Selfishness and Morality in Social and Organizational Behavior

Essays in Experimental Economics

Der Fakultät für Wirtschaftswissenschaften der
Universität Paderborn

zur Erlangung des akademischen Grades
Doktor der Wirtschaftswissenschaften
- Doctor rerum politicarum -

vorgelegte Dissertation
von

Nina Lucia Stephan, M.Sc.

geboren am 29.05.1990
in Hamburg

Januar 2019

Contents

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, 1976, p.3)

Humans are social beings who do not spend their life in isolation. Their decision-making is influenced by others, and their decisions impact more than just themselves. Every individual is exposed to a unique social environment, which is first made up by the people in the individual's family and private life, and later on in his school and professional life. Besides an individual's selfish survival instinct and personal innate characteristics, another part of every individual's preferences, perceptions, and values evolves in this social context that the individual is exposed to. An individual's personal goals are thus also changing together with the social environment, and are not stable over time.

Just as Nobel Laureate Gary Becker, I understand the economic approach as a powerful method of analysis (Becker, 1993), rather than as a mindset restricted by a fixed set of traditional assumptions about selfishness or material gain. Knowing that as social beings all individuals care about social problems, it is self-evident that making use of economic methods to study such issues is of interest to many. Moreover, puzzling results from empirical studies trigger improvements in economic theories, as prevailing models are challenged by new insights regarding what drives human decision-making. This dissertation adds empirically-based behavioral insights that may not only help advance economic models, but may also inform policy makers and organizations to better understand the society or their employees, respectively.

Whenever other people play a role in the decision-making, either because they are involved in deciding or because they are affected by the decision, I refer to these different circumstances as interpersonal factors. It is self-evident that interpersonal factors play a role when decisions may impact others. However, it is much less straightforward to explain how these factors impact choices, because additional factors may be at work at the same time. Such factors may include, amongst others, the selfish pursuit of maximizing own benefits (Gneezy, 2005), as well as caring for the well-being of others (Andreoni, 1995), or striving for a positive social image (Fehr and Falk, 2002) or self image (Akerlof and Kranton, 2000; Gino et al., 2011). This challenge of disentangling what drives certain choices has been taken up by experimental economists, who aim at separating alternative explanations by design. The aim of this dissertation is to approach a better comprehension of interpersonal

1

factors influencing choices in social and organizational behavior, making use of controlled experiments. Each of the four chapters contributes to the understanding of one specific decision situation, in which decisions have consequences not only for the decision maker, but also for others, and thus for society and the economy as a whole. In each chapter, I shed light on another particular issue related to the bigger question of how interpersonal factors shape decision-making.

To start off, chapter one is concerned with the effect of an interpersonal experience on the decision to engage in embezzlement. In a laboratory experiment, participants receive endowments of different heights, either within an interpersonal experience or without any other person playing a role, and the different effects on immoral behavior are measured. In the following two chapters, I take my understanding of interpersonal factors a step further, by additionally identifying the nature of the relationship between the decision maker and other people. For instance, in chapter two it is of interest to distinguish between emotionally positive and emotionally negative experiences with other people: We are concerned with negative interpersonal experiences that may be passed on to other, innocent people, which we refer to as the 'chain of unkindness'. Again, we use a laboratory experiment to review a remedy that promises to stop chains of unkindness: writing a letter to the person who first started the chain. The third chapter continues to investigate this potential remedy. It examines the drivers of the favorable effect of letter writing even closer, by also taking into account how it influences the passing on of kind behavior. In doing so, the chapter aims at revealing the general pro-social effect of letter writing, rather than its curative effect. In the next and final chapter of this dissertation, it becomes relevant to distinguish whether a decision maker is related to another person. The fourth chapter takes a different angle, as it tackles the origin of interpersonal factors starting to play a role in decision-making, rather than investigating a specific decision situation that may also be encountered in private or professional everyday life: This last chapter explores whether interpersonal factors already influence decision-making early in life, namely among children. Here, a lab-in-the-field experiment is used to investigate children's decisions in a more natural environment. The central question in this study is whether the moral decision to cheat changes when cheating does not benefit the child itself, but rather benefits another child instead.

These four chapters are related not only because they take into account the role of some interpersonal experience — they are also concerned with the immoral nature of popular choices that can harm both social and professional life and thus adversely affect employers or the economy as a whole. To give some examples of the economic consequences of immoral behavior and their dimensions, we can refer to numerous empirical studies not only in economics but also in related fields such as law or criminology. For instance, according to Cialdini et al. (2004), about 400 billion US dollars are lost due to fraud within US companies. Bressler (2009) estimates

that around 30% of all small business failures in 1995 could be traced back to employee dishonesty. Dreher and Herzfeld (2005) review literature on white-collar crimes, focusing on corruption, and estimate that a 1-point increase in the corruption index significantly reduces GDP per capita by around 425 US dollars. Fan et al. (2010) focus on embezzlement, another form of white-collar crime, and estimate that embezzlement by Chinese officials amounted to a cost of 143 billion US dollars in 2006. Understanding what drives the decision to embezzle is also the topic of the first chapter of this dissertation. In the following, I provide a more detailed summary of each of the four chapters of my dissertation.

Chapter 1: "Crumbling morals? An experimental study on how a social framing can affect embezzlement"

In this first chapter of my dissertation I am concerned with the effect of a social framing on the decision to engage in embezzlement. With data from a laboratory experiment, I aim at answering the question whether an interpersonal experience can either encourage or reduce subsequent immoral decision-making; in this case the immoral decision to engage in embezzlement.

The experiment consists of two parts. In the first part, subjects are not making any decisions yet, but they receive initial endowments, either as a share from a previous dictator (social framing treatment) or by random determination (neutral framing treatment). In the second part of the experiment, subjects can increase their payoff by embezzling, i.e. taking away part of a donation that they are entrusted with, before forwarding the rest to Médecins Sans Frontières.

I find that after receiving a less-than-half share in the social framing treatment, in comparison to receiving an equally large amount in the neutral framing treatment, subjects are significantly more likely to embezzle. Thus, depending on the height of endowment, the social framing encourages forfeiting moral behavior. I argue that this effect is driven by self-deception: observing the dictator's decision to share less than half facilitates excusing one's own immoral choice.

From this study, I conclude that socially framing a moral decision situation, even though this may remind individuals of what is socially acceptable, does not have an unanimously beneficial effect on the moral decision to embezzle.

Chapter 2: "Revisiting a remedy against the chain of unkindness" - with Wendelin Schnedler

The second chapter of this dissertation is co-authored with Wendelin Schnedler. In our study, we revisit a remedy against the passing on of unkindness to an innocent third person, starting a chain of negative downward reciprocity. Again, we make use of a laboratory experiment to investigate whether such a negative interpersonal experience is successfully stopped from being passed on.

Previous experimental studies show that subjects who receive little in a dictator game, pass on less to a third person when they become dictators themselves. However, Strang et al. (2016) find that when subjects can write a letter to their dictator, they act more kindly. The prevailing explanation for this phenomenon is that letter writing helps them to emotionally 'close the case' (closure explanation). In our study, we suggest an alternative explanation for the observed effect: being asked to write a letter may be seen as a signal that something is wrong with the dictator's behavior and lead to less imitation of this behavior (signal explanation).

In our experiment we examine whether letter writing also helps to decrease unkindness when the second explanation is excluded by design. More specifically, subjects face choices in different decision domains, so that imitation is not possible. Instead of deciding how to split money, the participants' first decision consists of assigning an annoying work task (the unkind option) instead of an enjoyable task to another person.

We find that letter writing nevertheless increases giving in a subsequent dictator game. Our results confirm the effectiveness of letter writing as a remedy against the chain of unkindness. From the results we conclude that the closure explanation is perhaps more suitable. There is, however, one caveat: while writing the letter may make people emotionally 'close the case', this is not reflected in how happy people rate themselves. Thus, it may be that closure does not describe the channel through which letter writing works best. This open question is followed up on in the next chapter of this dissertation.

Chapter 3: "When letter writing increases kindness: Regulating emotions or activating pro-social thinking?" - with Wendelin Schnedler

The third chapter of this dissertation builds directly on the second, and is also co-authored with Wendelin Schnedler. Here, we dig deeper into the effect of the remedy against the chain of unkindness (letter writing). We aim at answering the question whether the effect of letter writing is actually curative, or whether it has a general

socially desired effect instead. For this purpose, we take additional data into account, namely data from participants who experience kindness, instead of unkindness. We make use of data collected in the same experiment as the data collected for chapter two of this dissertation.

We re-consider the effect of letter writing as described in chapter two, and describe two different channels that may drive this effect. First, we recall from chapter two that letter writing may help writers to regulate their emotions and thereby to 'close the case'. We propose an alternative explanation for this behavior: letter writing might activate more pro-social modes of thinking - irrespective of how the letter writer was treated before. Pro-social modes of thinking simply arise when addressing a specific other person, in this case by thinking about writing to someone, and how to address this other person.

The additional data we use from our experiment is data from kindly treated subjects. We examine decisions of these additional subjects to test how letter writing affected them in comparison to the subjects who were treated unkindly, as described in chapter two. Our results reveal that letter writing increases giving to an innocent third person in both groups: among the kindly and the unkindly treated. This result suggests that the effect of letter writing is indeed not curative. Instead, it seems to activate more pro-social modes of thinking and to have a general pro-social effect.

Chapter 4: "Do children stop cheating when someone else reaps the benefits?" - with Julia Kramer and Silvia Lübbecke

This fourth chapter is co-authored with Julia Kramer and Silvia Lübbecke. It is concerned with the role of interpersonal factors in yet another type of immoral behavior: cheating. Here, we use a lab-in-the-field experiment to answer the question whether children's cheating behavior changes when the motives for cheating change, and motives are no longer purely selfish.

A large amount of existing literature from the economic and related disciplines regards various reasons as well as consequences of cheating. However, little is known on how such behavior originates. Especially, when cheating is not purely selfish, as it is among very young children.

In this study, we describe a lab-in-the-field experiment with 512 children. In the experiment, children privately toss a fair coin with a 50% chance of winning a prize. Subsequently, they report the outcome. Children are partnered with another child. In the treatment group, the prize in case of winning is received by the passive partner, instead of the actively participating child. By introducing this variation regarding who receives the benefit from cheating, we are able to observe what happens when

changing the motive for cheating: cheating in the treatment group is no longer purely selfish.

We find that the share of prize winning outcomes is significantly reduced when the prize is received by the partner, and motives can no longer be purely selfish. We also control for whether this depends on the social relationship to the respective partner. In this respect, we find that the difference between the treatments is largest when the partner is a stranger. In line with existing literature, the difference between the treatments becomes smaller the older the children, and eventually disappears among older children. We argue that part of the effect is explained by children cheating to prevent the partner from receiving the prize, and hence, engaging in downward cheating.

Chapter 1

Crumbling morals? An experimental study on how a social framing can affect embezzlement

Nina Lucia Stephan
Paderborn University

1.1 Introduction

When an employee decides to take a bundle of printing paper home for private use, this may seem mildly inappropriate. However, according to the Association of Certified Fraud Examiners (2016), such cases of employee stealing or embezzlement[1] on their own add up to an economic loss of 3.7 trillion US dollars a year. Ultimately, embezzlement, even though it often remains undetected, has far-reaching consequences, from injustice and exploitation (Lindebaum et al., 2017) over loss of trust and reputation (Toca Torres and Rodríguez, 2013) to business failure (Bressler, 2009). Therefore, as already argued by Falco et al. (2016) or Lindebaum et al. (2017), it should be of serious interest not only to scholars, but also to practitioners, to learn more about the circumstances under which embezzlement happens.

Existing research shows that various factors influence the decision to embezzle, or to engage in other types of employee theft. Greenberg (2002) distinguishes between two factor types: individual factors and situational factors. We understand the first type of factors (individual factors) as inherent to the decision maker, for example describing his demographics or character. The second type of factors (situational factors) are inherent to the structure of the decision situation, such as the risk of being detected or the height of potential gains. Studying both individual and situational factors helps explain who may engage in embezzlement and in which situation. However, studying these factors does not explain why one individual will sometimes behave immorally, e.g. by engaging in embezzlement, and sometimes not. Mazar et al. (2008), Gibson et al. (2013) and Gneezy et al. (2014) have shown that those who behave immorally do not behave immorally all the time. Therefore, we also consider a third type of factors that may influence embezzlement: interpersonal factors. Whether an experience preceding the decision situation involves interacting with others, is an interpersonal factor. For example, the decision maker can either be partnered with another human being, or with a computer. Another example is, whether another individual treats the decision maker in a certain way that influences the state of mind he has in a subsequent decision situation.

Examples from recent literature show that socially framing an experience, by letting another social being play a role, affects moral decisions. Houser et al. (2012) find that whether the person that a subject interacts with is a real person or not affects the likelihood that this subject subsequently engages in cheating. Subjects who feel treated unfairly by another social being, rather than by a computer, subsequently behave less morally. Our idea is that adding or avoiding a social framing also has

---

[1]This paper regards embezzlement as a special type of stealing. According to the United States Uniform Crime Reports and the Federal Bureau of Investigation, embezzlement is a theft-after-trust crime, defined as the "misappropriation or misapplication of money or property entrusted to one's care, custody or control" (Nicholson, 1999, p.62).

potential to prevent or to produce other forms of immoral behavior next to cheating, such as embezzlement. Right after being handed out a bonus pay check from his boss, who is a social being, rather than receiving the same amount of money via bank transfer, and no social being playing a role, would employees not be less likely to take home the printing paper? This leads to the central research question: How does socially framing an experience affect the subsequent decision on embezzlement?

We consider two potential answers to this question, both described by Mazar and Ariely (2006). They argue that immoral behavior, and thus also the decision to embezzle, can be driven through different channels. This paper will investigate two of these channels as potential drivers of an effect of social framing on embezzlement: lack of self-awareness and self-deception. On the one hand, self-awareness may be restored when a social framing helps individuals to remind themselves of the society that they are part of, and the moral rules that this society is comprised of. If a social framing reduces embezzlement due to increased self-awareness, this social framing should decrease overall engagement in embezzlement. On the other hand, a social framing may facilitate self-deception, when this social framing provides circumstances which an individual may use as an excuse, to justify own wrongdoings in front of himself. If a social framing affects embezzlement through facilitated self-deception, this social framing should lead to an increase, rather than a decrease, of embezzlement.

In real life, the question whether a social framing affects embezzlement, and in which way, is difficult to answer, as most cases of embezzlement remain undetected and would hence not show in observational data (Bressler, 2009). Moreover, organizations detecting cases of embezzlement among their own have an interest in hiding or even ignoring them due to image concerns. Therefore, the present study makes use of a laboratory experiment, which not only allows observing individual level decisions on embezzlement, but also allows constructing a setting with a social framing and a comparable control group, to rule out other factors that may influence the decision to embezzle. Here, the control group participates in the 'neutral framing treatment'. An experimental study also allows gathering additional data in order to gain insights with respect to the channel at work.

The experiment consists of two parts. In the first part, participants are not yet making any decisions, but they all receive an endowment. In the neutral framing treatment, endowments are received without any framing. We create a social framing treatment in which receiving an endowment consists in receiving money from a participant of a previous experiment, that contained a dictator game. Hence, the endowment varies, and can also be zero. But more importantly, this variation allows to distinguish cases in which participants in the social framing treatment perceive the dictator's decision as either moral or immoral. Following the self-deception explanation, the social framing effect should decrease embezzlement

only if a decision that is perceived as complying with moral standards is observed. Otherwise it should increase embezzlement. However, following the self-awareness explanation, the social framing effect should decrease embezzlement irrespective of the height of endowment. In the second part of the experiment, we test whether experiencing the social framing treatment affects the likelihood and magnitude of embezzlement. Here, all participants act as intermediaries, forwarding donation money (provided and entrusted to them by the experimenter) that is meant to benefit a charity (Médecins Sans Frontières). In their role as intermediary, participants have the option to manipulate the amount to be donated before forwarding it.

The results reveal an endowment-specific social framing effect: Only after receiving an average or low endowment from a previous dictator, rather than receiving an amount of equal height as an endowment in the neutral framing treatment, is the likelihood of embezzlement increased.

The remainder of this paper is structured as follows. After an introduction to existing literature in Section 1.2, Section 1.3 explains the design of the experiment. In Section 1.4, we derive our central hypotheses. Next, Section 1.5 provides information on the sample and implementation, followed by a summary of the results in Section 1.6. Section 1.7 discusses the main results and finally, Section 1.8 concludes.

1.2   Related Literature

A variety of disciplines has contributed to our understanding of why and under which circumstances people engage in embezzlement. However, many studies focus on the more general phenomenon of immoral behavior, or look at embezzlement together with other types of white-collar crimes, such as fraud or corruption. Here, we first consider relevant studies on individual factors that may influence embezzlement, followed by situational and then interpersonal factors.

