the number drawn is between 1 and 40, and it pays three pounds (£3) if the number is between 41 and 100. The blue color in the pie chart corresponds to 40% of the area and illustrates the chances that the number drawn will be between 1 and 40 and your prize will be £1. The green area in the pie chart corresponds to 60% of the area and illustrates the chances that the number drawn will be between 41 and 100 and your prize will be £3.

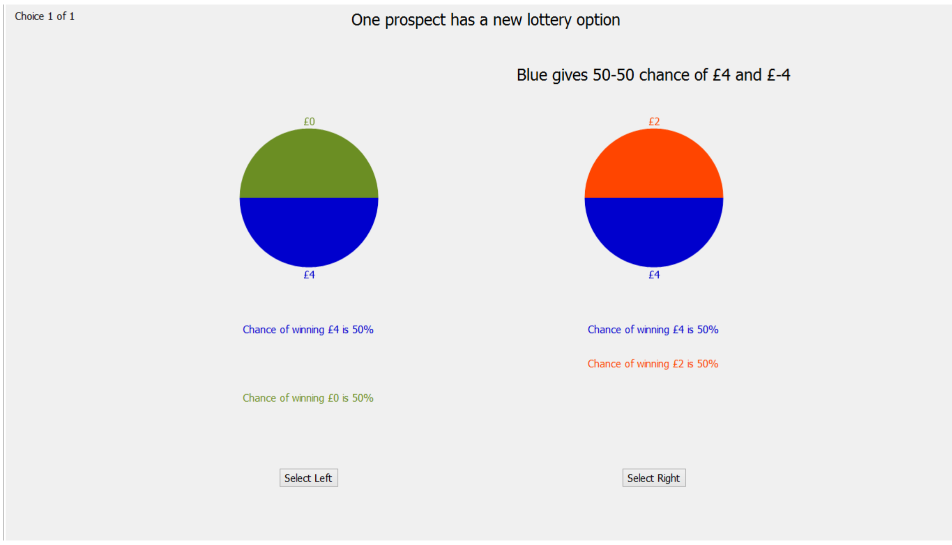
Now look at the pie in the chart on the right. It pays 1 pound (£1) if the number drawn is between 1 and 50; it pays two pounds (£2) if the number is between 51 and 90; and it pays three pounds (£3) if the number is between 91 and 100. As with the prospect on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the £3 pie slice is 10% of the total pie.

Each pair of prospects is shown on a separate screen on the computer. On each screen, you should indicate which prospect you prefer to play by clicking on one of the buttons beneath the prospects.

You could also get a pair of prospects in which one of the prospects will give you a new lottery option. For instance, the right prospect in the following screen image (Image

B.2) pays a new lottery if the Blue area is selected, which happens if the number drawn is between 51 and 100. The right pie chart indicates that if the number is between 1 and 50 you get £2. If the number is between 51 and 100 you get £4 and a new lottery which gives you £ – 4 if Heads comes up and £4if Tails comes up. The prizes listed underneath each pie refer to the amounts before any coin toss.

Figure B.2: Example of on-screen prospects



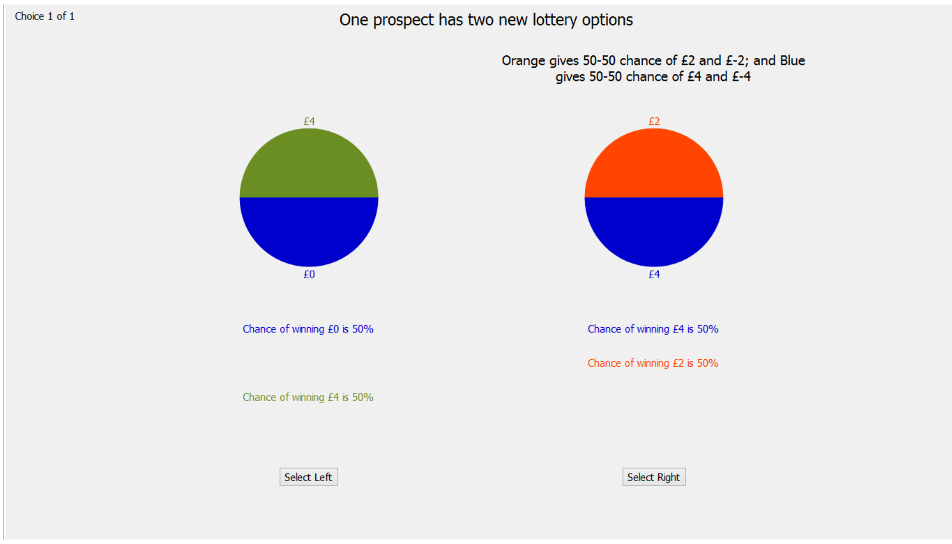
After you have worked through all of the 40 pairs of prospects, raise your hand and an experimenter will come over. You will then roll two 10-sided dice until a number between 1 and 40 comes up to determine which pair of prospects will be played out. Since there is a chance that any of your 40 choices could be played out for real, you should approach each pair of prospects as if it is the one that you will play out. Finally, you will roll the two ten-sided dice to determine the outcome of the prospect you chose, and if necessary you will then toss a coin to determine the outcome of the new lottery option.

For instance, suppose you picked the prospect on the left in the last example. If the random number was 37, you would win £0; if it was 93, you would get £4.

If you picked the prospect on the right and drew the number 37, you would get £2; if it was 93, you would have to toss a coin to determine the outcome of the new lottery option. If the coin comes up Heads then you get £4, and if it comes up Tails you get £ – 4.

It is also possible that you will be given a prospect in which there is a new lottery option no matter what the outcome of the random number. The screen B.3 illustrates this possibility.

Figure B.3: Example of on-screen prospects



Therefore, your payoff is determined by four things:

- by which prospect you selected, the left or the right, for each of these 40 pairs;
- by which prospect pair is chosen to be played out in the series of 40 such pairs using the two 10-sided dice;
- by the outcome of that prospect when you roll the two 10-sided dice;
- by the outcome of a coin toss if the chosen prospect outcome is of the new option type.

  Which prospects you prefer is a matter of personal taste. The people next to you may be presented with different prospects, and may have different preferences, so their responses should not matter to you. Please work silently, and make your choices by thinking carefully about each prospect. All payoffs are in cash, and are in addition to the £5 show-up fee that you receive just for being here.

Beliefs about Choices over Prospects

A group of 10 students at Durham University have previously participated in this experiment. Of these 10 participants, 5 are men and 5 are women.

This is a task where you will indicate your beliefs about the choices made by 1 of these 10 previous participants. First, you will randomly select 1 of the 10 previous participants. The previous participants chose between prospects with varying prizes and chances of winning. You will be presented with 40 pairs of prospects. For each pair of prospects, you should choose the prospect that you think was chosen by the randomly selected previous participant. You will actually be paid based on the accuracy of your belief about the randomly selected previous participant's choice for one of the pairs of prospects, so you should think carefully about which prospect you think was chosen.

The exact instructions that were given to the 10 previous participants have been provided to you. Please read them carefully before making your predictions.

Your payoff will be determined by the following four things:

- by which prospect you believe was chosen, the left or the right, for each of the 40 pairs;
- by which prospect pair is selected to be used for payment in the series of 40 such pairs using the two 10-sided dice;
- by which of the 10 previous participants is randomly selected;
- by which prospect the previous participant chose in the prospect pair that is selected for payment.

You will be paid £5 if your belief is correct in the randomly selected prospect pair. If your belief is incorrect you will be paid £0.

When you randomly select 1 of the 10 previous participants, you will be given an envelope that contains a list of the 40 choices by the person. On the outside of the envelope, you will see an ID number for the previous participant, labelled 1 to 10. Odd numbers indicate that the person is a male, and even numbers indicate that the person is a female. Inside the envelope is a printed table of the choices between prospects made by the randomly selected previous participant. Once you have indicated your beliefs for all 40 pairs, an experimenter will open the envelope so that the randomly selected previous participant's choices can be verified.

Which prospects your think were chosen by the randomly selected previous participant is a matter of your personal beliefs. The people next to you may be presented with different prospects and decisions by different previous participants, so their beliefs should not matter to you. Please work silently and think carefully about each prospect before indicating your belief. All payoffs are in cash, and are in addition to the £5show-up fee that you receive just for being here.

Demographic Survey

In this survey most of the questions asked are descriptive. We will not be grading your answers and your responses are completely confidential. Please think carefully about each question and give your best answers.

1. What is your age?

2. What is your sex?

    - Male
    - Female

3. Which of the following categories best describes you?

    - White
    - Mixed
    - Asian or Asian British
    - Chinese or other ethnic group
    - Prefer not to say

4. What is your main field of study?

    - Accounting
    - Economics
    - Finance
    - Business Administration, other than Accounting, Economics, or Finance
    - Education
    - Engineering
    - Health and Medicine
    - Biological and Biomedical Sciences
    - Math, Computer Sciences, or Physical Sciences
    - Social Sciences or History
    - Law

- Psychology
- Modern Languages and Cultures
- Other Fields

5. What is your year of studies?

- First year
- Second year
- Third year
- Masters
- Doctoral

6. What is the highest level of education you expect to complete?

- Bachelor's degree
- Master's degree
- Doctoral degree
- Professional qualification

7. As a percentage, what is your current average mark if you are doing a Bachelor's degree, or what was it when you did a Bachelor's degree? This mark should refer to all your years of study for this degree, not just the current year. Please pick one by rounding up or down to the nearest number:

- Above 70%
- Between 60 and 69%
- Between 50 and 59%
- Between 40 and 49%
- Less than 40%
- Have not taken courses for which grades are given.

8. What is your citizenship status?

- British Citizen
- EU Citizen (non-British Citizen)

- Non-EU Citizen

9. Are you currently:

   - Single and never married?
   - Married?
   - Separated, divorced or widowed?

10. How many people live in your household? Include yourself, your spouse and any dependents. Do not include your parents or roommates unless you claim them as dependents.

11. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2014 by the people in your household (as ‚Äúhouse-hold‚Äù is defined in question 10). [Consider all forms of income, including salaries, tips, interest and dividend payments, scholarship support, student loans, parental support, social security, alimony, and child support, and others.]

   - Less than £10,000
   - £10,000 and £19,999
   - £20,000 and £29,999
   - £30,000 and £49,999
   - Over £50,000

12. Please circle the category below that describes the total amount of income before tax earned in the calendar year 2014 by your parents. [Consider all forms of income, including salaries, tips, interest and dividend payments, social security, alimony, and child support, and others.]

   - Less than £10,000
   - £10,000 and £19,999
   - £20,000 and £29,999
   - £30,000 and £49,999
   - Over £50,000
   - Don't Know

13. Do you currently smoke cigarettes? (Circle one number.)

    - No
    - Yes
    - If yes, approximately how much do you smoke in one day? ... cigarettes.

APPENDIX C

Appendix C

C.1 Instructions

Participants were given written and on-screen instructions. Before starting the task, they were instructed to try to roll the die a couple of times. Differently from Weisel and Shalvi (2015), the written instructions were not only general, but they also included the treatment-specific instructions presented on the screen.

Printed Instructions

Welcome and thank you for participating in this experiment! You have received DKK 50 for having shown up on time. This amount will not be affected by the experiment results. Please, read carefully the printed and the on-screen instructions. The instructions are the same for all participants. Before starting each phase, there will be some control questions to verify that you understood the instructions. Should you have any questions about the instructions, raise your hand and a lab assistant will come to your place. During the experiment, your earnings are calculated in points. At the end of the experiment, the total amount of points you have earned will be converted into DKK at the following exchange rate:

$$1 point = 1 DKK$$

The decisions made by you and by the other participants are anonymous. Anonymity is maintained throughout and after the experiment. At the end of the experiment, you will be privately paid in another room. The experiment will consist of two parts and two questionnaires.

First phase In the first phase, you will be asked to roll the die in the cup and to report the outcome on the computer. To roll, hold the cup in your hand, shake a few times, and place the cup on your desk. You will look at the result and you will report it on the computer. After reporting, you will place the die inside the cup to avoid losing it. Roll the die a few times now to be sure you have understood the procedure.

On-screen instructions

In this phase, your task will be to roll the 6-sided die and report the outcome.