The influence of individual factors on embezzlement has been studied empirically in different disciplines, often making use of data from questionnaires developed to elicit specific personal characteristics. For example, Hassan and Jan (2017) use questionnaire data to investigate the relationship between psychological capital and white-collar crime in Pakistan, including embezzlement. They find that factors of an employee's psychological capital, such as hope, optimism, self-efficacy, or resilience, are negatively correlated with white-collar crime. However, their method does not allow establishing causal links. Moreover, questionnaire data can only measure immoral behavior indirectly, either by assuming that answers to trick-questions reveal participants' immorality, or by relying on participants' truthfulness. Similar studies suggesting that the likelihood of immoral decision-making, such as the decision to engage in immoral behavior or white-collar crime, including embezzlement, depends

on different characteristics of the person who makes the decision, are also found in related disciplines such as business ethics (Crossan et al., 2013) or psychology (Schuessler and Cressey, 1950; Doris, 2002). However, already among the earlier studies, evidence was mixed, and scholars concluded that using empirical data it seems "impossible to conclude [...] that criminality and personality elements are associated" (Schuessler and Cressey, 1950, p.1). In contrast to these studies on individual factors influencing embezzlement, the experimental data we use here measure immoral behavior directly. Moreover, by controlling the design of the experiment, we are able to establish a causal link between our variable of interest, social framing, and embezzlement.

Next, we turn to studies on situational factors influencing the decision to embezzle. Again, existing literature from different disciplines has contributed to our understanding of relevant situational factors, including studies from the field of law and criminology (e.g. Hochstetler et al., 2002; Agnew et al., 2009), management or business ethics (e.g. Schnatterly, 2003), psychology, and economics. Economic studies on situational factors include Makowsky and Wang (2018), who use experiments to study the effect of different organizational architectures on embezzlement, and find horizontal and pyramid structures to be most suitable to avoid embezzlement. Similarly, Falco et al. (2016) find experimental evidence of transparency influencing embezzlement. And Chlaß et al. (2015) show that when intermediaries are involved, there is less charitable giving. Even though such results are very relevant for organizations, re-considering their organizational architecture or the involvement of additional intermediaries may not be feasible for all of them. Falco and his co-authors also point out another deficiency of many existing studies on embezzlement. They criticize the lack of studies investigating embezzlement with real negative externalities.

This deficiency also holds for studies such as Crumpler and Grossman (2008) or Luccasen and Grossman (2017). For different reasons, such as allowing to investigate very specific behavioral motives, they are designed such that the embezzled amounts are replenished by the experimenters later on. While there are reasons for these studies to be designed that way, we argue that designing real negative externalities is necessary when the aim is to learn how to get embezzlement under control in real life.

Some existing experimental studies on embezzlement already involve real negative externalities. Attanasi et al. (2018) study how different types of guilt aversion affect embezzlement. However, embezzlement does not hurt a party in need, but only leads to lower earnings for other participants within the experiment. This stands in contrast to most relevant cases of real-life embezzlement, as well as embezzlement in our study. The present study will contribute to this literature by looking at embezzlement of 'real' donation money, where embezzlement is harming a good cause, as it reduces

real donations made to Médecins Sans Frontières.

One example for an experimental study already involving real negative externalities at the expense of a party in need, namely the potential loss of measles vaccines donated by UNICEF, is Kirchler et al. (2015). They provide evidence of the effect of yet two other situational factors on embezzlement: introducing punishment and removal of anonymity.

Studying the aforementioned or other situational factors may contribute to reducing embezzlement. However, it does not answer the question why the same individuals, repeatedly living the same situations, will sometimes embezzle and sometimes not. Therefore, we will next turn to interpersonal factors that may influence embezzlement. The general idea is that while individual as well as situational factors remain the same, the type of interpersonal experience within a given situation may lead any individual to behave sometimes more and sometimes less in line with moral standards.

The idea of interpersonal factors being relevant for moral decision-making is supported by neuroscientific research making use of Magnetic Resonance Imaging (MRI) to explain choice making. For example, MRI studies have shown that when one specific decision situation involves another human, instead of a computer player, there are more positive feelings that turn the situation into a more pleasant experience, all else equal. Many studies (Rilling et al., 2002, 2004; Rilling and Sanfey, 2011; Fareri et al., 2012) provide evidence, of such additional emotional benefits, stemming from other humans, instead of computers or moves of nature, being involved in one way or the other. This is also in line with studies on business ethics (e.g. Lindebaum et al., 2017) that describe how decision makers may blend out moral concerns, depending on social context.

The present paper relates to this neuroscience literature by investigating whether the influence of making an experience involving another human, and the potentially resulting positive feelings, also has an effect on the likelihood of the subsequent decision to embezzle. By producing such an emotional benefit, the social framing may make potential material gains from embezzlement relatively less important. This would explain a generally lower level of embezzlement in the social framing treatment, but would not help understanding how such effects interact with different levels of endowment. The latter will therefore be addressed in the paper at hand.

Our study is even more closely related to some recent studies on the effects of social framing in the field of economics. Capraro and Marcelletti (2014) study the effect of receiving an endowment by a dictator, rather than a random mechanism, on the recipients' willingness to donate. While investigating the same interpersonal factor, we are interested in its effect on a different outcome variable: the decision to engage in embezzlement. Moreover, the study by Capraro and Marcelletti is carried out on Amazon Mechanical Turk (MTurk), an online crowdsourcing

marketplace. According to the authors themselves, this came with the drawback that some participants did not believe in the existence of their fellow participants. Since the experiment in the present paper is carried out in an experimental laboratory, we are able to avoid this problem. Additionally, in the design of this experiment, special care is put into ensuring that participants understand and believe the real consequences that their decisions on embezzlement have for the affected charity Médecins Sans Frontières.

The present paper most closely relates to Houser et al. (2012). Their study includes a treatment in which participants receive endowments from a previous dictator, just as in our social framing treatment, and a treatment in which similar endowments are distributed without any framing ('no intentions treatment'), just as in our neutral framing treatment. However, the authors focus on a different outcome variable: cheating, measured in a coin-flip game. The decision to engage in embezzlement, and thereby creating a real negative externality that is harming a good cause, is not part of the design. Moreover, in a coin-flip game with only two possible outcomes, participants can only decide on abstaining from immoral behavior, or going all the way. In contrast to that, we are able to measure the different magnitudes of immoral behavior. Despite these differences regarding the aim of the study and the details of the design, the overall setup of Houser et al. (2012) still closely resembles the design of our experiment. Therefore, it is not surprising that they find that individuals who believe they were treated unfairly by another real person are subsequently more likely to cheat in a coin-flip game, and are thus more likely to behave immorally. This is in line with the negative effect of the social framing treatment on embezzlement, in the case of receiving a less-than-half share from the dictator, as reported in the study at hand.

To the best of our knowledge, this study is the first to investigate the effect of a social framing on 'real' embezzlement. Even though embezzlement is similar to other forms of immoral behavior, such as lying or cheating, the literature presented above provides various reasons for studying embezzlement in the particular way it is done in the study at hand. First, we contribute by investigating a factor that can potentially alter subsequent decision situations and reduce both the likelihood and the magnitude of embezzlement. Second, we are interested in embezzlement of donation money, where embezzlement is indisputably harmful and hence adequately reflects embezzlement in real life. Thereby, the contribution of this study also differs from most experimental studies on cheating, or other forms of immoral behavior, where either other participants or the experimenter are affected negatively, but no 'real' harm done. Third, this study distinguishes between general effects of social framing and interaction effects with different levels of endowment, and provides evidence suitable for discussing different channels that may drive the various effects. Thereby, the results contribute to research on embezzlement in the economic as well

as related disciplines.

1.3 Experiment

This study uses a controlled laboratory experiment designed to test whether, and in which way, a social framing can be used to influence participants' subsequent decision to enrich themselves by embezzling donation money. The experiment consists of two main parts, described in the following. Thereafter, this section explains additional questions regarding the participants' happiness and perception of norms, that are part of the design. The experiment ends with a short questionnaire. An English translation of the instructions is provided in the Appendix.

1.3.1 Part I

Part I is the treatment stage. Participants are either in a socially framed or a neutrally framed setting. Within this setting, they receive an initial endowment.

In the social framing treatment, participants receive endowments in the form of money given up by a participant of a previous different experiment. This previous experiment took place in May and June 2017 in the Paderborn University BaER-Lab.[2] As part of this previous experiment, participants played a dictator game in which they could leave a share of 10 € to a 'future participant of a different experiment'. Participants of this previous experiment were not invited to participate in the experiment at hand. This information is also provided to the social framing treatment participants in the experiment at hand. They are informed that receiving a positive amount of money means that the respective previous participant has actually given up this same amount of money. Hence, depending on the decision of the respective previous participant, they can receive between 0 € and 10 € in increments of 1 €. This variation in endowments is a crucial part of the design, because it allows to study whether there are interaction effects between the endowment and the social framing treatment. As will be described in the Section 1.4, there are reasons to expect interaction effects between the endowment and the social framing treatment. Therefore, this variation in endowments is key in our design, because it enables identifying the general social framing effect.

Since receiving money from a dictator is a situation that involves another human, a social context is created that can remind participants of the society they are part of, and may therefore affect moral decision-making. To make sure that potential social framing treatment effects do not only stem from participants having received money, and thus being richer, there is a control group: the neutral framing treatment. In the

---

[2]Further details on this previous experiment are available upon request.

neutral framing treatment, participants receive the initial endowment as a starting budget, and thus in a neutral setting not involving another human.

In order to ensure that the existence of a social framing is the only difference between the treatments, two more measures are taken. First, the distribution of initial endowments in the neutral framing treatment is designed to be the same as that of the social framing treatment. Consequently, following the decisions taken by the dictators in the previous experiment, the very same amounts as shared by the dictators are handed out as initial endowments in the neutral framing treatment. Second, to avoid differences in upfront expectations regarding the endowments, participants in both treatments receive the same information about the distribution of initial endowments. They learn that the amounts can be between 0 € and 5 € in increments of 1 €. In order to rule out that participants in the neutral framing group assume a uniform distribution, which is unlikely to be expected from amounts shared by dictators in the social framing treatment, instructions also reveal the median. In order to keep explanations simple, both groups are informed that in 55 out of 100 cases, the amount is 2 € or smaller.

1.3.2 Part II

In Part II of the experiment, participants act as intermediaries, forwarding real donation money. Knowing that the willingness to give may significantly differ, depending on the type of receiver (Eckel and Grossman, 1996; Fong, 2007; Cappelen et al., 2017), the donations the intermediaries are entrusted with go to an established charity: Médecins Sans Frontières. Since this experiment is concerned with the decision whether or not to embezzle donation money, rather than with the decision to donate, the donation money is provided by the experimenter, and all participants are entrusted with handling a donation of the same amount (3.50 €).[3] On their screens, participants are shown an envelope addressed to Médecins Sans Frontières and containing a 3.50 € donation. In order to ensure that participants trust in donations being real donations, the real donation money is collected in cash in one real envelope in the front of the laboratory, and participants are informed that they can inspect it at the end of the experiment. Moreover, during each session, one participant is randomly selected to witness the transfer of the donations from that session, and receipts are made available online to all participants.

In order to guarantee that participants do not make an error when deciding, as well as to reduce potential demand effects, this second part of the experiment consists

---

[3]Despite this not being a large amount, it is arguably representative of the most frequently occurring types of embezzlement in real life. Moreover, just as in real life, costs from single cases of embezzlement accumulate to large sums. According to Bressler (2009) a total loss of 1.5 million US dollars accumulated due to petty theft in one cafeteria, and a loss of 300,000 US dollars accumulated due to diversion of postage stamps at a single university.

of a two-step procedure. First, participants decide whether to immediately confirm the forwarding of the donation, by clicking on a closed envelope, or to manipulate[4] the donation, by clicking on an envelope that is cut open, as illustrated in Figure 1.1.

Teil 2: (Spenden)

Bitte wählen Sie duch Anklicken eines Bildes aus, ob Sie den Spendenbetrag von 3,50 € an Ärzte ohne Grenzen bestätigen oder manipulieren (also vorab etwas für sich selbst entnehmen) wollen.

enthält 3,50 €
für Ärzte ohne Grenzen

enthält 3,50 €
für Ärzte ohne Grenzen

manipulieren   bestätigen

Figure 1.1: Screenshot from the first stage of Part II. By clicking on the respective envelope, participants decide to either confirm and forward the donation immediately, or to first manipulate the donation.

Second, in the case that the open envelope is selected, participants are asked to enter the amount of money they want to take out of the envelope. This way, we avoid measuring embezzlement when participants actually clicked on the wrong envelope either by mistake, out of curiosity, or because they perceive a demand to do so. They can enter any amount in increments of $0.10 \in$, up to the maximum of $3.50 \in$. Manipulating the donation comes with a loss of $0.20 \in$, which is subtracted from the money taken from the donation. To the participants, this is explained as the loss stemming from the extra effort of having to open and close the envelope and adjust the sum. Hence, embezzling is inefficient. [5]

---

[4]The term "manipulate" was used in this part of the experiment to avoid that participants perceive the task as simply splitting up money between themselves and the charity.

[5]Making embezzlement inefficient reflects real-life consequences of embezzlement. Moreover, it reduces the likelihood of participants deciding to maximize profit in order to donate the profits to another charity later on, outside of the experiment. Such behavior could otherwise be mistaken for selfish and immoral behavior.

We are also interested in the extent to which participants embezzle. Evidence from different previous experiments shows that often, participants decide not to go all the way when it comes to immoral decisions, such as the decision to lie (e.g. Fischbacher and Föllmi-Heusi, 2013) or to cheat (e.g. Mazar et al., 2008; Pruckner and Sausgruber, 2013). Allowing participants to decide on the magnitude of embezzlement comes with several advantages. First, even though embezzlement does not involve any risk, participants may shy away because of image concerns. They may feel observed by the experimenter and may want to hide or at least downplay their true type. By allowing partial embezzlement, participants can choose the amount such that their total payment will not stand out during the payment, which may help them appear less greedy. Second, just as is possible in real life, taking a very small amount may also help participants downplay their wrongdoing in front of themselves, and maintain a positive self image. Third, partial embezzlement also captures the general willingness to engage in embezzlement by those for whom the arbitrary jump from taking nothing to taking 3.50 € would be slightly too extreme. This way, we avoid mis-measuring the general willingness to embezzle due to imperfections in setting the design parameters.

1.3.3 Additional Control Variables

In order to gain more insight into the reasons behind the participants' decisions, we gather information on participants' emotions, as well as their perception of behavioral norms.

According to Motro et al. (2016), emotions often influence immoral behavior. More specifically, Schweitzer and Gibson (2008) argue that after observing perceived violations of moral behavior by others, individuals are more likely to derive happiness from engaging in immoral behavior themselves. In order to learn whether emotions play a role in explaining potential social framing effects, we elicit participants' happiness and measure how it changes during the experiment, based on a method first used by Bradley and Lang (1994). In the beginning of each part of the experiment, as well as at the end, participants are asked to describe their current mood on a nine-point Likert scale by clicking on a respective *Self-Assessment Manikin*.

Irlenbusch and Villeval (2015) argue that when people apply self-deception to persuade themselves of an immoral decision, this is facilitated by believing that, under the same circumstances, peers would do the same. Thus, if people are not acting upon their emotions, an alternative factor driving differences in embezzlement between treatments may be that beliefs regarding the behavioral norm differ between groups. In this case, which decision to make, in order to fit in with the group, will also

differ between treatments and heights of endowment. In order to shed light on this potential driving factor for our results, we elicit participants' perceived behavioral norm regarding the decision to embezzle or not, by adding a bonus question at the end of the experiment. Participants are informed about the number of participants in the laboratory. They are asked to guess in how many cases a fellow participant decided to forward the donation without taking anything. A correct guess is rewarded with a bonus of 1 €. From these guesses, we calculate the guessed share of moral peers, which we will refer to as 'guessed share'.[6]

1.4 Hypotheses

In this section, we build on theoretical considerations from existing literature to formulate our expectations regarding the results of our experiment. As mentioned in the introduction, Mazar and Ariely (2006) distinguish general drivers of dishonest behavior. These four drivers are: low external costs compared to benefits, a lack of social norms, a lack of self-awareness, and self-deception. Embezzlement, as a special type of dishonest behavior, may be driven by any of these factors. In the case at hand, external costs, such as costs from being detected and punished, are mostly absent[7], and the benefits from dishonest behavior are held constant. Social norms may differ among the participants in our experiment. However, assuming successful randomization, on average there should not be any systematic differences between different conditions. Hence, in this study, differences in dishonest behavior may be driven either by differences in self-awareness or by differences in self-deception. We will consider both types of drivers in turn.