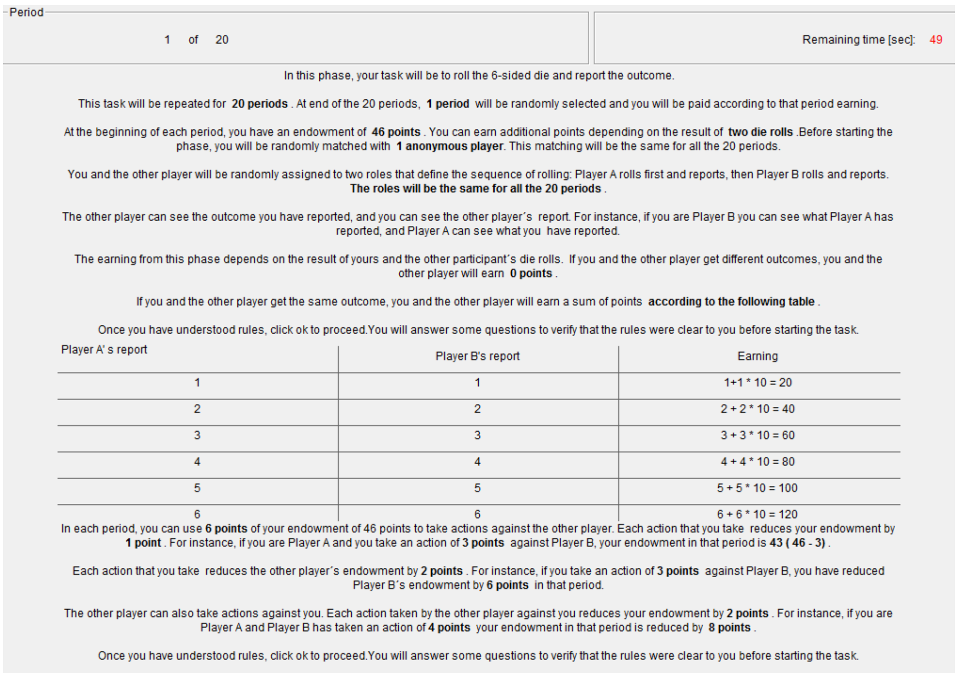
This task will be repeated for 20 periods. At end of the 20 periods, 1 period will be randomly selected and you will be paid according to that period earning. At the beginning of each period, you have an endowment of 46 points. You can earn additional points depending on the result of two die rolls. Before starting the phase, you will be randomly matched with 1 anonymous player. This matching will be the same for all the 20 periods. You and the other player will be randomly assigned to two roles that define the sequence of rolling: Player A rolls first and reports, then Player B rolls and reports. The roles will be the same for all the 20 periods. The other player can see the outcome you have reported, and you can see the other player's report. For instance, if you are Player B you can see what Player A has reported, and Player A can see what you have reported. The earning from this phase depends on the result of yours and the other participant's die rolls. If you and the other player get different outcomes, you and the other player will earn 0 points. If you and the other player get the same outcome, you and the other player will earn a sum of points according to the following table.

Table C.1: Earnings

Player A's outcome	Player B's outcome	Outcome Earning
1	1	$1+1\cdot 10=20$
2	2	$2+2\cdot 10=40$
3	3	$3+3\cdot 10=60$
4	4	$4+4\cdot 10=80$
5	5	$5+5\cdot 10=100$
6	6	$6+6\cdot 10=120$

In each period, you can use 6 points of your endowment of 46 points to take actions against the other player. Each action that you take reduces your endowment by 1 point. For instance, if you are Player A and you take an action of 3 points against Player B, your endowment in that period is $43(46-3)$. Each action that you take reduces the other player's endowment by 2 points. For instance, if you take an action of 3 points against Player B, you have reduced Player B's endowment by 6 points in that period. The other player can also take actions against you. Each action taken by the other player against you reduces your endowment by 2 points. For instance, if you are Player A and Player B has taken an action of 4 points your endowment in that period is reduced by 8 points.

Figure C.1: Example on-screen Instructions



Period

1 of 20

Remaining time [sec]: 49

In this phase, your task will be to roll the 6-sided die and report the outcome.

This task will be repeated for **20 periods** . At end of the 20 periods, **1 period** will be randomly selected and you will be paid according to that period earning.

At the beginning of each period, you have an endowment of **46 points** . You can earn additional points depending on the result of **two die rolls** .Before starting the phase, you will be randomly matched with **1 anonymous player**. This matching will be the same for all the 20 periods.

You and the other player will be randomly assigned to two roles that define the sequence of rolling: Player A rolls first and reports, then Player B rolls and reports. **The roles will be the same for all the 20 periods** .

The other player can see the outcome you have reported, and you can see the other player´s report. For instance, if you are Player B you can see what Player A has reported, and Player A can see what you have reported.

The earning from this phase depends on the result of yours and the other participant´s die rolls. If you and the other player get different outcomes, you and the other player will earn **0 points** .

If you and the other player get the same outcome, you and the other player will earn a sum of points **according to the following table** .

Once you have understood rules, click ok to proceed.You will answer some questions to verify that the rules were clear to you before starting the task.

Player A' s report	Player B's report	Earning
1	1	1+1 * 10 = 20
2	2	2 + 2 * 10 = 40
3	3	3 + 3 * 10 = 60
4	4	4 + 4 * 10 = 80
5	5	5 + 5 * 10 = 100
6	6	6 + 6 * 10 = 120

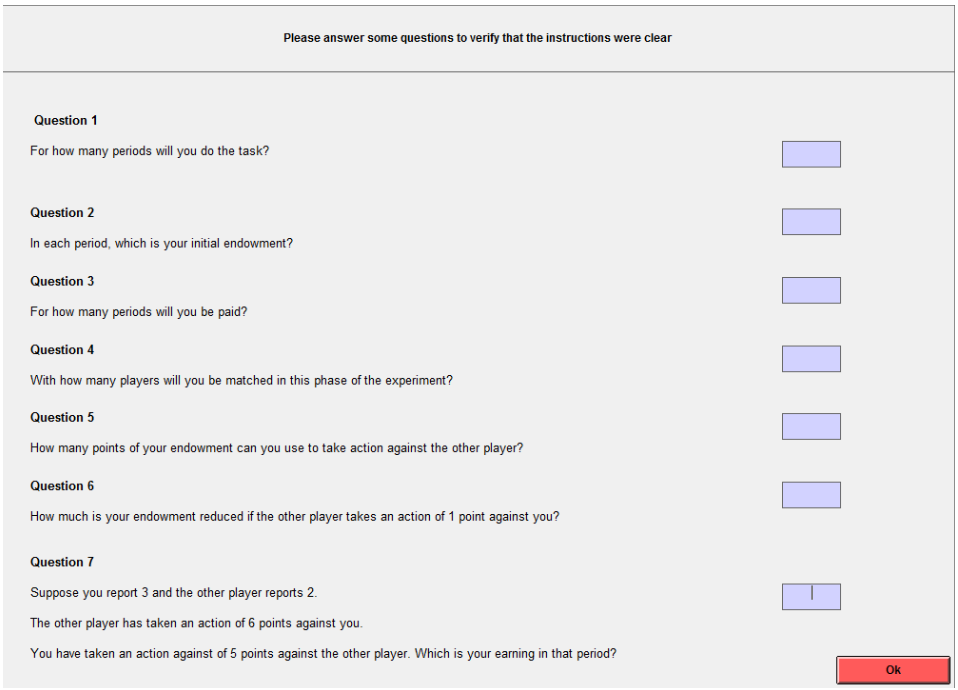
In each period, you can use **6 points** of your endowment of 46 points to take actions against the other player. Each action that you take reduces your endowment by **1 point** . For instance, if you are Player A and you take an action of **3 points** against Player B, your endowment in that period is **43 (46 - 3)** .

Each action that you take reduces the other player´s endowment by **2 points** . For instance, if you take an action of **3 points** against Player B, you have reduced Player B´s endowment by **6 points** in that period.

The other player can also take actions against you. Each action taken by the other player against you reduces your endowment by **2 points** . For instance, if you are Player A and Player B has taken an action of **4 points** your endowment in that period is reduced by **8 points** .

Once you have understood rules, click ok to proceed.You will answer some questions to verify that the rules were clear to you before starting the task.

Figure C.2: Control questions



C.1.1 Decision screens

Figure C.3: Player A enters the report at Period 4

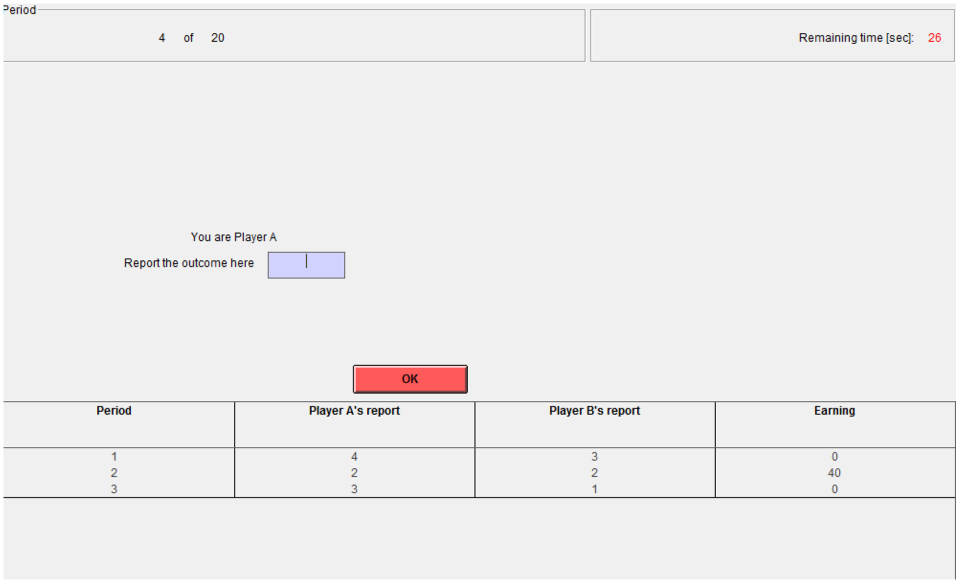


Figure C.4: After Player B has made the decision, Player A has the option to punish Player B

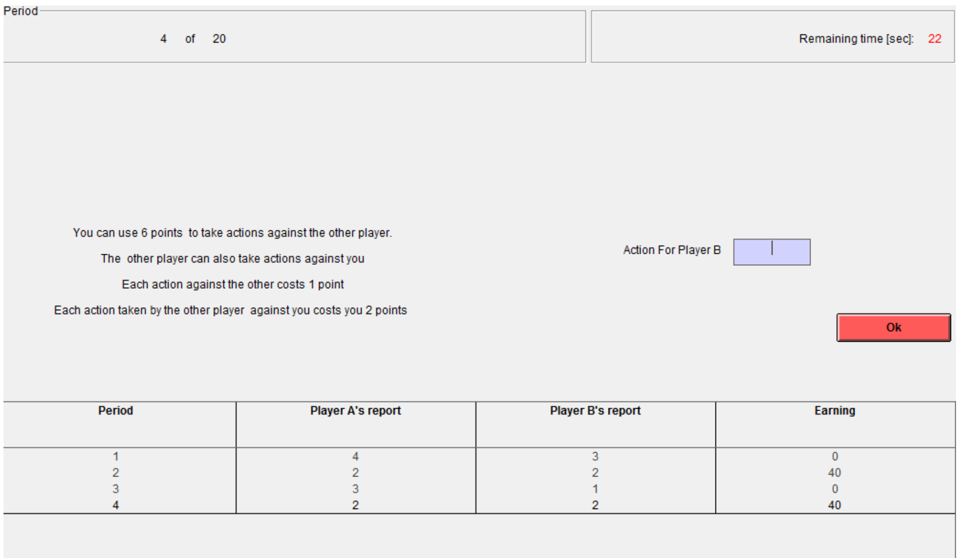
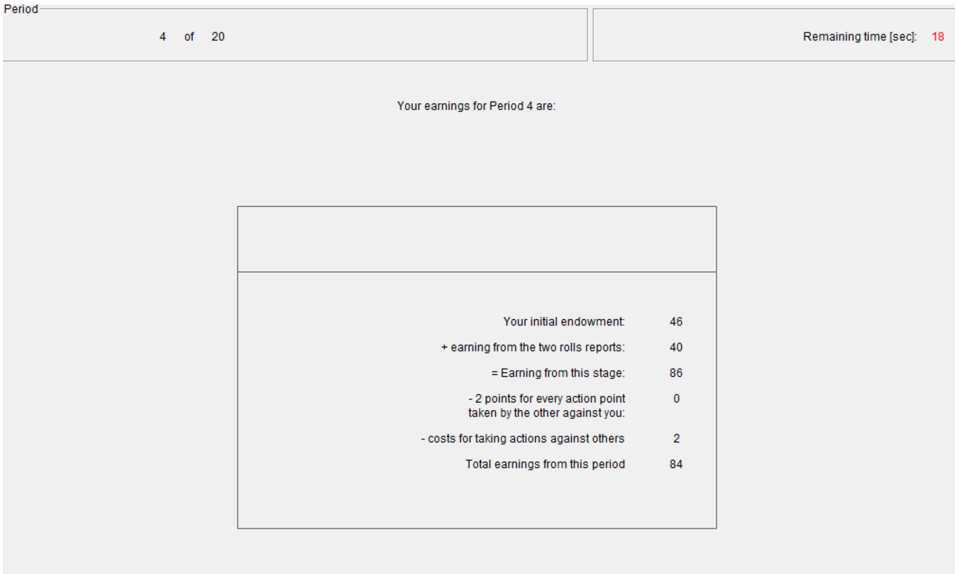