1.4.1 Lack of Self-Awareness as a Driver of Embezzlement

Mazar and Ariely (2006) argue that independent of the exact way how an individual's internal reward system functions, any such system influences actions only when decision makers make active use of it. Several studies have shown that by manipulating the circumstances of a decision situation, and thereby activating self-awareness, dishonest behavior may be reduced. For example, reduced dishonest behavior is found by Carver and Scheier (1998) when letting subjects stand in front of a mirror, or by Duval and Wicklund (1972) in presence of a real versus an implied audience.

Mazar and Ariely (2006) explain that when such measures to increase self-awareness are in place, subjects become aware of themselves as objects in the world.

---

[6]The participant's guess, divided by the number of participants in the laboratory.

[7]We are aware of potential psychological costs from being detected by the experimenter; however, we have no reason to believe that such costs differ across treatments.

This raises awareness for noticing discrepancies between one's own actions and one's ideal self. The authors further argue that noticing such discrepancies is often accompanied by emotional arousal. These feelings increase the motivation to adopt a behavior more in line with one's own standards. This is why increased self-awareness may lead to more honest behavior. In support of this channel, Schachter and Latane (1964) show that under the influence of a tranquilizer inhibiting sympathetic feelings, dishonesty is more likely to survive.

In the present experiment, the social framing treatment involves one specific other individual, namely the previous dictator. By involving this other individual, the social framing treatment invites participants to think about other humans in general. While thinking about other humans, we are reminded of our own moral standards according to which we live amidst other humans. We may generally have more sympathetic feelings (Schachter and Latane, 1964), and we may feel that it is harder to ignore own moral standards. Hence, we expect self-awareness to be increased in the social framing treatment in comparison to the neutral framing treatment, resulting in fewer immoral choices. Under the assumption that in the present experiment the social framing treatment increases participants' self-awareness, we derive the following hypothesis:

H1a: The average likelihood of embezzlement is lower in the social framing treatment in comparison to the neutral framing treatment.

Even though increased self-awareness may discourage embezzlement in the social framing treatment, this is not the only factor driving the decision (Mazar and Ariely, 2006). For example, within both the social framing and the neutral framing treatment, there may be individual differences regarding the value participants place in the potential benefits from embezzlement. On the individual level, participants may still trade off subjectively valued benefits against conflicting moral standards. When self-awareness is high, moral standards become more salient, which may affect this trade-off. However, self-awareness does not necessarily override all other factors. If there is increased self-awareness in the social framing treatment, some participants may thus refrain from embezzlement all along, but others may end up only reducing the extent to which they engage in embezzlement. Therefore, under the assumption that the social framing treatment increases self-awareness, we derive the following hypothesis with respect to the magnitude of embezzlement:

H1b: The average magnitude of embezzlement is lower in the social framing treatment in comparison to the neutral framing treatment.

1.4.2 Self-Deception as a Driver of Embezzlement

According to Mazar and Ariely (2006), self-deception should be considered as another, or an additional, potential driver of dishonest behavior. We therefore consider self-deception as a potential driver of embezzlement. The authors speak of self-deception when decision makers use the given circumstances to re-frame their actions, such that they no longer interpret these actions as dishonest. In other words, decision makers may use the circumstances and recent history of the situation as an excuse to justify their choices.

For example, Cojoc and Stoian (2014) show that cheating increases when subjects learn that there will be an opportunity to subsequently donate part of their gains — though donations do not actually increase. Similarly, Croson and Shang (2008) show that announcing reduced donations recently made by other participants affects subsequent contributions in a public goods game. Keizer et al. (2008) show that signs of norm violations trigger further violations of legal regulations. Similarly, with the help of a business ethics questionnaire, Schweitzer and Gibson (2008) show that subjects are more likely to indicate that they would be willing to engage in unethical behaviors when they perceive their counterparts to violate fairness standards.

In the case at hand, people may also engage in self-deception by evaluating their own choice in light of others' choices in recent history. In the social framing treatment, the obvious other choice to relate to is that of the previous participant, who acted as a dictator. Depending on the dictator's decision to share, this may be used as an excuse to justify one's own immoral behavior. In other words, when someone else did already break moral standards, participants may feel less obliged to uphold moral standards themselves. They may also interpret the situation as if the other person's immoral behavior resulted in a moral license for themselves to also behave immorally.[8]

Of course, using someone else's choice as an excuse is conditional on perceiving this other choice as violating a moral standard. Hence, the existence of a previous participant in the social framing treatment can facilitate self-deception only if the previous participant's sharing decision is perceived as violating a moral standard. Following Andreoni and Bernheim (2009), this study assumes that everything but a fair share of 50% may be regarded as a violation of moral standards. Consequently, engaging in self-deception should be easier in the social framing treatment, at least when the previous dictator shares less than half. Thus, under the assumption that the social framing treatment in the present experiment facilitates self-deception, we expect an effect on embezzlement, dependent on the behavior of the dictator, and derive the following hypothesis:

---

[8]Usually, moral licensing refers to a person feeling free to misbehave, conditional on having behaved morally at another point in time. A recent review of moral licensing literature has been published by Blanken et al. (2015).

H2a: The likelihood of embezzlement is higher in the social framing treatment in comparison to the neutral framing treatment, given that the dictator shares less than half.

Following the same logic as before, facilitating self-deception may not only affect the decision whether to embezzle at all, but also the decision to which extent to embezzle. For example, if the dictator decides to share something in between zero and less than half, it may seem appropriate to also engage in embezzling partly. Assuming again, that in the present experiment the social framing treatment facilitates self-deception, we therefore also expect an effect on the magnitude of embezzlement, and derive the following hypothesis:

H2b: The magnitude of embezzlement is higher in the social framing treatment in comparison to the neutral framing treatment, given that the dictator shares less than half.

In the next section, in order to answer the question how a social framing actually affects the decision to embezzle, we turn to the data gathered in our experiment and to testing the hypotheses.

1.5 Implementation and Sample

The experiment was conducted in July 2017 in the Paderborn University BaER-Lab in Germany. The experiment was designed using the software z-Tree (Fischbacher, 2007), and ORSEE (Greiner, 2015) was used for recruitment. A total of 131 participants were recruited at Paderborn University, excluding no-shows. Sessions lasted approximately 30 minutes with an average payment of 6.25 €.

	neutral framing	social framing
age	22.91	23.70
male (in %)	42.19	37.31
undergrads (in %)	68.75	64.18
economics students (in %)	35.94	41.79
initial happiness	5.91	5.81
observations	64	67

Table 1.1: To check randomization, average values across treatments are compared using Mann-Whitney-U tests and Fisher exact tests.We do not find significant differences and presume successful randomization.

As shown in Table 1.1, randomization seems successful. Using a Mann-Whitney-U test, or Fisher exact test in case of binary variables, does not reveal significant differences between the groups regarding any of the listed variables. Due to fewer no-shows, there are slightly more observations in the treatment group, and the distribution of the amounts of money received as initial endowments in the group without social framing is not perfectly equal to that in the group with social framing. Still, the amounts received are very similar in both groups.[9] The average shared amount, 2.37 €, is within the range of amounts typically shared in dictator games.[10]

1.6 Results

To describe the results, this section investigates each hypothesis in turn. It presents summary statistics as well as results from multivariate analyses. Then, it moves to ancillary findings.

1.6.1 Assuming Lack of Self-Awareness as a Driver of the Likelihood of Embezzlement

We begin by testing the hypotheses derived from the assumption that self-awareness is driving embezzlement behavior. First, we investigate whether the average likelihood of embezzlement is lower in the social framing treatment (*H1a*). To determine whether the likelihood of embezzlement, as displayed in Figure 1.2, actually differs across treatments, we use a Mann-Whitney-U test.

---

[9]We use amounts shared by 67 dictators who have participated in the previous experiment in May and June 2017, as described in Section 1.4.

[10]A literature review of dictator games is provided by Camerer (2003) or Engel (2011), who show that in many games dictators share around 24% of the available amount.
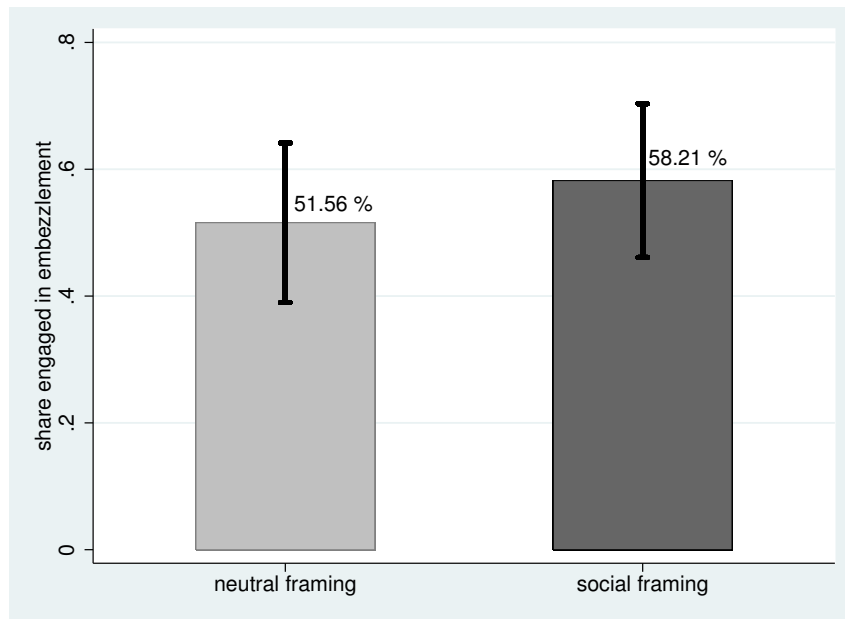
Figure 1.2: Likelihood of embezzlement in the neutral framing and the social framing treatment.

There is no significant difference between the overall average shares of participants engaging in embezzlement. Hence, we reject Hypothesis H1a and conclude:

Result 1a: The average likelihood of embezzlement is not lower among participants in the social framing treatment in comparison to the neutral framing treatment.

1.6.2 Assuming Lack of Self-Awareness as a Driver of the Magnitude of Embezzlement

Next, we investigate the hypothesis that the average magnitude of embezzlement is lower in the social framing treatment (*H1b*). To do so, we focus only on those 72 subjects who decided to engage in embezzlement. Figure 1.3 displays average magnitudes of embezzlement, and Figure 1.3 displays the distribution of embezzled amounts in the neutral framing and the social framing treatment.
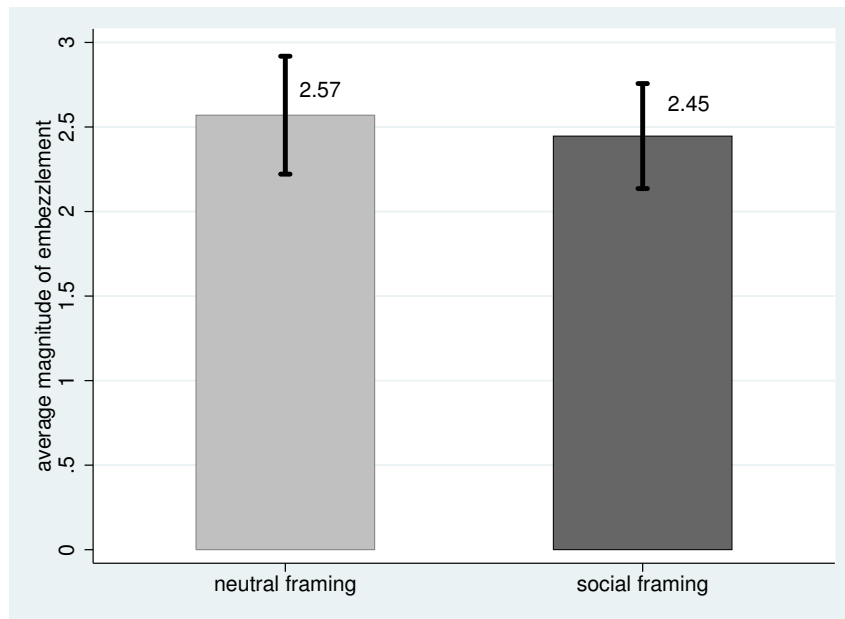
Figure 1.3: Average magnitudes of embezzlement, excluding participants who do not engage in embezzlement.
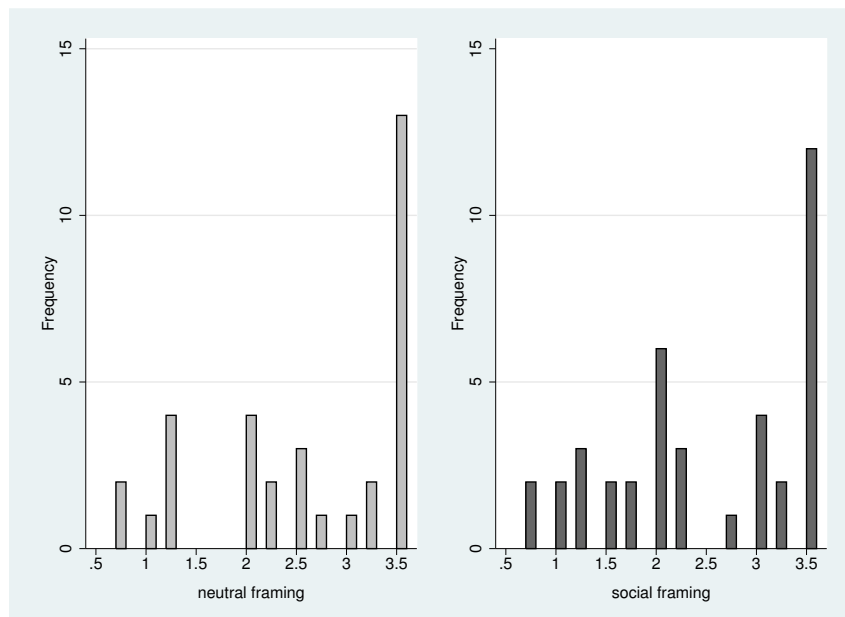

Figure 1.4: Distribution of embezzled amounts of those engaging in embezzlement in the neutral framing and social framing treatment.

In both treatments, distributions are left skewed, with about one third of participants (13 out of 33 in the neutral framing, and 12 out of 39 in the social framing treatment) choosing to embezzle to the full extent. Average magnitudes of embezzlement, according to a Mann-Whitney-U test, are not significantly different between the social framing and the neutral framing treatment, so we conclude:

Result 1b: The average magnitude of embezzlement is not lower among participants in the social framing treatment in comparison to the neutral framing treatment.

1.6.3 Assuming Self-Deception as a Driver of the Likelihood of Embezzlement

We turn to the hypotheses derived from the assumption that self-deception is a driver of embezzlement. First, we investigate the hypothesis that the likelihood of embezzlement is higher in the social framing treatment in comparison to the neutral framing treatment, given that the dictator shares less than half (H2a). Of course, participants in the neutral framing group did not receive endowments from dictators; hence, we compare averages conditional on the endowments in both treatments being lower than $5 \, €$, as shown in Figure 1.5.
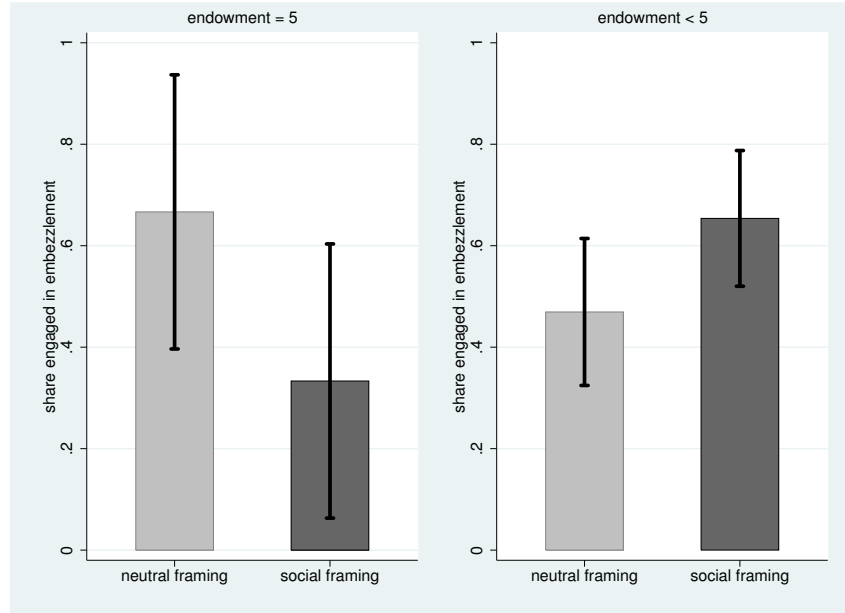

Figure 1.5: Likelihood of embezzlement in the neutral framing and the social framing treatment, conditional on either having received an endowment of $5 \, €$ or below.

It becomes clear from looking at Figure 1.5 that in the case of an endowment below 5 €, the likelihood of embezzlement is higher in the social framing treatment. A Mann-Whitney-U test shows that this difference is statistically significant at the 10% level (p-value 0.0630). Hence, we cannot reject Hypothesis H2a and conclude:

Result 2a: The likelihood of embezzlement is significantly higher when receiving a less-than-half share from a dictator in comparison to receiving less than 5 € in the neutral framing treatment.