Figure C.5: After Player B has decided to punish Player A, Player A is shown the final earning from Period 4



Bibliography

Abeler, Johannes, Becker, Anke, & Falk, Armin. 2014.
Representative evidence on lying costs.
Journal of Public Economics, **113**, 96–104.

Abeler, Johannes, Nosenzo, Daniele, & Raymond, Collin. 2016.
Preferences for truth-telling.

Adams, J Stacy. 1965.
Inequity in social exchange.
Advances in experimental social psychology, **2**(267-299).

Akerlof, George A, & Kranton, Rachel E. 2000.
Economics and identity.
Quarterly journal of Economics, 715–753.

Akerlof, George A, & Kranton, Rachel E. 2005.
Identity and the Economics of Organizations.
Journal of Economic perspectives, 9–32.

Akerlof, George A, & Kranton, Rachel E. 2008.
Identity, supervision, and work groups.
The American Economic Review, 212–217.

Alempaki, Despoina, Doğan, Gönül, & Saccardo, Silvia. 2016.
Deception and Reciprocity.

Ambrose, Maureen L, Seabright, Mark A, & Schminke, Marshall. 2002.
Sabotage in the workplace: The role of organizational injustice.
Organizational Behavior and Human Decision Processes, **89**(1), 947–965.

Ames, Daniel R. 2004.

Strategies for social inference: a similarity contingency model of projection and stereotyping in attribute prevalence estimates.
Journal of personality and social psychology, **87**(5), 573.

Andersen, Steffen, Harrison, Glenn W, Lau, Morten Igel, & Rutström, E Elisabet. 2006.
Elicitation using multiple price list formats.
Experimental Economics, **9**(4), 383–405.

Ashton, Michael C, & Lee, Kibeom. 2007.
Empirical, theoretical, and practical advantages of the HEXACO model of personality structure.
Personality and Social Psychology Review, **11**(2), 150–166.

Ashton, Michael C, & Lee, Kibeom. 2009.
The HEXACO–60: A short measure of the major dimensions of personality.
Journal of personality assessment, **91**(4), 340–345.

Ball, Sheryl, Eckel, Catherine C, & Heracleous, Maria. 2010.
Risk aversion and physical prowess: Prediction, choice and bias.
Journal of Risk and Uncertainty, **41**(3), 167–193.

Balliet, Daniel, Wu, Junhui, & De Dreu, Carsten KW. 2014.
Ingroup favoritism in cooperation: A meta-analysis.
American Psychological Association.

Barr, Abigail, Miller, Luis, & Ubeda, Paloma. 2016.
Moral consequences of becoming unemployed.
Proceedings of the National Academy of Sciences, **113**(17), 4676–4681.

Baumgartner, Thomas, Schiller, Bastian, Rieskamp, Jörg, Gianotti, Lorena RR, & Knoch, Daria. 2013.
Diminishing parochialism in intergroup conflict by disrupting the right temporo-parietal junction.
Social cognitive and affective neuroscience, nst023.

Becker, Gary S. 1968.
Crime and punishment: An economic approach.
Journal of Political economy, **76**(2), 169–217.

Becker, Gordon M, DeGroot, Morris H, & Marschak, Jacob. 1964.
Measuring utility by a single-response sequential method.
Behavioral science, **9**(3), 226–232.

Bénabou, Roland, & Tirole, Jean. 2011.
Identity, morals, and taboos: Beliefs as assets.
The Quarterly Journal of Economics, **126**(2), 805–855.

Binswanger, Hans P. 1980.
Attitudes toward risk: Experimental measurement in rural India.
American journal of agricultural economics, **62**(3), 395–407.

Binswanger, Hans P. 1981.
Attitudes toward risk: Theoretical implications of an experiment in rural India.
The Economic Journal, **91**(364), 867–890.

Bolton, Gary E, & Ockenfels, Axel. 2000.
ERC: A theory of equity, reciprocity, and competition.
American economic review, 166–193.

Bowles, Samuel. 1998.
Endogenous preferences: The cultural consequences of markets and other economic institutions.
Journal of economic literature, **36**(1), 75–111.

Bucciol, Alessandro, & Piovesan, Marco. 2011.
Luck or cheating? A field experiment on honesty with children.
Journal of Economic Psychology, **32**(1), 73–78.

Bucciol, Alessandro, Landini, Fabio, & Piovesan, Marco. 2013.
Unethical behavior in the field: Demographic characteristics and beliefs of the cheater.
Journal of Economic Behavior & Organization, **93**, 248–257.

Cadsby, C Bram, Song, Fei, & Tapon, Francis. 2010.
Are you paying your employees to cheat? An experimental investigation.
The BE Journal of Economic Analysis & Policy, **10**(1).

Cappelen, Alexander W, Eichele, Tom, Hugdahl, Kenneth, Specht, Karsten, Sørensen, Erik Ø, & Tungodden, Bertil. 2014.
Equity theory and fair inequality: A neuroeconomic study.