Figure 1.5 also reveals a weakly significant difference across treatments (Mann-Whitney-U test p-value 0.072) regarding the likelihood of embezzlement when endowments are equal to 5 €, and dictators are sharing half. However, in this case the effect of the social framing goes the other way.

To better describe this overall endowment-specific effect of the social framing treatment in comparison to the neutral framing treatment, we provide results from a Probit regression analysis, with engaging in embezzlement as the dependent variable. We include the interaction term *social framing x endowment* to separate the endowment-specific effect of the social framing treatment. In order to get a closer estimate of the effect size, we also control for gender, since existing experimental evidence[11] shows that gender has a significant infuence in many different forms of both pro-social and immoral decision-making. Ignoring potential gender effects could therefore lead to biased results. Table 1.2 displays marginal effects.

Generally, the likelihood of embezzlement is 29.4 percentage points higher in the social framing treatment in comparison to the neutral framing treatment (see Table 1.2). On average, this general social framing effect is reduced by 9.3 percentage points for every additional 1 € of endowment. Hence, the effect is stronger the lower the endowment, and even turns negative for endowments of 4 € or 5 €. However, this does not imply that the likelihood of embezzlement is higher only when endowments are less than 4 €, since there is also a general and significant endowment effect.

---

[11]E.g. Eckel and Grossman (1998); Andreoni and Vesterlund (2001); Dufwenberg and Muren (2006); Houser et al. (2012); Reinstein and Riener (2012); Capraro and Marcelletti (2014).

	Probit1	Probit2
social framing	0.302**	0.294**
	(0.129)	(0.131)
endowment	0.070**	0.065**
	(0.033)	(0.033)
social framing x endowment	-0.101**	-0.093**
	(0.045)	(0.046)
male		0.159*
		(0.088)
_cons	-0.393	-0.537
	(-1.52)	(-1.96)
n=131, (se),	* p<0.1, ** p<0.05, *** p<0.01	

Table 1.2: Estimates from Probit regressions with the dependent variable taking the value 1 if the participant did engage in embezzlement. The table displays marginal effects with standard errors in parentheses.

Interestingly, there is a significantly positive marginal effect of 6.5 percentage points. This means that within the neutral framing treatment, the likelihood for embezzlement is larger among participants with higher endowments. Concluding, the overall evidence is completely in line with Result 2a. On a side note, all results are also robust to controlling for gender. In line with existing literature (e.g. Eckel and Grossman, 1998; Andreoni and Vesterlund, 2001; Dufwenberg and Muren, 2006; Houser et al., 2012; Reinstein and Riener, 2012; Capraro and Marcelletti, 2014), the gender being male leads to a 15.9 percentage point (p<0.1) increase in the likelihood of embezzlement.

1.6.4 Assuming Self-Deception as a Driver of the Magnitude of Embezzlement

Finally, we investigate the hypothesis that the magnitude of embezzlement is higher in the social framing treatment in comparison to the neutral framing treatment, given that endowments are below 5 € (H2b). We start by comparing average embezzled amounts across treatments by those who did engage in embezzlement, as displayed in Figure 1.6.
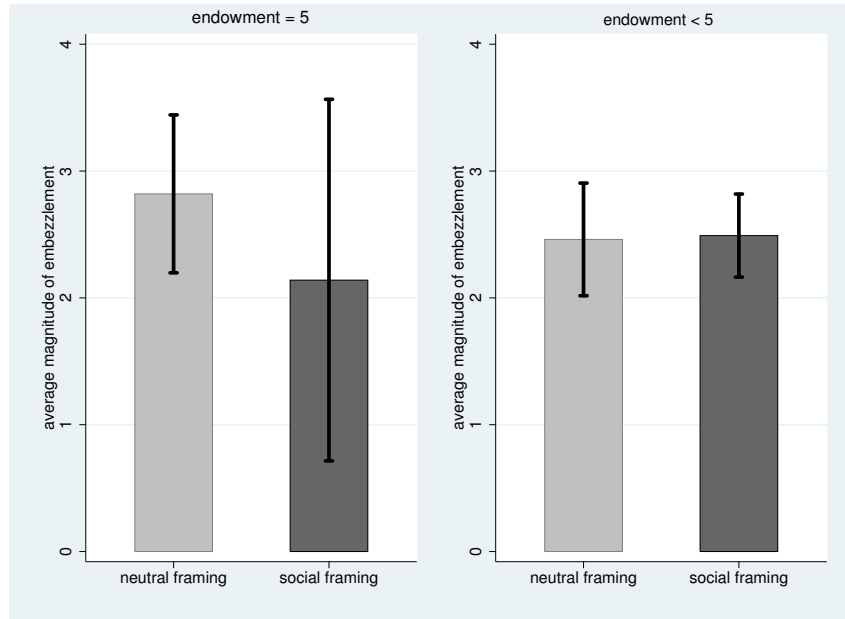
Figure 1.6: Average magnitudes of embezzlement of those engaging in embezzlement in the neutral framing and the social framing treatment, conditional on either having received an endowment of 5 € or less.

Figure 1.6 shows that in the case of an endowment below 5 €, the average magnitudes of embezzlement do not differ across treatments. Hence, we can reject Hypothesis H2b and conclude:

Result 2b: The magnitude of embezzlement when receiving a less-than-half share from the dictator does not differ from the magnitude of embezzlement when receiving less than 5 € in the neutral framing treatment.

Moreover, also the small difference in average embezzled amounts among those who received an endowment of 5 €, displayed in Figure 1.6, is not statistically significant. To make sure that there is no other endowment-specific effect of the type of framing on the magnitude of embezzlement, Table 1.3 provides results from a Tobit regression analysis, with the embezzled amount as the dependent variable.

	Tobit1	Tobit2	Tobit3
social framing	-0.244	-0.041	0.109
	(0.34)	(0.57)	(0.56)
endowment	-0.126	-0.061	-0.054
	(0.08)	(0.13)	(0.13)
social framing x endowment		-0.108	-0.135
		(0.18)	(0.17)
male	0.582*		0.602*
	(0.34)		(0.34)
_cons	2.975***	3.098***	2.747***
	(0.39)	(0.46)	(0.49)
sigma _cons	1.312***	1.340***	1.306***
	(0.13)	(0.12)	(0.13)
n=72, (se)		* p<0.1, ** p<0.05, *** p<0.01	

Table 1.3: Estimates from Tobit regressions with the embezzled amounts (up to 3.50 € in increments of 0.10 €) as a dependent variable, right-censored with the upper limit set to 3.5, and robust standard errors displayed in parentheses. Observations with 0 € embezzled are omitted.

Estimates from Table 1.3 do not provide any evidence in favor of H2b either. We conclude that, given that participants engage in embezzlement, the social framing treatment in comparison to the neutral framing treatment does not influence the embezzled amount. The magnitude of embezzlement is not higher in the social framing compared to the neutral framing treatment, given that the dictator shares less than half. This Result 2b is robust to controlling for endowment-specific effects, as well as to gender.

1.6.5 Ancillary Findings

Before discussing the results, we provide an overview of further insights from our data that may help understand the results. First, we describe elicited happiness values, and changes in happiness, then we describe the participants' perceived behavioral norms.

	neutral framing	social framing
initial happiness	5.91	5.81
happiness1	5.06	5.40
happiness2	5.73	6.24

Table 1.4: Average elicited happiness values (min 1, max 9) across treatments.

29

As shown in Table 1.4, while starting from similar levels, the happiness values elicited after receiving the endowments (happiness1) are a bit higher in the socially framed setting. In both treatments, the height of the endowment is significantly positively correlated with the resulting change in happiness, with a stronger significance and higher correlation in the social framing treatment (0.658, p<0.01) in comparison to the neutral one (0.306, p<0.05). Intuitively, it seems that people reward receiving money that another human has given up for them higher than when receiving an equal amount by chance.

Despite these differences across treatments, elicited happiness does not help explain the treatment effect on the decision to engage in embezzlement. In neither treatment are levels or changes in happiness correlated with the subsequent decisions to embezzle.

However, focusing only on those participants who do engage in embezzlement, we find a negative correlation (-0.1765, p-value <0.05) between happiness1 and the magnitude of embezzlement. Hence, whereas emotions may not explain whether embezzlement is taking place, they may well influence to what extent embezzlement is taking place.

	neutral framing	social framing
guessed share (embezzlement=0)	66.32	67.96
guessed share (embezzlement=1)	28.67	28.92

Table 1.5: Averages of guessed shares of how many fellow participants abstain from embezzlement. These are averages of guessed shares by those who do embezzle themselves (embezzlement=0) and those who do do not embezzle themselves (embezzlement=1).

Table 1.5 shows that in both treatments, guessed shares are higher among those who themselves do not embezzle either, and lower among those who do engage in embezzlement. Hence, in both treatments, many participants guess that their decision on engaging in embezzlement is also the more common decision among their fellow participants. Thus, many participants seem to act in line with their perceived norm.

Interestingly, there may be a difference between treatments regarding how this perceived norm of engaging in embezzlement originates. Whereas the correlation between guessed shares and endowments is weakly significantly positive (0.2167, p<0.10) in the social framing treatment, it is highly significantly negative (-0.3213, p<0.01) in the neutral framing treatment. The next section will further discuss in what way these findings support and help explain the results.

1.7 Discussion

From Results 1a and 1b we learn that when considering the entire sample without controlling for endowment effects, we cannot detect any general effect of the social framing treatment, as would have been expected from assuming that the social framing strengthens self-awareness. One reason may be that self-awareness may already be quite high to start with, and can thus not be increased even further. This stands in contrast to results from MRI studies (e.g. Rilling and Sanfey, 2011) or criminology studies (e.g. Schachter and Latane, 1964), according to which situations involving another human should in general be influenced as more positive, that contributed to deriving Hypothesis H1a. However, in light of Result 2a, we cannot rule out that in some cases, self-awareness may still be increased and at least partly add to the effects that are detected when controlling for endowments, as suggested by Mazar and Ariely (2006): especially because results from the Probit regression analysis reveal that next to the interaction effect of endowment and social framing, there remains a general social framing effect. This underlines the importance of the design of the experiment. For revealing a social framing effect, it is crucial to consider settings with different levels of endowments, as well as to control for pure endowment effects by also varying endowments in the neutral framing treatment.

Result 2b best captures the central findings of this paper. This result is in line with the assumption that self-deception, as also suggested by Mazar and Ariely (2006) works as a driver of embezzlement. We argue that the existence of the dictator, whose decision can be used as an excuse for the own decision, facilitates self-deception. The ancillary finding on guesses of shares of peers who abstain from embezzlement support this argument. When participants find it difficult to self-deceive and re-frame their wrongdoing, because the perceived norm is to abstain from embezzlement, we should find higher guessed shares among those who decide not to embezzle. This is clearly the case in our experiment, as shown in Table 1.5. This result is in line with a lot of previously mentioned examples from existing literature, showing that when individuals believe that someone else has just made an immoral choices, they are subsequently less reluctant to also engage in wrongdoings themselves, such as Croson and Shang (2008) or Keizer et al. (2008). These are not the only recent study describing that immoral behavior is contagious (Croson et al., 2009; Robert and Arnab, 2013; d'Adda et al., 2017). Just as argued by Irlenbusch and Villeval (2015), it seems that participants in our experiment applied self-deception to excuse their immoral actions in front of themselves, by believing that peers do the same. Our additional control variable, measuring 'guessed shares' of peers engaging in embezzlement, helps confirm this argument. Overall, in the case at hand, guessed shares correctly predict embezzlement in both groups: participants tend to act according to their perceived behavioral norm. In both groups, many of those

who decide to embezzle expect this to be the more common choice. The correlation between guessed shares and engaging in embezzlement is -0.6608 (p<0.01) in the social framing treatment, and -0.7349 (p<0.01) in the neutral framing treatment.

It is straightforward to argue why and in which cases the norm to abstain from embezzlement originates in the social framing treatment (namely dependent on the previous dictator's sharing decision). In contrast, explaining why and in which cases this behavioral norm originates in the neutral framing treatment is less straightforward. In the neutral framing treatment, no particular choice of any other participant is observed that could serve as a justification for one's own choices. Clearly, other drivers of the decision-making than the ones investigated in the study at hand are at play here. We find the opposite from what we find in the social framing treatment, regarding the correlation between endowments and perceived norms: In the neutral framing treatment, when receiving a high endowment, the perceived norm seems to be to engage in embezzlement. It may be that being lucky (compared to being shared with) creates a feeling of entitlement, as suggested by Schurr and Ritov (2016) and/or produces greed. In turn, it may be that being unlucky (as opposed to not being shared with) creates a feeling of resignation, and/or produces frugality. At this point, we do not have evidence to draw any conclusions on the reason for this perceived behavioral norm. However, having detected this perceived norm still helps understand the results: Just as in the social framing treatment, participants' decisions in the neutral framing treatment seems to be in line with what they express as their perceived behavioral norm.

Further research is necessary to investigate the unexpected negative correlation between endowments and the perceived norm to abstain from embezzlement in the neutral framing treatment. Future investigations may then also shed light on why higher endowments are related with a higher likelihood of embezzlement within the neutral framing treatment.

1.8   Conclusion

The present experimental study investigated whether receiving an initial endowment in a socially framed setting has an influence on subsequent immoral behavior in the form of embezzling donation money. We created a social framing treatment by letting half of the participants receive initial endowments as a share from previous participants, rather than by luck. Subsequently, we observed participants' decisions whether to manipulate a donation (entrusted to them by the experimenter) before forwarding it to a charity. By taking away part of the donation participants could enrich themselves, which we refer to as engaging in embezzlement.

We found a significant effect of the social framing on the likelihood of embezzlement, given that received endowments were low. Using data on perceived norms,

gathered during the experiment, we argued that self-deception may be the driving channel behind this social framing effect. With these results, we contribute to understanding the effects of social framing in immoral behavior. To the best of our knowledge, this is the first study to describe the effect of a social framing on the subsequent decision to engage in embezzlement that is harmful to a good cause, and therefore closely resembling real-life embezzlement.

Our central results are in line with existing literature on cheating, another form of immoral behavior. Future research will show, whether social framing generally has this effect on immoral behavior. The lesson to be learned for now is that people seem more willing to behave immorally when some other human's intent, rather than bad luck, is behind a recent negative experience. The most secure take-away is that there is no such thing as a generally favorable effect of socially framing an experience.

1.9 Appendix

On the following pages we display the instructions used in this experiment in the following order:

- The original instructions form the neutral framing treatment in German.

- An English translation of the instructions form the neutral framing treatment.

- The original instructions form the social framing treatment in German.

- An English translation of the instructions form the social framing treatment.
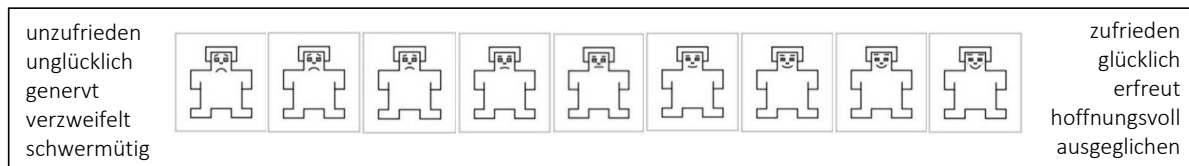
ANLEITUNG

Sie nehmen nun an einem wissenschaftlichen Experiment teil. Diese Anleitung erklärt Ihnen alles, was Sie für die Teilnahme wissen müssen. Falls Sie Fragen haben, melden Sie sich bitte per Handzeichen, ansonsten gilt während des gesamten Experiments ein absolutes Kommunikationsverbot.

Im Laufe des Experiments können Sie Geld verdienen. Alle Beträge werden direkt in Euro angegeben. Sie starten mit einem Guthaben von 0,00€. Am Ende des Experiments wird Ihnen Ihr Guthaben zusätzlich zum Teilnahmeentgelt von 2,50€ in bar ausgezahlt.

Stimmungsabfragen

Während des Experiments werden Sie mehrmals gebeten, Ihre Stimmungslage durch die Auswahl von Symbolen auszudrücken. Bitte geben Sie an, wie Sie sich im entsprechenden Moment fühlen, indem Sie eines der neun Symbole auswählen. Dabei stehen die äußeren Symbole für extreme und die Symbole dazwischen für entsprechend weniger eindeutige Stimmungslagen. **Es gibt keine richtigen oder falschen Antworten. Antworten Sie bitte ehrlich und nur Ihrer aktuellen Stimmungslage in dem Augenblick entsprechend.** Die Stimmungsabfragen haben keinerlei Auswirkung auf Ihre Auszahlungen.

| unzufrieden unglücklich genervt verzweifelt schwermütig | | | | | | | | | | zufrieden glücklich erfreut hoffnungsvoll ausgeglichen |

Ablauf des Experiments

Teil 1: Empfangen

In Teil 1 erhalten Sie eine Ihnen zufällig zugeloste Anfangsausstattung. Sie erfahren direkt die Höhe des Betrags. Dieser Betrag wird sofort Ihrem Guthaben hinzugefügt. Der Betrag kann eine Höhe von 0,00€, 1,00€, 2,00€, 3,00€, 4,00€ oder 5,00€ haben. In 55 von 100 Fällen hat dieser Betrag eine Höhe von 2,00€ oder weniger.

Teil 2: Spenden

Im Rahmen des heutigen Experiments soll pro Teilnehmer ein Betrag von 3,50€ an *Ärzte ohne Grenzen* gespendet werden. Die Spendengelder für diese Sitzung liegen gesammelt in einem Umschlag, der Ihnen im Rahmen der Auszahlung gezeigt wird.

In Teil 2 wird Ihnen der Spendenbetrag von 3,50€ angezeigt. Sie haben dann zwei Möglichkeiten:

Betrag bestätigen: Dann beträgt die Spende an *Ärzte ohne Grenzen* 3,50€, Ihr Guthaben bleibt unverändert.

Betrag vorab verringern: Sie können einen Teil des Geldes (0,20€; 0,30€; 0,40€; … bis maximal 3,50€) für sich selbst entnehmen. Die Spende an *Ärzte ohne Grenzen* wird entsprechend verringert. Dadurch entstehen Verwaltungskosten von 0,20€, da der gesamte Spendenbetrag dieser Sitzung angepasst werden muss. Diese 0,20€ gehen zurück an die Experimentatoren. Sie können also nicht weniger als 0,20€ entnehmen. Das Geld was Sie entnommen haben, abzüglich 0,20€, wird Ihrem persönlichen Guthaben hinzugefügt.

Der Umschlag mit den gesammelten Spendengeldern aller Teilnehmer liegt vorne bei den Experimentatoren. Nach Ende der Auszahlung werden die Spendengelder entsprechend der Entscheidungen aller Teilnehmer angepasst. Die angepassten Spendengelder werden innerhalb einer Woche an *Ärzte ohne Grenzen* überwiesen. Ein Teilnehmer wird vor der Auszahlung ausgelost, um den gesamten Spendenbetrag dieser Sitzung zu quittieren. Dieser Teilnehmer wird außerdem eingeladen, als Stellvertreter für diese Sitzung bei der Überweisung der Spende an *Ärzte ohne Grenzen* anwesend zu sein. Die Überweisungsbestätigung finden Sie eine Woche nach dem Experiment auf der Homepage des BaER-Lab (http://www.baer-lab.org). Damit wird für Sie sichergestellt, dass die Spendenbeträge tatsächlich bei *Ärzte ohne Grenzen* ankommen.

Teil 3: Schätzen

Ihnen wird angezeigt, wie viele Teilnehmer an der laufenden Sitzung teilnehmen. Dann sollen Sie einschätzen, wie viele dieser Teilnehmer den Spendenbetrag von 3,50€ in Teil 2 direkt bestätigt haben. Wenn Sie richtig liegen, wird Ihr Guthaben um 1,00€ erhöht.

Auszahlung

Die Auszahlungen für dieses Experiment setzen sich wie folgt zusammen:

- **Aus Teil 1:** Den Geldbetrag, den Sie in Teil 1 empfangen haben.
- **Aus Teil 2:** Das Geld was Sie der Spende entnommen haben, abzüglich 0,20€ (zwischen 0,00€ und 3,30€).
- **Aus Teil 3:** 1,00€ falls Sie mit Ihrer Einschätzung genau richtig lagen.
- **Teilnahmeentgelt:** 2,50€.

Nach dem Experiment

Im Anschluss an das Experiment werden Sie gebeten, noch einige Fragen zu beantworten. Bitte beantworten Sie alle Fragen ehrlich und vollständig. Ihre Antworten sind anonym und haben keinen Einfluss auf Ihre Auszahlung in diesem Experiment.

Bitte beachten Sie:
- Während des gesamten Experiments ist keine Kommunikation mit anderen Teilnehmern gestattet.
- Alle Handys müssen während der kompletten Experimentdauer ausgeschaltet sein.
- Wenn Sie eine Frage haben, bleiben Sie bitte an Ihrem Platz sitzen und heben die Hand. Stellen Sie bitte Ihre Frage so, dass kein anderer Teilnehmer Ihre Frage mithören kann.
- Sämtliche Entscheidungen erfolgen anonym, d.h. keiner der anderen Teilnehmer erfährt die Identität desjenigen, der eine bestimmte Entscheidung getroffen hat.
- Auch die Auszahlung erfolgt anonym, d.h. kein Teilnehmer erfährt, wie hoch die Auszahlung eines anderen Teilnehmers ist.
- Bitte bleiben Sie bis zum Ende des Experiments an Ihrem Platz sitzen, Sie werden zur Auszahlung mittels der Ihnen zugeordneten Platznummer aufgerufen.

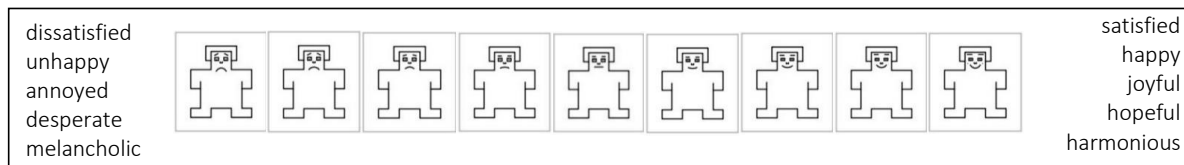**Viel Erfolg und Danke für Ihre Teilnahme an diesem Experiment!**

INSTRUCTIONS

You are now going to participate in an economic experiment. These instructions provide you with all relevant information regarding your participation. If you have any questions, please raise your hand. During the whole experiment, we ask you to remain silent and not talk to other participants.

During the experiment you can earn money. Amounts will directly be displayed in Euro. You start with a budget of 0.00€. After the experiment you will receive the money you earned during your participation, additional to a fixed show-up fee of 2.50€ in cash.

Happiness Ratings

During the experiment you will repeatedly be asked to rate your mood by selecting a suitable symbol. Please, indicate how you feel in that very moment, by choosing one of the nine symbols displayed below. Symbols on the extreme right or left represent rather extreme moods, while the ones towards the center represent more ambiguous moods. **There are no right or wrong answers. Please be honest and answer only according to your mood at that very moment.** Your mood ratings will not affect your payment in any way.

| dissatisfied unhappy annoyed desperate melancholic | | satisfied happy joyful hopeful harmonious |

Sequence of the Experiment

Part 1: Receiving

In Part 1 you receive an initial endowment, randomly selected for you. You will instantly be informed about the height of this amount of money, and the money will instantly be added to your experimental account. The amounts distributed as initial endowments are of different heights (either 0.00€, 1.00€, 2.00€, 3.00€, 4.00€, 5.00€). In 55 out of 100 cases, the amount is 2.00€ or less.

Part 2: Forwarding Donations

As part of today's experiment, a donation of 3.50€ per participant is supposed to be donated to *Médecins Sans Frontières*. The donation money for the current session of the experiment is collected in one envelope, which will be shown to you after the experiment when you receive your payment.

In Part 2, an donation amount of 3.50€ will be displayed on your screen. You have two options:

Confirming the amount: The donation to *Médecins Sans Frontières* will then be 3.50€, and your personal experimental account remains unchanged.

Prior reduction of the amount: You can take part of the amount to be donated (0.20€, 0.30€, 0.40€, … up to max. 3.50€) for yourself. The donation to *Médecins Sans Frontières* will then be reduced accordingly. Doing so leads to additional administrative costs of 0.20€, since the total amount to be donated after this session needs to be adjusted. Therefore, you cannot take less than 0.20€ from the donation. The amount you take from the donation, minus 0.20€, will be added to your personal experimental account.

The envelope containing the collected amounts to be donated from all participants lies in the front of the laboratory. After the experiment, all amounts will be adjusted according to the participants' decisions. The adjusted amounts to be donated will be transferred to *Médecins Sans Frontières* within one week. One randomly selected participant will be asked to acknowledge the collected amounts to be donated from this session. This participant will also be invited to witness the transferal of the donations to *Médecins Sans Frontières*, acting as a representative of the participants of this session. One week from today, you will be able to find a receipt of the transfers on the homepage of the BaER-Lab (http://www.baer-lab.org), so you can check for yourself that the donations actually did arrive at *Médecins Sans Frontières*.

Part 3: Guessing

You will be shown the number of participants of the current session. Next, you are asked to guess how many of these participants decided to immediately confirm the amount of 3.50€ to be donated in Part 2. If your guess is correct, you receive an additional 1.00€.

Payment

Your payment for participating in the current experiment consists of the following components::

- **From Part 1:** The amount of money you received in Part 1.

- **From Part 2:** The amount you took from the donation, minus 0.20€ (hence, between 0.00€ and 3.30€).

- **From Part 3:** 1.00€ in case of a correct guess.

- **Show-up fee:** 2.50€.

After the Experiment

After the experiment you will be asked to answer a post-experimental questionnaire. Please, do give honest and complete answers to all questions. Your answers are anonymous and will not affect your payment in any way.

Important remarks:
- Throughout the experiment you may not communicate with other participants.
- Cellphones must be turned off during the experiment.
- In case of any questions, please raise your hand and remain seated. Please ask your question in a way that does not allow other participants to overhear them.
- All of your decisions are anonymous. Consequently, none of the other participants will learn the identity of anyone who made a particular decision during the experiment.
- Payment is anonymous; hence, the participants will not observe the payment of any other participant.
- Please remain seated until the end of the experiment. You will be called by your seat number to receive your payment.

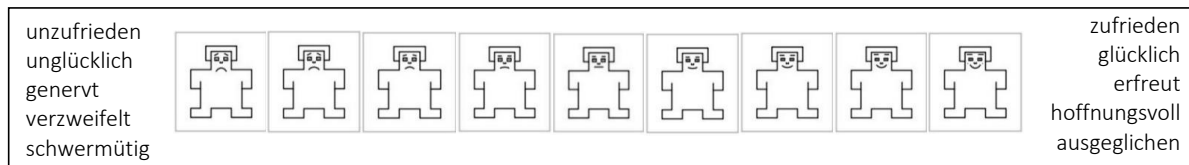**Lots of success and thank you for your participation!**

ANLEITUNG

Sie nehmen nun an einem wissenschaftlichen Experiment teil. Diese Anleitung erklärt Ihnen alles, was Sie für die Teilnahme wissen müssen. Falls Sie Fragen haben, melden Sie sich bitte per Handzeichen, ansonsten gilt während des gesamten Experiments ein absolutes Kommunikationsverbot.

Im Laufe des Experiments können Sie Geld verdienen. Alle Beträge werden direkt in Euro angegeben. Sie starten mit einem Guthaben von 0,00€. Am Ende des Experiments wird Ihnen Ihr Guthaben zusätzlich zum Teilnahmeentgelt von 2,50€ in bar ausgezahlt.

Stimmungsabfragen

Während des Experiments werden Sie mehrmals gebeten, Ihre Stimmungslage durch die Auswahl von Symbolen auszudrücken. Bitte geben Sie an, wie Sie sich im entsprechenden Moment fühlen, indem Sie eines der neun Symbole auswählen. Dabei stehen die äußeren Symbole für extreme und die Symbole dazwischen für entsprechend weniger eindeutige Stimmungslagen. **Es gibt keine richtigen oder falschen Antworten. Antworten Sie bitte ehrlich und nur Ihrer aktuellen Stimmungslage in dem Augenblick entsprechend.** Die Stimmungsabfragen haben keinerlei Auswirkung auf Ihre Auszahlungen.

| unzufrieden unglücklich genervt verzweifelt schwermütig | | | | | | | | | | zufrieden glücklich erfreut hoffnungsvoll ausgeglichen |

Ablauf des Experiments

Teil 1: Empfangen

Im Rahmen von Sitzungen eines früheren Experiments, welches von Mai bis Juni 2017 im BaER-Lab stattgefunden hat, konnten Teilnehmer 10,00€ aufteilen: Diese *früheren Teilnehmer* konnten 0,00€, 1,00€, 2,00€, 3,00€, 4,00€, 5,00€, 6,00€, 7,00€, 8,00€, 9,00€ oder 10,00€ an einen *zukünftigen Teilnehmer* eines anderen Experiments abgeben, und haben den Restbetrag für sich selbst behalten. In 55 von 100 Fällen haben *frühere Teilnehmer* 2,00€ oder weniger abgegeben. *Frühere Teilnehmer* haben nie mehr als 5,00€ abgegeben.

Sie sind einer dieser *zukünftigen Teilnehmer* und wurden einem *früheren Teilnehmer* zugelost.

In Teil 1 erhalten Sie den Betrag, den der *frühere Teilnehmer* Ihnen überlassen hat, und erfahren direkt die Höhe des Betrags. Dieser Betrag wird sofort Ihrem Guthaben hinzugefügt.

Teil 2: Spenden

Im Rahmen des heutigen Experiments soll pro Teilnehmer ein Betrag von 3,50€ an *Ärzte ohne Grenzen* gespendet werden. Die Spendengelder für diese Sitzung liegen gesammelt in einem Umschlag, der Ihnen im Rahmen der Auszahlung gezeigt wird.

In Teil 2 wird Ihnen der Spendenbetrag von 3,50€ angezeigt. Sie haben dann zwei Möglichkeiten:

Betrag bestätigen: Dann beträgt die Spende an *Ärzte ohne Grenzen* 3,50€, Ihr Guthaben bleibt unverändert.

Betrag vorab verringern: Sie können einen Teil des Geldes (0,20€; 0,30€; 0,40€; … bis maximal 3,50€) für sich selbst entnehmen. Die Spende an *Ärzte ohne Grenzen* wird entsprechend verringert. Dadurch entstehen Verwaltungskosten von 0,20€, da der gesamte Spendenbetrag dieser Sitzung angepasst werden muss. Diese 0,20€ gehen zurück an die Experimentatoren. Sie können also nicht weniger als 0,20€ entnehmen. Das Geld was Sie entnommen haben, abzüglich 0,20€, wird Ihrem persönlichen Guthaben hinzugefügt.

Der Umschlag mit den gesammelten Spendengeldern aller Teilnehmer liegt vorne bei den Experimentatoren. Nach Ende der Auszahlung werden die Spendengelder entsprechend der Entscheidungen aller Teilnehmer angepasst. Die angepassten Spendengelder werden innerhalb einer Woche an *Ärzte ohne Grenzen* überwiesen. Ein Teilnehmer wird vor der Auszahlung ausgelost, um den gesamten Spendenbetrag dieser Sitzung zu quittieren. Dieser Teilnehmer wird außerdem eingeladen, als Stellvertreter für diese Sitzung bei der Überweisung der Spende an *Ärzte ohne Grenzen* anwesend zu sein. Die Überweisungsbestätigung finden Sie eine Woche nach dem Experiment auf der Homepage des BaER-Lab (http://www.baer-lab.org). Damit wird für Sie sichergestellt, dass die Spendenbeträge tatsächlich bei *Ärzte ohne Grenzen* ankommen.

Teil 3: Schätzen

Ihnen wird angezeigt, wie viele Teilnehmer an der laufenden Sitzung teilnehmen. Dann sollen Sie einschätzen, wie viele dieser Teilnehmer den Spendenbetrag von 3,50€ in Teil 2 direkt bestätigt haben. Wenn Sie richtig liegen, wird Ihr Guthaben um 1,00€ erhöht.

Auszahlung

Die Auszahlungen für dieses Experiment setzen sich wie folgt zusammen:

- **Aus Teil 1:** Den Geldbetrag, den Sie in Teil 1 empfangen haben.

- **Aus Teil 2:** Das Geld was Sie der Spende entnommen haben, abzüglich 0,20€ (zwischen 0,00€ und 3,30€).

- **Aus Teil 3:** 1,00€ falls Sie mit Ihrer Einschätzung genau richtig lagen.

- **Teilnahmeentgelt:** 2,50€.

Nach dem Experiment

Im Anschluss an das Experiment werden Sie gebeten, noch einige Fragen zu beantworten. Bitte beantworten Sie alle Fragen ehrlich und vollständig. Ihre Antworten sind anonym und haben keinen Einfluss auf Ihre Auszahlung in diesem Experiment.