Proceedings of the National Academy of Sciences, **111**(43), 15368–15372.

Chakravarty, Sujoy, Harrison, Glenn W, Haruvy, Ernan E, & Rutström, E Elisabet. 2011. Are you risk averse over other people's money? *Southern Economic Journal*, **77**(4), 901–913.

Charness, Gary, & Dufwenberg, Martin. 2006. Promises and partnership. *Econometrica*, **74**(6), 1579–1601.

Charness, Gary, & Jackson, Matthew O. 2007. Group play in games and the role of consent in network formation. *Journal of Economic Theory*, **136**(1), 417–445.

Chen, Roy, & Chen, Yan. 2011. The potential of social identity for equilibrium selection. *The American Economic Review*, 2562–2589.

Chen, Yan, & Li, Sherry Xin. 2009. Group identity and social preferences. *The American Economic Review*, 431–457.

Chen, Yan, Li, Sherry Xin, Liu, Tracy Xiao, & Shih, Margaret. 2014. Which hat to wear? Impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, **84**, 58–86.

Cherry, Todd L, Frykblom, Peter, & Shogren, Jason F. 2002. Hardnose the dictator. *American Economic Review*, 1218–1221.

Cohn, Alain, Fehr, Ernst, & Maréchal, Michel André. 2014. Business culture and dishonesty in the banking industry. *Nature*.

Conrads, Julian, Irlenbusch, Bernd, Rilke, Rainer Michael, & Walkowitz, Gari. 2013. Lying and team incentives. *Journal of Economic Psychology*, **34**, 1–7.

Conte, Anna, Levati, M Vittoria, & Nardi, Chiara. 2016. Risk preferences and the role of emotions. *Economica*.

Crosetto, Paolo, & Filippin, Antonio. 2013.
The ‚Äúbomb‚Äù risk elicitation task.
Journal of Risk and Uncertainty, **47**(1), 31–65.

Croson, Rachel, & Gneezy, Uri. 2009.
Gender differences in preferences.
Journal of Economic literature, 448–474.

Daruvala, Dinky. 2007.
Gender, risk and stereotypes.
Journal of Risk and Uncertainty, **35**(3), 265–283.

Dengler-Roscher, Kathrin, Montinari, Natalia, Panganiban, Marian, Ploner, Matteo, Werner, Benedikt, *et al.* 2015.
On the Malleability of Fairness Ideals: Order Effects in Partial and Impartial Allocation Tasks.
Tech. rept. Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

DiMaggio, Paul. 1997.
Culture and cognition.
Annual review of sociology, 263–287.

Dreber, Anna, & Johannesson, Magnus. 2008.
Gender differences in deception.
Economics Letters, **99**(1), 197–199.

Eckel, Catherine C, & Grossman, Philip J. 2002.
Sex differences and statistical stereotyping in attitudes toward financial risk.
Evolution and human behavior, **23**(4), 281–295.

Eckel, Catherine C, & Grossman, Philip J. 2005.
Managing diversity by creating team identity.
Journal of Economic Behavior & Organization, **58**(3), 371–392.

Eckel, Catherine C, & Grossman, Philip J. 2008.
Forecasting risk attitudes: An experimental study using actual and forecast gamble choices.
Journal of Economic Behavior & Organization, **68**(1), 1–17.

Ellingsen, Tore, & Johannesson, Magnus. 2004.
Promises, threats and fairness.
The Economic Journal, **114**(495), 397–420.

Ellingsen, Tore, Johannesson, Magnus, Lilja, Jannie, & Zetterqvist, Henrik. 2009.
Trust and Truth*.
The Economic Journal, **119**(534), 252–276.

Engelmann, Dirk, & Strobel, Martin. 2004.
Inequality aversion, efficiency, and maximin preferences in simple distribution experiments.
American economic review, 857–869.

Erat, Sanjiv, & Gneezy, Uri. 2012.
White lies.
Management Science, **58**(4), 723–733.

Falk, Armin, & Tirole, Jean. 2016.
Narratives, Imperatives and Moral Reasoning1.

Fehr, Ernst, & Fischbacher, Urs. 2004.
Third-party punishment and social norms.
Evolution and human behavior, **25**(2), 63–87.

Fehr-Duda, Helga, De Gennaro, Manuele, & Schubert, Renate. 2006.
Gender, financial risk, and probability weights.
Theory and decision, **60**(2-3), 283–313.

Feri, Francesco, Irlenbusch, Bernd, & Sutter, Matthias. 2009.
Efficiency gains from team-based coordination–Large-scale experimental evidence.

Fischbacher, Urs. 2007.
z-Tree: Zurich toolbox for ready-made economic experiments.
Experimental economics, **10**(2), 171–178.

Fischbacher, Urs, & Föllmi-Heusi, Franziska. 2013.
Lies in disguise,Äîan experimental study on cheating.
Journal of the European Economic Association, **11**(3), 525–547.

Gächter, Simon, & Schulz, Jonathan F. 2016.
Intrinsic honesty and the prevalence of rule violations across societies.
Nature.

Gill, David, & Prowse, Victoria. 2012.
A structural analysis of disappointment aversion in a real effort competition.
The American Economic Review, **102**(1), 469–503.

Gill, David, Prowse, Victoria, & Vlassopoulos, Michael. 2013.
Cheating in the workplace: An experimental study of the impact of bonuses and productivity.
Journal of Economic Behavior & Organization, **96**, 120–134.

Gino, Francesca, & Pierce, Lamar. 2009.
Dishonesty in the name of equity.
Psychological Science, **20**(9), 1153–1160.

Gino, Francesca, & Pierce, Lamar. 2010a.
Lying to level the playing field: Why people may dishonestly help or hurt others to create equity.
Journal of Business Ethics, **95**(1), 89–103.

Gino, Francesca, & Pierce, Lamar. 2010b.
Robin Hood under the hood: Wealth-based discrimination in illicit customer help.
Organization Science, **21**(6), 1176–1194.

Gino, Francesca, Ayal, Shahar, & Ariely, Dan. 2013.
Self-serving altruism? The lure of unethical actions that benefit others.
Journal of economic behavior & organization, **93**, 285–292.

Glätzle-Rützler, Daniela, & Lergetporer, Philipp. 2015.
Lying and age: An experimental study.
Journal of Economic Psychology, **46**, 12–25.