Bitte beachten Sie:

- Während des gesamten Experiments ist keine Kommunikation mit anderen Teilnehmern gestattet.
- Alle Handys müssen während der kompletten Experimentdauer ausgeschaltet sein.
- Wenn Sie eine Frage haben, bleiben Sie bitte an Ihrem Platz sitzen und heben die Hand. Stellen Sie bitte Ihre Frage so, dass kein anderer Teilnehmer Ihre Frage mithören kann.
- Sämtliche Entscheidungen erfolgen anonym, d.h. keiner der anderen Teilnehmer erfährt die Identität desjenigen, der eine bestimmte Entscheidung getroffen hat.
- Auch die Auszahlung erfolgt anonym, d.h. kein Teilnehmer erfährt, wie hoch die Auszahlung eines anderen Teilnehmers ist.
- Bitte bleiben Sie bis zum Ende des Experiments an Ihrem Platz sitzen, Sie werden zur Auszahlung mittels der Ihnen zugeordneten Platznummer aufgerufen.

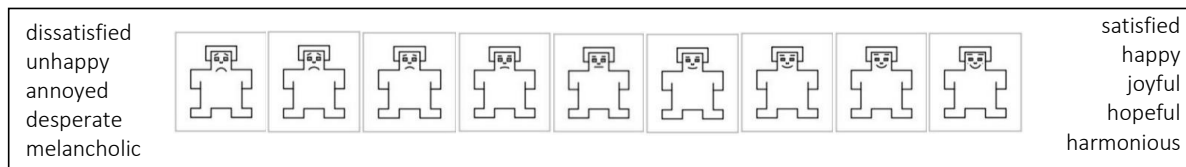**Viel Erfolg und Danke für Ihre Teilnahme an diesem Experiment!**

INSTRUCTIONS

You are now going to participate in an economic experiment. These instructions provide you with all relevant information regarding your participation. If you have any questions, please raise your hand. During the whole experiment, we ask you to remain silent and not talk to other participants.

During the experiment you can earn money. Amounts will directly be displayed in Euro. You start with a budget of 0.00€. After the experiment you will receive the money you earned during your participation, additional to a fixed show-up fee of 2.50€ in cash.

Happiness Ratings

During the experiment you will repeatedly be asked to rate your mood by selecting a suitable symbol. Please, indicate how you feel in that very moment, by choosing one of the nine symbols displayed below. Symbols on the extreme right or left represent rather extreme moods, while the ones towards the center represent more ambiguous moods. **There are no right or wrong answers. Please be honest and answer only according to your mood at that very moment.** Your mood ratings will not affect your payment in any way.

Sequence of the Experiment

Part 1: Receiving

Participants of a different experiment, which took place in May and June 2017 in the BaER-Lab, were asked to split up 10.00€: These *previous participants* could give part of the money (either 0.00€, 1.00€, 2.00€, 3.00€, 4.00€, 5.00€, 6.00€, 7.00€, 8.00€, 9.00€, or 10.00€) to a *future participant* of a different experiment, while keeping the leftover money for themselves. In 55 out of 100 cases, these *previous participants* gave away 2.00€ or less. None of the *previous participants* gave more than 5.00€.

You are one of the *future participants* and you were randomly assigned to one of the *previous participants*.

In Part 1, you receive the amount of money that this *previous participant* gave away to you. You will instantly be informed about the height of this amount of money, and the money will instantly be added to your experimental account.

Part 2: Forwarding Donations

As part of today's experiment, a donation of 3.50€ per participant is supposed to be donated to *Médecins Sans Frontières*. The donation money for the current session of the experiment is collected in one envelope, which will be shown to you after the experiment when you receive your payment.

In Part 2, an donation amount of 3.50€ will be displayed on your screen. You have two options:

Confirming the amount: The donation to *Médecins Sans Frontières* will then be 3.50€, and your personal experimental account remains unchanged.

Prior reduction of the amount: You can take part of the amount to be donated (0.20€, 0.30€, 0.40€, … up to max. 3.50€) for yourself. The donation to *Médecins Sans Frontières* will then be reduced accordingly. Doing so leads to additional administrative costs of 0.20€, since the total amount to be donated after this session needs to be

41

adjusted. Therefore, you cannot take less than 0.20€ from the donation. The amount you take from the donation, minus 0.20€, will be added to your personal experimental account.

The envelope containing the collected amounts to be donated from all participants lies in the front of the laboratory. After the experiment, all amounts will be adjusted according to the participants' decisions. The adjusted amounts to be donated will be transferred to *Médecins Sans Frontières* within one week. One randomly selected participant will be asked to acknowledge the collected amounts to be donated from this session. This participant will also be invited to witness the transferal of the donations to *Médecins Sans Frontières*, acting as a representative of the participants of this session. One week from today, you will be able to find a receipt of the transfers on the homepage of the BaER-Lab (http://www.baer-lab.org), so you can check for yourself that the donations actually did arrive at *Médecins Sans Frontières*.

Part 3: Guessing

You will be shown the number of participants of the current session. Next, you are asked to guess how many of these participants decided to immediately confirm the amount of 3.50€ to be donated in Part 2. If your guess is correct, you receive an additional 1.00€.

Payment

Your payment for participating in the current experiment consists of the following components::

- **From Part 1:** The amount of money you received in Part 1.

- **From Part 2:** The amount you took from the donation, minus 0.20€ (hence, between 0.00€ and 3.30€).

- **From Part 3:** 1.00€ in case of a correct guess.

- **Show-up fee:** 2.50€.

After the Experiment

After the experiment you will be asked to answer a post-experimental questionnaire. Please, do give honest and complete answers to all questions. Your answers are anonymous and will not affect your payment in any way.

Important remarks:
- Throughout the experiment you may not communicate with other participants.
- Cellphones must be turned off during the experiment.
- In case of any questions, please raise your hand and remain seated. Please ask your question in a way that does not allow other participants to overhear them.
- All of your decisions are anonymous. Consequently, none of the other participants will learn the identity of anyone who made a particular decision during the experiment.
- Payment is anonymous; hence, the participants will not observe the payment of any other participant.
- Please remain seated until the end of the experiment. You will be called by your seat number to receive your payment.

Lots of success and thank you for your participation!

Chapter 2

Revisiting a remedy against the chain of unkindness

Wendelin Schnedler
Paderborn University

Nina Lucia Stephan
Paderborn University

2.1 Introduction

People who have been treated unkindly may treat others unkindly. Consider a manager whose supervisor just ordered her to work overtime. Perhaps she will afterwards have to pull herself together to avoid hurling abuse at a colleague. Passing on of unkind behavior may create a bad atmosphere and ultimately reduce productivity. Several experiments document that unkind behavior gets passed on (Diekmann, 2004; Gray et al., 2014; Leimgruber et al., 2014), a phenomenon also referred to as negative generalized reciprocity (Nowak and Sigmund, 2005; Stanca, 2009).

In a recent contribution, Strang et al. (2016) suggest a remedy against the chain of unkindness: after having received little money from a subject A, the recipients of the unkind behavior, say B, are allowed to write a letter to A. These 'letter writers' later pass on much more of a common pot of money to an innocent third subject, say C, than recipients of unkind behavior who could not write a letter. This suggests that letter writing helps subjects to 'close the case' and separate the person who harmed them (A) from the person who they are about to harm (C). In other words: letter writing helps them cooling down.

In the experimental setting, A and B both have to decide on how much money to share with another person. In contrast to that, real life examples of chains of unkindness often involve decisions that are not of the same type. The manager can decide whether to hurl abuse and her supervisor whether to impose overtime work. If letter writing actually 'closes the case', it should work even if the decisions of A and B are of a different type.

There is, however, a plausible reason why letter writing might only have the effect observed by Strang et al. (2016) when decisions are of the same type. Suppose that B has the intention to imitate A. Being given the chance to write a letter to A may change B's intention to imitate A. In particular, being given the opportunity to write a letter may be seen by B as an indication that A's behavior was not adequate. As a result, B may be less likely to imitate A. Imitation would then be less likely for subjects who have the opportunity to write a letter than for those who do not. Accordingly, this group would pass on more money, which would explain the outcome observed by Strang et al. (2016). This argument only applies as long as A and B face the same type of decision. Otherweise, the opportunity to write a letter cannot signal to B that A should not be imitated and it would be wrong to conclude that letter writing helps against the chain of unkindness.

In this paper, we report on an experiment that re-examines whether letter writing interrupts the chain of unkindness by considering a setting where A and B's decisions are taken in very different domains. We find that nevertheless letter writing has the intended effect. From 10 € received, the amount B passes on to C significantly

increases from 2.12 € to 2.76 € for subjects (B) who have the opportunity to first write a letter to A. The size of the effect is similar to that found by (Strang et al., 2016) when there is not change of the decision domain.

The key innovation in our experimental design is that we prevent B from imitating A by confronting A and B with choices from very different domains. The difference is very salient as A's choice has non-monetary and B's choice monetary effects. More precisely, A decides whether B is allowed to watch funny movies or has to engage in an annoying encryption task instead. We ensure that subjects in the role of A assign enough subjects in the role of B to the annoying task, by paying them 10 € whenever they assign half of the subjects to the annoying task. The decision of B, on the other hand, is the dictator game decision also used by Strang et al.: subjects B can decide how much of a sum of 10 € they want to allocate to a participant of some other future experiment (subject C).

There is vast evidence from a variety of disciplines, such as psychology (Bertenthal et al., 2006; Delfino et al., 2016; Heyes, 2011; Leighton et al., 2010), neuroscience and biology (Byrne and Russon, 1998; Bien et al., 2009), or economics (Alos-Ferrer and Schlag, 2009; Apesteguia et al., 2007; Conlisk, 1980; Day, 1984; Matthey, 2010; Pingle, 1995), showing that imitation is common. It is thus very plausible that subjects try to imitate the dictator rather than to pass on unkind behavior.

Here, we exploit that the passing on of unkindness does not have to occur within one domain, to investigate the passing on of unkindness, or generalized negative reciprocity, in a setting where imitation cannot play a role. To the best of our knowledge, this distinguishes our work not only from Strang et al. (2016), but also from any other experimental studies on the chain of unkindness or negative generalized reciprocity, where imitation is generally possible, such as Ben-Ner et al. (2004) or Diekmann (2004). In theoretical models of reciprocity, such as Gouldner (1960) or more recently Falk and Fischbacher (2006) reciprocity is the rewarding of a kind or punishing of an unkind action. Nothing suggests that the rewarding or punishing needs to take place in the same domain as the initial kind or unkind action. In fact, existing literature has called for using more realistic decision situations, which our design is aiming at, and even turned to study the behavior in the field and within non-monetary domains (Gray et al., 2014; Mujcic and Leibbrandt, 2017; van Apeldoorn and Schram, 2016).

We thus contribute experimental evidence to the following central question: Does letting an actor write a letter to the initial unkind actor remedy a chain of unkindness that is crossing domains? More specifically, we will show that when given the chance to write a letter, subjects B pass on significantly more money to subject C, even in a setting with a domain change.

As the passing on of unkindness is a very general concept, frequently observed in anyone's real life, existing literature on the topic is not restricted to any specific

group of subjects (Ben-Ner et al., 2004; Diekmann, 2004; Gray et al., 2014; Herne et al., 2013; Leimgruber et al., 2014; Mujcic and Leibbrandt, 2017; Stanca, 2009). In contrast to most literature, Strang et al. (2016) solely investigate behavior of female subjects. To underline the robustness of our results, we thus investigate whether the overall effect we find is the same within the subgroup of our female subjects. Confirming the robustness of our overall results, we do not find a gender specific effect of letter writing on the passing on of money.
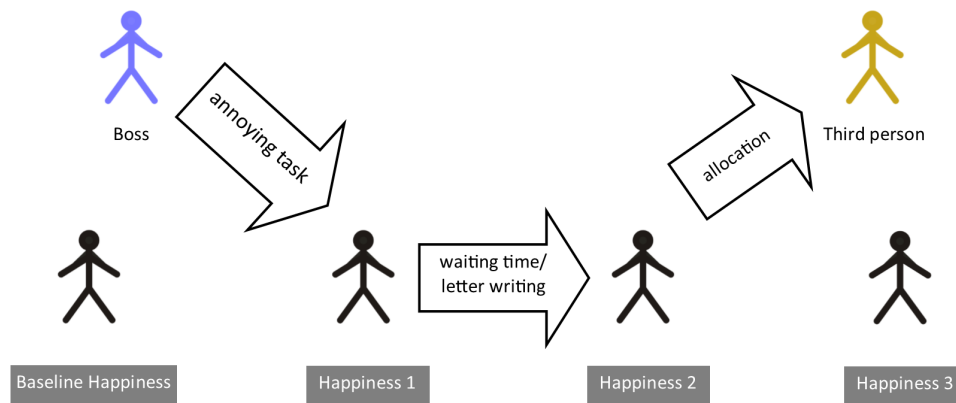
2.2 Method

Figure 2.1: Experimental Design: Participants were assigned (by the boss) to participate in an annoying, instead of an enjoyable task. Next, they went through a waiting time (control group) or letter writing (treatment group) task. Subsequently, they made allocations to a third person as dictators in a dictator game.

All participants within the same treatment receive the same set of instructions.[1] After checking for their understanding, roles are assigned. In each session, one participant is randomly chosen as the boss. The remaining participants are randomly put into pairs, where one is called player X and one player Y. Inititally X and Y are assigned to an enjoyable default task. The boss, however, can receive a payment of 10 € if he assigns half of the participants to an annoying task and exposes them to constant reminders to hurry up instead of letting them do the enjoyable default task. If deciding to do so,[2] the boss has to determine whom in each pair should carry out

---

[1]Original instructions, as well as an English translation are displayed in the Appendix.

[2]In all session, the boss did indeed decide to annoy subjects in in order to earn the 10 €.

the annoying task: X or Y.

As a result, participants are either assigned to carry out an enjoyable task (as raters) or an annoying task (as encrypters), depending solely on the boss's decision. Without the existence of such an alternative default task, an assignment to the encryption task would not be perceived as unkind. This is why an alternative and relatively more enjoyable task, as well as a real possibility to be assigned to it, must exist, even though for our analysis we only need the data from the encrypters' decisions. All encrypters are aware that the boss is responsible for the task assignment and that the other half of the participants are assigned to a more enjoyable task. The tasks, both rewarded by a small, symbolic fixed payment, start right thereafter during a work phase, and are described in the following.

While the raters are watching and rating funny videos, the encrypters have to encrypt a combination of numbers into a random combination of letters (password). Entering the correct password on the first screen leads to a new encryption task on the next screen. According to Benndorf et al. (2018) encrypting of this type only allows for minimal learning, which reduces the chances that participants might enjoy their work as they get better at it. Moreover, the encrypters are repeatedly reminded to speed up by a reminder box popping up and blocking their screens, which leads to their latest password entry being deleted. The work phase is controlled to be of equal length in each session.

The next part of the experiment depends on the treatment. In the letter writing treatment, subjects are given three minutes to write a letter to the boss. This letter is later by us and handed to the boss after the experiment. In the control treatment, subjects simply wait for three minutes.

Finally, encrypters and raters act as dictators in a dictator game, splitting up $10 \, €$ between themselves and a participant of a future, unrelated experiment. Dictators know that none of the participants of the current experiment will be invited to participate in this future experiment and that this experiment will also take place in the BaER-Lab at Paderborn University. [3]

In order to check whether the task assignment worked as intended, as well as to survey how participants' emotions evolve throughout the experiment and within the different treatments, we measure the participants' subjective level of happiness. This is done by letting participants rate their own happiness on a 9-point Likert scale, using *Self-Assessment Manikins* as described by Bradley and Lang (1994). Happiness is measured at the beginning of the experiment, before assigning tasks (baseline happiness), as well as after the work phase (happiness 1), after the letter writing or waiting (happiness 2) and after the dictator game (happiness 3).

---

[3]As announced in the instructions, all money from the 308 dictators (including raters) was passed on to participants who did not take part in the present experiment. The sessions, where these receivers did participate, took place in July and September 2017, as part of different, unrelated experiments.

In line with Strang et al. (2016), we expect participants in the letter writing treatment to allocate more to a third person in a subsequent dictator game. If the reason for the letter writing effect in Strang et al.'s experiment is that it signals subjects not to imitate, however, we would expect that the letter writing should not affect allocations.

2.3   Results

This study was programmed using z-Tree (Fischbacher, 2007) and carried out at the Business and Economics Research Laboratory (BaER-Lab) of the University of Paderborn, Germany, in May and June 2017. Participants were recruited at the University, using the software ORSEE (Greiner, 2015). In total, 154 subjects were assigned to encrypt (Table 2.1).

	Total	Control Treatment	Letter Writing Treatment
observations	154	99	55
average age (min:18; max:35)	21.98	23.21	21.98
female participants (%)	67.53	64.65	72.73

Table 2.1: Descriptive statistics show a successful randomization.