Gneezy, Uri. 2005.
Deception: The role of consequences.
American Economic Review, 384–394.

Goette, L., Huffman D. Meier S. 2006.

The impact of group membership on cooperation and norm enforcement: evidence using random assignment to real social groups.
American Economic Review, **2**(96), 212–6.

Goffman, Erving. 1963.
Stigma: Notes on the management of spoiled identity.
Simon and Schuster.

Gonzalez, Richard, & Wu, George. 1999.
On the shape of the probability weighting function.
Cognitive psychology, **38**(1), 129–166.

Greenberg, Jerald. 1990.
Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts.
Journal of applied psychology, **75**(5), 561.

Greiner, Ben. 2004.
An online recruitment system for economic experiments.

Greiner, Ben. 2015.
Subject pool recruitment procedures: organizing experiments with ORSEE.
Journal of the Economic Science Association, **1**(1), 114–125.

Grolleau, Gilles, Kocher, Martin G, & Sutan, Angela. 2016.
Cheating and Loss Aversion: Do People Cheat More to Avoid a Loss?
Management Science.

Grossman, Philip J. 2013.
Holding fast: the persistence and dominance of gender stereotypes.
Economic Inquiry, **51**(1), 747–763.

Harrison, Glenn W, & Rutstrom, E Elisabet. 2008.
Risk aversion in the laboratory.

Harrison, Glenn W, & Swarthout, J Todd. 2012.
The independence axiom and the bipolar behaviorist.
Georgia State University CEEL working paper, **1**.

Harrison, Glenn W, & Swarthout, J Todd. 2014.
Experimental payment protocols and the bipolar behaviorist.
Theory and Decision, **77**(3), 423–438.

Harrison, Glenn W, Martínez-Correa, Jimmy, & Swarthout, J Todd. 2015.
Reduction of compound lotteries with objective probabilities: Theory and evidence.
Journal of Economic Behavior & Organization, **119**, 32–55.

Herrmann, Benedikt, Thöni, Christian, & Gächter, Simon. 2008.
Antisocial punishment across societies.
Science, **319**(5868), 1362–1367.

Hey, John D. 2001.
Does repetition improve consistency?
Experimental economics, **4**(1), 5–54.

Hey, John D, & Orme, Chris. 1994.
Investigating generalizations of expected utility theory using experimental data.
Econometrica: Journal of the Econometric Society, 1291–1326.

Hilbig, Benjamin E, & Zettler, Ingo. 2015.
When the cat's away, some mice will play: A basic trait account of dishonest behavior.
Journal of Research in Personality, **57**, 72–88.

Hilbig, Benjamin E, Zettler, Ingo, & Heydasch, Timo. 2012.
Personality, punishment and public goods: Strategic shifts towards cooperation as a matter of dispositional honesty–humility.
European Journal of Personality, **26**(3), 245–254.

Hoff, Karla, & Stiglitz, Joseph E. 2016.
Striving for balance in economics: Towards a theory of the social determination of behavior.
Journal of Economic Behavior & Organization, **126**, 25–57.

Hollinger, Richard C, & Clark, John P. 1983.
Deterrence in the workplace: Perceived certainty, perceived severity, and employee theft.
Social Forces, **62**(2), 398–418.

Holt, Charles A, & Laury, Susan K. 2002.
Risk aversion and incentive effects.
American economic review, **92**(5), 1644–1655.

Homans, George C. 1958.
Social behavior as exchange.
American journal of sociology, 597–606.

Houser, Daniel, Vetter, Stefan, & Winter, Joachim. 2012.
Fairness and cheating.
European Economic Review, **56**(8), 1645–1655.

Hsee, Christopher, & Weber, Elke U. 1998.
Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk.
Management Science, **44**(9), 1205–1217.

Hsee, Christopher K, & Weber, Elke U. 1997.
A fundamental prediction error: Self–others discrepancies in risk preference.
Journal of experimental psychology: general, **126**(1), 45.

Hsee, Christopher K, & Weber, Elke U. 1999.
Cross-national differences in risk preference and lay predictions.
Journal of Behavioral Decision Making, **12**(2), 165–179.

John, Leslie K, Loewenstein, George, & Rick, Scott I. 2014.
Cheating more for less: Upward social comparisons motivate the poorly compensated to cheat.
Organizational Behavior and Human Decision Processes, **123**(2), 101–109.

Kahneman, Daniel, & Tversky, Amos. 1979.
Prospect theory: An analysis of decision under risk.
Econometrica: Journal of the econometric society, 263–291.

Kocher, Martin G, & Sutter, Matthias. 2005.
The decision maker matters: Individual versus group behaviour in experimental beauty-contest games.
The Economic Journal, **115**(500), 200–223.

Kocher, Martin G, Schudy, Simeon, & Spantig, Lisa. 2016.
I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups.

Kollock, Peter. 1998.
Transforming social dilemmas: Group identity and cooperation.

Modeling rational and moral agents, 186–210.

Konow, James. 1996.
A positive theory of economic fairness.
Journal of Economic Behavior & Organization, **31**(1), 13–35.

Konow, James. 2000.
Fair shares: Accountability and cognitive dissonance in allocation decisions.
American Economic Review, 1072–1091.

Konow, James. 2003.
Which is the fairest one of all? A positive analysis of justice theories.
Journal of economic literature, 1188–1239.

Kramer, Roderick M, & Brewer, Marilynn B. 1984.
Effects of group identity on resource use in a simulated commons dilemma.
Journal of personality and social psychology, **46**(5), 1044.

Krupka, Erin L, & Weber, Roberto A. 2013.
Identifying social norms using coordination games: Why does dictator game sharing vary?
Journal of the European Economic Association, **11**(3), 495–524.

Levitt, Steven D, & List, John A. 2007.
Viewpoint: On the generalizability of lab behaviour to the field.
Canadian Journal of Economics/Revue canadienne d'économique, **40**(2), 347–370.

Marx, K. 1904.
A contribution to the critique of political economy.

Mazar, Nina, & Ariely, Dan. 2006.
Dishonesty in everyday life and its policy implications.
Journal of Public Policy & Marketing, **25**(1), 117–126.

Mazar, Nina, Amir, On, & Ariely, Dan. 2008.
The dishonesty of honest people: A theory of self-concept maintenance.
Journal of marketing research, **45**(6), 633–644.

McLeish, Kendra N, & Oxoby, Robert J. 2007.
Identity, cooperation, and punishment.
IZA Discussion Papers.

McLeish, Kendra N, & Oxoby, Robert J. 2011.
Social interactions and the salience of social identity.
Journal of Economic Psychology, **32**(1), 172–178.

Meier, Stephan, Pierce, Lamar, Vaccaro, Antonino, & Cara, Barbara La. 2016.
Trust and In-Group Favoritism in a Culture of Crime.
Journal of Economic Behavior and Organization, –.

Mittone, Luigi, & Ploner, Matteo. 2012.
Asset legitimacy and distributive justice in the dictator game: an experimental analysis.
Journal of Behavioral Decision Making, **25**(2), 135–142.

Monin, Benoît, & Jordan, Alexander H. 2009.
The dynamic moral self: A social psychological perspective.
Personality, identity, and character: Explorations in moral psychology, 341–354.

Nese, Annamaria, Palomba, Arturo, Sbriglia, Patrizia, Scudiero, Maurizio, *et al.* 2013.
Third party punishment and criminal behavior: An experiment with the Italian Camorra prison inmates.
Economics Bulletin, **33**(3), 1875–1884.

Nikiforakis, Nikos, Noussair, Charles N, & Wilkening, Tom. 2012.
Normative conflict and feuds: The limits of self-enforcement.
Journal of Public Economics, **96**(9), 797–807.

Pirandello, Luigi. 1926.
Uno, nessuno e centomila.

Ploner, Matteo, & Regner, Tobias. 2013.
Self-image and moral balancing: An experimental analysis.
Journal of Economic Behavior & Organization, **93**, 374–383.

Plott, Charles R, & Zeiler, Kathryn. 2005.
The Willingness to Pay–Willingness to Accept Gap, the.
The American Economic Review, **95**(3), 530–545.

Quiggin, John. 1982.
A theory of anticipated utility.
Journal of Economic Behavior & Organization, **3**(4), 323–343.

Rawls, John. 1971.
A theory of justice.
Oxford University Press.

Rong, Rong, Houser, Daniel, & Dai, Anovia Yifan. 2016.
Money or friends: Social identity and deception in networks.
European Economic Review.

Ross, Lee, Greene, David, & House, Pamela. 1977.
The ‚Äúfalse consensus effect‚Äù: An egocentric bias in social perception and attribution processes.
Journal of experimental social psychology, **13**(3), 279–301.

Roth, Benjamin, & Voskort, Andrea. 2014.
Stereotypes and false consensus: How financial professionals predict risk preferences.
Journal of Economic Behavior & Organization, **107**, 553–565.

Sani, Fabio, Bowe, Mhairi, Herrera, Marina, Manna, Cristian, Cossa, Tiziana, Miao, Xiulou, & Zhou, Yuefang. 2007.
Perceived collective continuity: Seeing groups as entities that move through time.
European Journal of Social Psychology, **37**(6), 1118–1134.

Schindler, Simon, & Pfattheicher, Stefan. 2016.
The frame of the game: Loss-framing increases dishonest behavior.
Journal of Experimental Social Psychology.

Shalvi, Shaul, Dana, Jason, Handgraaf, Michel JJ, & De Dreu, Carsten KW. 2011.
Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior.
Organizational Behavior and Human Decision Processes, **115**(2), 181–190.

Shalvi, Shaul, Gino, Francesca, Barkan, Rachel, & Ayal, Shahar. 2015.
Self-Serving Justifications Doing Wrong and Feeling Moral.
Current Directions in Psychological Science, **24**(2), 125–130.

Siegrist, Michael, Cvetkovich, George, & Gutscher, Heinz. 2002.
Risk preference predictions and gender stereotypes.
Organizational Behavior and Human Decision Processes, **87**(1), 91–102.

Sturge-Apple, Melissa L, Suor, Jennifer H, Davies, Patrick T, Cicchetti, Dante, Skibo, Michael A, & Rogosch, Fred A. 2016.
Vagal Tone and Children‚Äôs Delay of Gratification Differential Sensitivity in Resource-Poor and Resource-Rich Environments.
Psychological science, 0956797616640269.

Sutter, Matthias. 2009.
Deception through telling the truth?! Experimental evidence from individuals and teams.
The Economic Journal, **119**(534), 47–60.

Tajfel, Henri, Billig, Michael G, Bundy, Robert P, & Flament, Claude. 1971.
Social categorization and intergroup behaviour.
European journal of social psychology, **1**(2), 149–178.

Turner, John C. 1985.
Social categorization and the self-concept: A social cognitive theory of group behavior.
Advances in group processes, **2**, 77–122.

Tversky, Amos, & Kahneman, Daniel. 1975.
Judgment under uncertainty: Heuristics and biases.
Pages 141–162 of: Utility, probability, and human decision making.
Springer.

Tversky, Amos, & Kahneman, Daniel. 1992.
Advances in prospect theory: Cumulative representation of uncertainty.
Journal of Risk and uncertainty, **5**(4), 297–323.

Von Neumann, John, & Morgenstern, Oskar. 1944.
Theory of games and economic behavior.
Princeton university press.

Walster, Elaine, Berscheid, Ellen, & Walster, G William. 1973.
New directions in equity research.
Journal of personality and social psychology, **25**(2), 151.

Weisel, Ori, & Shalvi, Shaul. 2015.
The collaborative roots of corruption.
Proceedings of the National Academy of Sciences, **112**(34), 10651–10656.

Weng, Qian, & Carlsson, Fredrik. 2015.
Cooperation in teams: The role of identity, punishment, and endowment distribution.
Journal of Public Economics, **126**, 25–38.

Wiltermuth, Scott S. 2011.
Cheating more when the spoils are split.
Organizational Behavior and Human Decision Processes, **115**(2), 157–168.

Wyer Jr, Robert S, & Wyer Jr, Robert S. 1998.
Stereotype activation and inhibition: Advances in social cognition.
Psychology Press.

Xiao, Erte, & Kunreuther, Howard. 2016.
Punishment and cooperation in stochastic social dilemmas.
Journal of Conflict Resolution, **60**(4), 670–693.