Figure 2.2 shows that the encrytption task actually reduced happiness. In both, the control group (waiting time) and the treatment group (letter writing), the median happiness ratings drop from 6 to 4 (on a 1-9 scale) as shown in Table 2.2. A Wilcoxon-signed-rank-test confirms that happiness ratings before and after the encryption task differ significantly ($p < 0.01$, $z = -7.954$). Even though the encryption task does not go along with a monetary loss, happiness is affected in the same way and as strongly as in Strang et al. (2016). In their experiment, where the first unkind action consisted in receiving an unfairly low amount of money from a dictator, median happiness also decreased from 6 to 4. This confirms that the unkind experience in the non-monetary domain in our setup, is comparable to the unkind experience in the monetary domain in the setup of Strang et al. (2016).
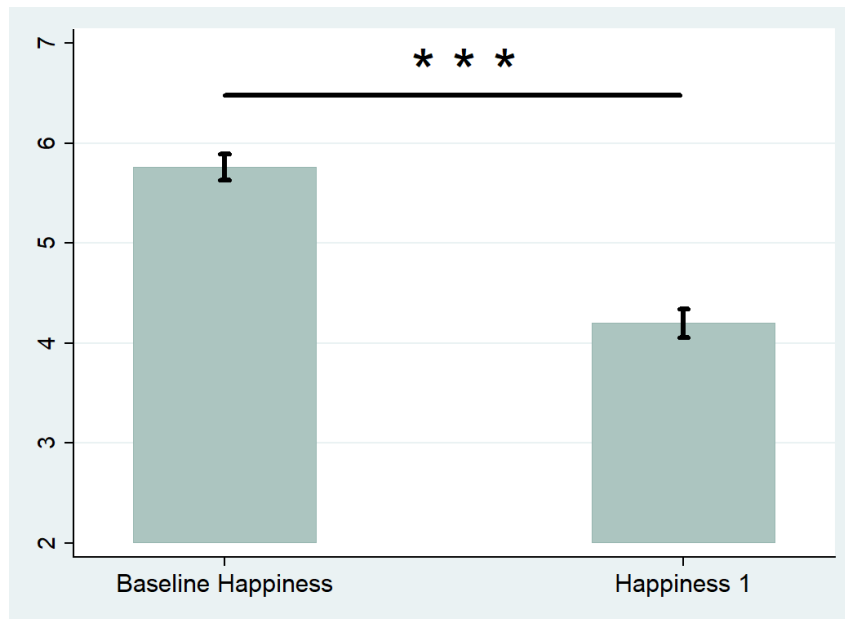
Figure 2.2

Figure 2.2: The mean values of encrypters' baseline happiness and happiness 1 confirm that reported happiness is significantly decreased after the work phase, and that the chosen task is actually annoying. (* p<0.05, ** p<0.01, *** p<0.001; error bars indicate standard error of mean)

	Total	Control	Letter Writing
baseline happiness	5.760 [6]	5.788 [6]	5.709 [6]
happiness 1	4.201 [4]	4.141 [4]	4.309 [4]
allocations	2.352 [2.5]	2.121 [2]	2.764 [3]

Table 2.2: Mean happiness and allocation (median values in brackets). In both groups, encrypters' happiness is affected negatively by the work phase. Neither baseline happiness nor happiness 1 levels differ significantly between the groups. However, allocations are larger in the letter writing group.
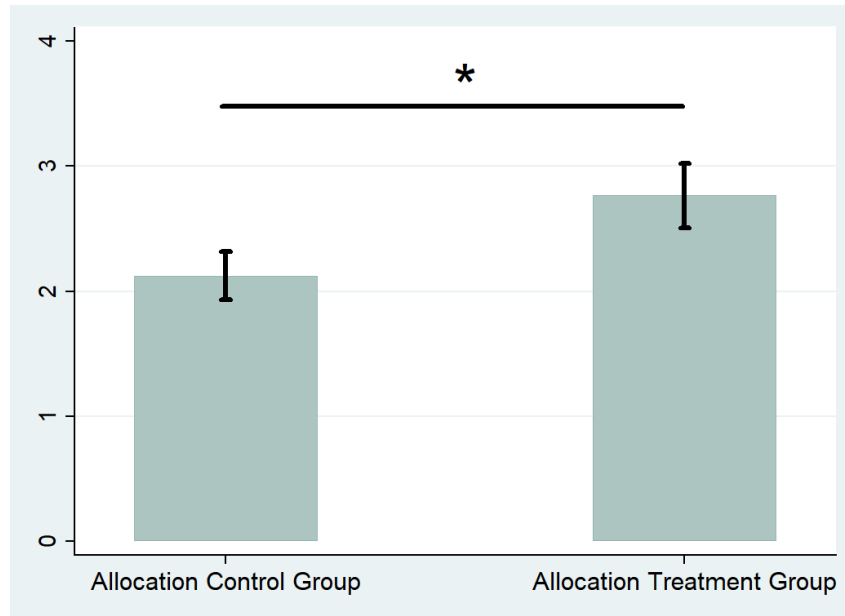
2.3.1 Main Result

Figure 2.3

Figure 2.3: Averages show that allocations are higher in the letter writing group, and hence, letter writing increases allocations. (* p<0.05, ** p<0.01, *** p<0.001; error bars indicate standard error of mean)

Next, we compare the effect of letter writing (treatment group) to that of waiting (control group). As shown in Figure 2.3, allocations are significantly higher in the treatment than in the control group (treatment group mean 2.764, median 3, control group mean 2.121, median 2). In other words, and in line with previous literature, letting annoyed participants write a letter to a responsible person makes them give more in a subsequent dictator game (p-value <0.05), even if the annoying task is non-monetary.

Result 1: *Subjects who were assigned to the annoying encryption task give significantly more to an independent third subject if they had the opportunity to first write a letter.*

2.3.2 Robustness of Results: Subgroup of Female Participants

In their experiment Strang et al. (2016) restrict participation to females, arguing that women typically have stronger emotional reactions. This is, for example, suggested by Ben-Ner et al. (2004); Bradley et al. (2001); Grossman and Wood (1993). Our experiment and results so far are not restricted to women but include data from both male and female subjects. The present section will show that our central findings are not gender specific. By looking at the subgroup of female subjects, we also enable direct comparison with Strang et al. (2016).

	Total	Control Treatment	Letter Writing Treatment
baseline happiness	5.817 [6]	5.875 [6]	5.725 [6]
happiness 1	4.115 [4]	3.969 [4]	4.350 [4]
allocation	2.519 [3]	2.219 [2]	3.000 [3]

Table 2.3: Average choices by female encrypters are displayed in the same way as in Table 2.2. Evidently, also among females, allocations are higher in the letter writing group, confirming the treatment effect.

As shown in Table 2.3, the results from the subgroup of female participants are very close to the overall results (Table 2.2). The encryption task decreases median levels of happiness from 6 to 4 (significant difference with p-value <0.01), just as in the complete sample, and just as reported by Strang et al. (2016). Regarding the effect of the remedy of letting participants write a letter to the boss, the results are again similar to those of the complete sample. In direct comparison, allocations by female dictators are slightly higher. More importantly though, the effect of allocations being larger in the treatment group is confirmed within the group of female subjects (p-value <0.05, z= -2.079). This result is also in line with Strang et al., 2016, p. 4, who find significantly larger allocations after letter writing (median 4.5) compared to waiting time (median 3). We conclude:

Result 2: *The positive effect of letting annoyed subjects write a letter, to increase their willingness to give in a subsequent dictator game also holds for female subjects.*
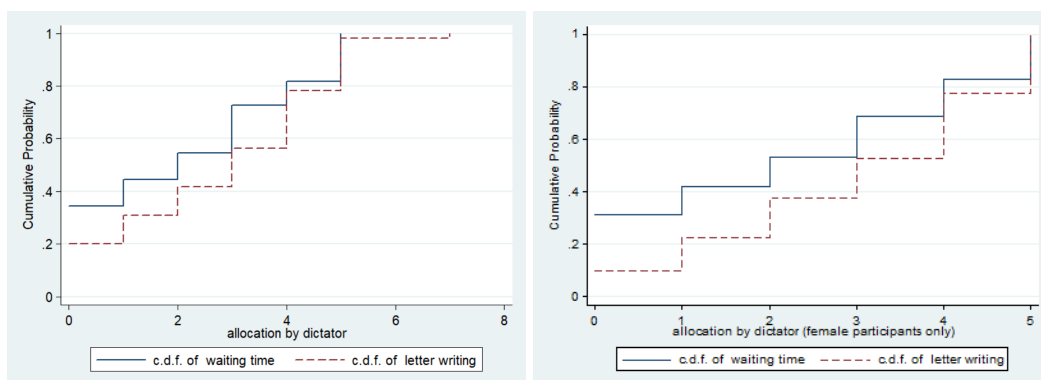
Figure 2.4: Allocations by Male and Female Dictators

Figure 2.5: Allocations by Female Dictators

Figure 2.4 & 2.5: C.d.f.s of dictators' allocations for the entire sample and subsample of females are displayed. Both figures show the positive effect of letter writing on allocations.

To visualize the results, Figures 2.4 and 2.5 show cumulative distribution functions (c.d.f.s) of the dictators' allocations to the third, previously uninvolved persons. As the c.d.f.s show, allocations are typically higher in the treatment group, that is in those cases where dictators have the chance to write a letter to the responsible boss, compared to the control group. This is shown for both, the entire sample and the subsample of female participants, confirming both, Result 1 and Result 2. As a side note it can be added that high allocations seem more likely within the subgroup of female dictators, irrespective of the treatment.

2.3.3   Ancillary Findings and Discussion

We have established the overall effect of letter writing on the passing on of unkindness across domains. Letter writing is confirmed to have the intended effect, which is in line with the 'closing the case' explanation. Yet, there is one caveat. As Strang et al. (2016) explain, letter writing is effective since it neutralizes the previously experienced negative emotional shock (closing the case), which in turn leads to higher allocations in a subsequent dictator game with a third person. In line with that explanation letter writing should lead to a stronger increase in happiness ratings than waiting. This, however, is not what we find. Instead, our data show that both letter writing and waiting time lead to slightly higher levels of happiness, with no evident difference in the change in happiness ratings between the two groups. In both groups, people just become slightly less emotional. This is the case for both male and female subjects.

Result 3: *The regulating measure of letter writing, compared to waiting time, has no significant effect on happiness ratings.*
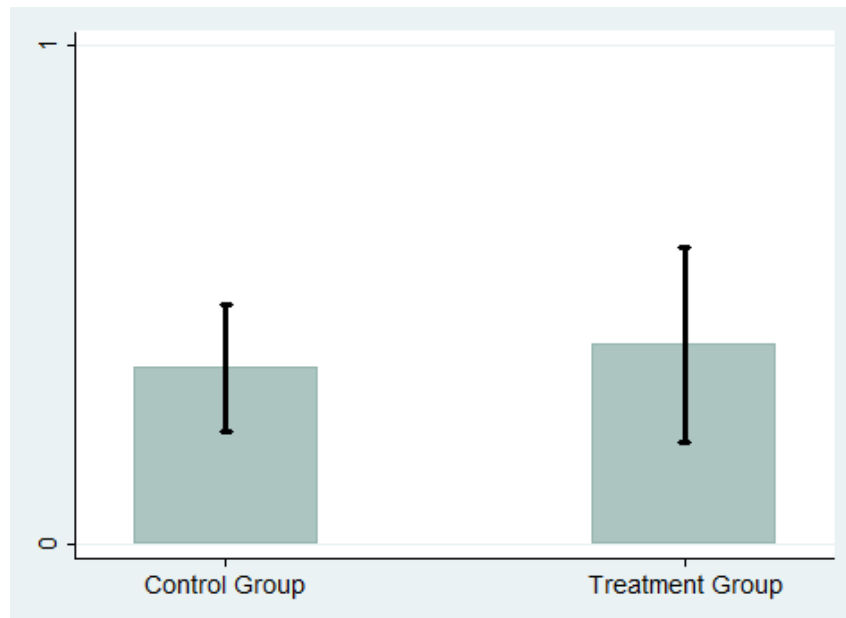

Figure 2.6

Figure 2.6: The change in happiness throughout the treatment phase shows that waiting time has almost no effect (increases of less than 0.5 unitson the 9-point Likert scale) on happiness, and neither does letter writing.

Figure 2.6 illustrates this result, showing the change in happiness ratings throughout the treatment phase. There is no significant difference between the effect of letter writing and waiting time. Moreover, the change in happiness is on average less than 0.5 points, and thus very small compared to the change from median happiness of 6 to 4 points after the annoying task. Hence, unlike expected, letter writing does not bring happiness back towards its baseline level. As the change in happiness, shown in Figure 2.6, is so small, it can actually not be distinguished from emotions simply cooling off over time (Oechssler et al., 2015).

Let us consider this puzzle again: Experiencing waiting time or the letter writing treatment both leads to slightly higher happiness with no evident difference in the change in happiness between the two groups. Writing a letter and spending the same time waiting both makes people just slightly less emotional, if at all. Happiness ratings can thus not explain significant differences in subsequent dictator game allocations. Looking for an explanation, we will now consider the content of the

letters, presenting results of a content element analysis as performed by Strang et al. (2016). Following their example, the letters were scanned for five specific content elements. For the subsequent analysis we consider correlations of the content elements and our variables of interest.

We observe that content elements related to emotional aspects like *expression of emotions* (0.2325; p<0.05), *unfairness criticism* (0.3579; p<0.01), and *questioning of motive* (0.3742; p<0.01) are significantly and positively correlated with a change in happiness, while there is no significant correlation with either *understanding* or *suggestion for usage*. Thus, those participants who make use of the letter to express their emotions, criticize the unfair behavior of the boss, or question his motive, are likely the same as those whose happiness ratings increase through letter writing. In contrast, these participants are rarely the same as those on whom the letter writing treatment has the effect of making them more willing to give. This is supported by the fact that the content element *expression of emotions*), positively correlated with the change in happiness as stated above, is at the same time negatively correlation with dictator allocations (-0.2048*; p<0.05). Moreover, those content elements showing the strongest correlation with a change in happiness are not correlated with dictator allocations at all. The missing correlation between changes in happiness and dictator allocations is also surprising, considering literature showing an effect of induced mood on dictator allocations or similar games (Capra, 2004; Kirchsteiger et al., 2006; Tan and Forgas, 2010). In contrast, our observations suggest that while some participants, by choosing certain content elements, use the letters to improve their happiness, there is another subgroup of participants for whom the letter writing treatment has the effect of increasing their willingness to give in a subsequent dictator game. With this puzzle, it is at this point not possible to draw final conclusions on the channel through which the letter writing treatment works.

2.4   Conclusion

We designed a controlled laboratory experiment where participants are treated badly by another participant, the boss, who assigns them an annoying effort demanding task. Previous research suggests that letting these participants write a letter to the boss on average leads to higher giving to an uninvolved third participant. Letter writing is thus proposed as a way to stop the passing on of unkindness, however, from previous literature it cannot be concluded that whether this is actually the case. With a novel experimental design, we confirm that letter writing has the intended effect. In order to do so, in our experiment the decision of the annoyed participants is in a completely different domain than that of the boss. Thereby, we not only contribute a new method for testing remedies against chains of unkindness, but also

for investigating chains of unkindness in general, distinguishing them from chains of imitating the predecessor. Moreover, our results allow us to confirm that the letter writing effect is robust to the gender of participants.

However, there remains doubt with regard to the channel at work. Our results show that letter writing is not more successful in bringing subjectively reported levels of happiness back to their initial pre-shock levels than simply waiting. Moreover, expressing emotions in the letter, which in theory should help to release them and hence lead to increased willingness to give, is negatively correlated with the amount of money passed on to the third participant. All of this suggests that although letter writing increases giving, the reason does not seem to be emotional regulation, which calls for future research into the matter.

2.5   Appendix

On the following pages we display the original instructions used in this experiment, followed by an English translation of the instructions.
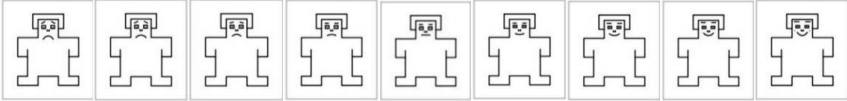
ANLEITUNG

Sie nehmen nun an einem wissenschaftlichen Experiment teil. Diese Anleitung ist für alle Teilnehmer gleich und erklärt Ihnen alles, was Sie für die Teilnahme wissen müssen. Falls Sie Fragen haben, melden Sie sich bitte per Handzeichen, ansonsten gilt während des gesamten Experiments ein absolutes Kommunikationsverbot.

Im Laufe des Experiments können Sie Geld verdienen. Alle Beträge werden direkt in Euro angegeben. Am Ende des Experiments wird Ihnen Ihr Guthaben zusätzlich zum Teilnahmeentgelt von 2,50€ in bar ausgezahlt.

Stimmungsabfragen

Während des Experiments werden Sie mehrmals gebeten, Ihre Stimmungslage durch die Auswahl von Symbolen auszudrücken. Bitte geben Sie an, wie Sie sich im entsprechenden Moment fühlen, indem Sie eines der neun Symbole auswählen. Dabei stehen die äußeren Symbole für extreme und die mittigeren Symbole für entsprechend weniger eindeutige Stimmungslagen. **Es gibt keine richtigen oder falschen Antworten. Antworten Sie bitte ehrlich und nur Ihrer aktuellen Stimmungslage entsprechend.** Die Stimmungsabfragen haben keinerlei Auswirkung auf Ihre Auszahlungen.

Bitte lesen Sie sich die Adjektive zu den Extremen genau durch. Versuchen Sie sich die Adjektive in Verbindung mit den Symbolen zu merken, da Ihnen während des Experiments nur die Symbole angezeigt werden. Sie sehen also nur auf diesem Blatt die Symbole in Kombination mit den Adjektiven.

unzufrieden unglücklich genervt verzweifelt schwermütig		zufrieden glücklich erfreut hoffnungsvoll ausgeglichen

Ablauf des Experiments

In diesem Experiment gibt es drei Rollen: Boss, Bewerter und Übersetzer. Zu Beginn des Experiments wird zufällig ein Teilnehmer aus dem Raum ausgelost, der als einziger die Rolle des Bosses übernimmt. Alle anderen Teilnehmer werden in Paare aufgeteilt, die einander für den Rest des Experiments zugeordnet bleiben. Sie alle sind zunächst Bewerter, und das Experiment besteht für sie aus drei Teilen: Arbeiten, Aufteilen und Einschätzen.

Der Boss erhält ein Fixgehalt, aber nur wenn er die folgenden Bedingungen voll erfüllt:
- Der Boss macht einen der Teilnehmer aus jedem Paar zu einem Übersetzer.
- Er entscheidet sich dafür, dass die Übersetzer zur Eile angetrieben werden.
- Er drückt durch Bestätigen per Klick jedes Mal seine Unterstützung für eine Unterbrechung aus. Bei jeder Unterbrechung werden alle Übersetzer daran erinnert sich zu beeilen und ihre aktuelle Eingabe wird gelöscht.

Teil 1: Arbeiten

Bewerter	Übersetzer
Ihre Aufgabe ist angenehm. Sie können lustige, unterhaltsame Filme anschauen und je nach Unterhaltungswert bewerten.	Ihnen werden die Zahlen 1 bis 26 und je ein zugeordneter Buchstabe angezeigt. Darunter sehen Sie eine Kombination von Zahlen, die Sie in Buchstaben übersetzen sollen. Geben Sie die entstehende Kombination von Buchstaben ein, um zur nächsten Übersetzungsaufgabe zu gelangen.
Sie bekommen dazu ein Video nach dem anderen angezeigt, und die Möglichkeit jedes Video durch einen Haken bei „Gefällt mir!" zu bewerten. Es werden so lange neue Videos angezeigt, bis die Übersetzer Teil 1 beendet haben.	Um die unangenehme Arbeitsphase zu verlassen, muss jeder Übersetzer eine Mindestzahl an Aufgaben richtig lösen. Diese Zahl liegt zwischen 3 und 10 und wird den Übersetzern zugelost, aber nicht mitgeteilt.
	Sie bekommen so lange neue Aufgaben angezeigt, bis alle Übersetzer ihre Mindestzahl an Aufgaben gelöst haben. In dem Ausnahmefall, dass alle Teilnehmer nur noch auf einzelne Übersetzer warten, wird Teil 1 vorzeitig beendet.
Gesamtauszahlung für Teil 1: 0,10€	**Gesamtauszahlung für Teil 1:** 0,10€

Als Bewerter denken Sie bitte daran, dass Sie Ihre Kopfhörer aufzusetzen sobald das erste Video erscheint.

Teil 2: Aufteilen

Der Boss nimmt nicht an Teil 2 teil. **Als Übersetzer** und **als Bewerter**, werden Sie gebeten jeweils einen Geldbetrag zwischen sich und einer anderen Person aufzuteilen. Der Geldbetrag ist für alle Teilnehmer gleich hoch. Wichtig: Diese andere Person ist nicht Ihr Partner, sondern ein Teilnehmer eines anderen Experiments. Dieses andere Experiment wird im Sommersemester 2017 im BaER-Lab stattfinden.

Teil 3: Einschätzen

Der Boss nimmt nicht an Teil 3 teil. **Als Übersetzer** und **als Bewerter** sollen Sie die Stimmung einschätzen, die der Ihnen zugeloste Partner unmittelbar nach Teil 1 (Arbeiten) angegeben hat. Liegen Sie mit Ihrer Einschätzung richtig, erhalten Sie 0,40€.

Auszahlung

Die Auszahlungen für dieses Experiment setzen sich wie folgt zusammen:

Bewerter	Übersetzer
Aus Teil 1: 0,10€ **Aus Teil 2:** der Betrag, den Sie bei der Aufteilung sich selbst zugeteilt haben **Aus Teil 3:** 0,40€, falls Ihre Einschätzung richtig war	**Aus Teil 1:** 0,10€ **Aus Teil 2:** der Betrag, den Sie bei der Aufteilung sich selbst zugeteilt haben **Aus Teil 3:** 0,40€, falls Ihre Einschätzung richtig war
Zusätzlich erhalten alle Teilnehmer das feste Teilnahmeentgelt von 2,50€.	

Der Boss erhält das volle Fixgehalt ohne Abzüge, gegeben dass er die oben genannten Bedingungen ohne Einschränkungen erfüllt. Die Höhe des Fixgehalts erfährt er direkt zu Beginn des Experiments. Falls der Boss sich dagegen entscheidet, die oben genannten Bedingungen zu erfüllen, wird sein Fixgehalt gekürzt. Dann ist dem Boss nur das feste Teilnahmeentgelt von 2,50€ sicher, welches er wie alle anderen Teilnehmer in jedem Fall erhält.

Nach dem Experiment

Im Anschluss an das Experiment werden Sie gebeten, noch einige Fragen zu beantworten. Bitte beantworten Sie alle Fragen ehrlich und vollständig. Ihre Antworten sind anonym und haben keinen Einfluss auf Ihre Auszahlung in diesem Experiment.

Bitte beachten Sie:
- Während des gesamten Experiments ist keine Kommunikation mit anderen Teilnehmern gestattet.
- Alle Handys müssen während der kompletten Experimentdauer ausgeschaltet sein.
- Wenn Sie eine Frage haben, bleiben Sie bitte an Ihrem Platz sitzen und heben die Hand. Stellen Sie bitte Ihre Frage so, dass kein anderer Teilnehmer Ihre Frage mithören kann.
- Sämtliche Entscheidungen erfolgen anonym, d.h. keiner der anderen Teilnehmer erfährt die Identität desjenigen, der eine bestimmte Entscheidung getroffen hat.
- Auch die Auszahlung erfolgt anonym, d.h. kein Teilnehmer erfährt, wie hoch die Auszahlung eines anderen Teilnehmers ist.
- Bitte bleiben Sie bis zum Ende des Experiments an Ihrem Platz sitzen, Sie werden zur Auszahlung mittels der Ihnen zugeordneten Platznummer aufgerufen.

**Viel Erfolg und Danke für Ihre Teilnahme an diesem Experiment!**

INSTRUCTIONS

Welcome to this experiment. These instructions are the same for all participants and provide you with all relevant information regarding your participation in this experiment. If you have any questions, please raise your hand. The experimenter will then come to answer your question. During the whole experiment, we ask you to remain silent and not to talk to other participants.

During the experiment you can earn money. Amounts will directly be displayed in Euro. After the experiment you will receive the money you earned in your experimental account, additional to a fix show-up fee of 2.50€, in cash.

Happiness Ratings

During the experiment you will repeatedly be asked to rate your mood by selecting a suitably symbol. Please, indicate how you feel in that very moment, by choosing one out of the nine symbols displayed below. Symbols on the extreme right or left represent rather extreme moods, while the ones towards the center represent more ambiguous moods. **There are no right or wrong answers. Please be honest and answer only according to your mood at that very moment.** Your mood ratings will not affect your payment in any way.

Please, do carefully study the adjectives describing the moods at the extreme left and right. Please, try to remember the adjectives corresponding to these symbols, as during the experiment only the symbols will be displayed. Hence, you can only study the symbols together with the adjectives on this sheet of instructions.

dissatisfied unhappy annoyed desperate melancholic		satisfied happy joyful hopeful harmonious

Sequence in the Experiment

In this experiment there are three roles: Boss, Rater and Encrypter. At the beginning of the experiment, one participant will be randomly chosen for the role of Boss. Initially, all other participants are assigned the role of Rater. They will be split into pairs of two, and these pairs remain intact until the end of the experiment. For these participants, the experiment consist of three parts: working, sharing, and guessing.

The Boss receives a fix payment, conditional on fulfilling the following requirements without restrictions:
- The Boss converts one participant of each pair into an Encrypter.
- He decides in favor of the Encrypters being pressed to hurry.
- By clicking on a button, he confirms his support of Encrypters being interrupted by reminders to hurry up. With each interruption, all Encrypters are reminded to hurry up and their current entry will be deleted.

Part I: Working

Rater	Encrypter
Your task is enjoyable. You will watch short, funny videos, and have the possibility to rate them according to their value of entertainment.	You will be shown the numbers 1 to26 and one letter of the alphabet corresponding to each number. Below, you will see a combination of numbers, that you will have to translate into the corresponding letters. Enter the resulting combination of letters to get to the next task.
To do so, one video after the other will be shown to you, and you will be able to rate a video by ticking „I like!". Until the Encrypters complete their task, you will always be shown yet another video.	For leaving the annoying work phase, each Encrypter has to solve a minimum number of tasks. This number is a randomly determined number between 3 and 10 and will remain secret. As long as any Encrypter is still working on his minimum of tasks, you will always receive yet another task yourself. In case that all participants end up waiting for a few Encrypters to finish, Part I will automatically stop after a fixed amount of time.
Total payment for part I: 0.10€	**Total payment for part I:** 0.10€

As a Rater, do not forget to put on your headphones as soon as you can see the first video.

Part II: Sharing

The Boss does not participate in part II. **As an Encrypter** or **Rater**, you will be asked to split up an amount of money between yourself and another person. This amount of money is of the same height for all participants. Please not: This other person is not your partner, but instead a different participant of a future experiment. This future experiment will be taking place in the BaER-Lab during the summer term of 2017.

Part III: Guessing

The Boss does not participate in part III. **As an Encrypter** or **Rater** you are asked to guess the mood that the participant who is paired to you has indicated right after finishing part I (working). If your guess is correct, you will receive 0.40€.

Payment

The final payment that you will receive for participating in the present experiment will be composed as follows:

Rater	Encrypter
From part I: 0.10€ **From part II:** the amount of money that you assigned to yourself when splitting up the money between yourself and the future participant of a different experiment **From part III:** 0.40€, in case of a correct guess	**From part I:** 0.10€ **From part II:** the amount of money that you assigned to yourself when splitting up the money between yourself and the future participant of a different experiment **From part III:** 0.40€, in case of a correct guess
Additionally, all participants receive a fixed show-up fee of 2.50€.	

The Boss will receive the fix payment without any reductions, given that he fulfills the requirements as specified above without any restrictions. The height of the fix payment will be revealed to him right at the beginning of the experiment, before any decision needs to be taken. In case that the Boss decides against fulfilling the requirements as specified above, reductions from his fix payment will be made. In that case, the Boss can only rely on the fix show-up fee of 2.50€ that, just as for all other participants, he will receive unconditionally.

After the Experiment

After the experiment you will be asked to answer a post-experimental questionnaire. Please, do give honest and complete answers to all questions. Your answers are anonymous and will not affect your payment in any way.

Important remarks:
- Throughout the experiment you may not communicate with other participants.
- Cellphones must be turned off during the experiment.
- In case of any questions, please raise your hand and remain seated. Please ask your question in a way that does not allow other participants to overhear them.
- All of your decisions are anonymous. Consequently, none of the other participants will learn the identity of anyone who made a particular decision during the experiment.
- Payment is anonymous, hence, the participants will not observe the payment of any other participant.
- Please, remain seated until the end of the experiment. You will be called by your seat number to receive your payment.

Lots of success and thank you for your participation!

Chapter 3

When letter writing increases kindness: Regulating emotions or activating pro-social thinking?

Wendelin Schnedler
Paderborn University

Nina Lucia Stephan
Paderborn University

3.1 Introduction

Anecdotal as well as experimental evidence suggests that people pass on unkind behavior (Ben-Ner et al., 2004; Gray et al., 2014). For example, Strang et al. (2016) show that an individual (person *B*), after receiving little money from his boss (person *A*), also passes on little money to a completely innocent party (person *C*). Understanding chains of unkindness and what can be done against them is crucial for firms and institutions who want to prevent the spread of unfair, inefficient, or unproductive behavior within their organizations.

In a recent study, Strang et al. (2016) suggest a remedy against the chain of unkindness: using an emotion regulation strategy, such as message writing. Subjects in the role of the unkindly treated person *B*, who write a letter to the unkind person *A*, subsequently pass on significantly more money to an innocent third persons *C*, relative to those subjects who are not given the option to write a letter but simply wait. The authors argue that writing a letter helps person *B* to emotionally 'close the case'. Hence, letter writing should affect the mood of subjects in the role of *B* more positively than waiting. On the one hand, this is exactly what Strang et al. (2016) find. On the other hand, Schnedler and Stephan (2019) observe that self-reported happiness is not differently affected by letter writing, compared to waiting.

In the present paper, we consider an alternative explanation for the effect of letter writing that is consistent with subjects remaining unhappy: Any letter is addressing a specific audience. Thus, at the time of writing, subjects are more likely to think about others, compared to subjects who spend their time waiting. In order to get their message across, they need to put themselves into the readers' shoes, i.e. by selecting the right language. Therefore, being given the opportunity to write a letter might activate more pro-social modes of thinking than having to wait. This difference may then translate into larger amounts of money passed on by writers, compared to subjects who wait instead.[1]

If social modes of thinking are activated, not only unkindly but also kindly treated subjects should give more when they write rather than wait. In contrast, the 'closing the case' explanation by Strang et al. (2016) predicts that writing a letter has different effects on kindly treated subjects compared to unkindly treated subjects: if anything, kindly treated subjects should give less after regulating their (positive) emotions with the letter.

Our experiment directly tests which explanation is more suitable by exploiting the different predictions. We find that the effect of letter writing is similar for unkindly and kindly treated subjects: in both groups, letter writing increases the

---

[1]Studies using functional Magnetic Resonance Imaging show that when certain brain regions, associated with perspective-taking, are active, this is associated with increased altruistic giving (Morishima et al., 2012) but they do not use letter writing as a manipulation.

amount of money being passed on to the innocent third party. This is consistent with the explanation that letter writing activates pro-social modes of thinking, and rather not in line with the 'closing the case' explanation.

In addition to this central finding, our data offer an interesting unexpected finding. While being treated unkindly, rather than kindly, strongly and persistently reduces the happiness of subjects in the role of *B*, it does not affect how much money is passed on to the innocent third person. We thus find no evidence for a chain of unkindness.

In our experiment, kind and unkind behavior is induced by giving a randomly chosen boss (*A*) the authority to assign a relatively pleasant or unpleasant task to the other subjects (in the role of *B*).[2] The pleasant task consists in watching and rating a series of funny video clips, while the unpleasant task involves encrypting meaningless codes. After finishing their respective tasks, subjects in our *L* treatment are given the opportunity to write a letter to *A*, while those in the *W* treatment have to wait. Finally, subjects are endowed with money, and decide whether to pass on part of the money to an innocent person (*C*), who is not taking part in the present experiment. A detailed description of the experimental design follows in Section 3.2.

Our design involves a sequence of three players, similar to experiments on third-party punishment (e.g Fehr and Fischbacher, 2004; Rand and Nowak, 2013; Nelissen and Zeelenberg, 2009). The crucial difference is that the third 'player' *C* is not active or even present in our experiment. More importantly, our key question is not how *C*'s possible punishment affects *A*'s behavior toward *B*, but how *B*'s behavior toward *C* is affected by letter writing.

We measure the passing on of unkindness using the dictator game. The widespread use of the dictator game (Engel, 2011) has been criticized as 'searching near the lamppost' (Oechssler, 2010). While there is some debate on whether giving in the dictator game actually measures altruism (e.g. Bardsley, 2008; List, 2007), giving is clearly kind to the receiver. This justifies its use here and more generally in the literature on the chain of unkindness. Moreover, our results become directly comparable to this literature (e.g. Ben-Ner et al., 2004; Diekmann, 2004; Bahr and Requate, 2014).

Since in our experiment we observe how letter writing affects pro-social choices in a dictator game, our study adds to a growing literature using experiments to investigating how pro-social decision making in the monetary domain is affected by soft factors outside the monetary domain, such as context or emotions. For example, Fukui and Toyoshima (2014) show that listening to preferred music increases pro-social behavior in a dictator game. Or, Capra (2004) finds that non-monetarily

---

[2]In real life, any task, depending on the subjective experience, can be perceived as either kind or unkind. The respective kind and unkind tasks in this study are thus meant to be more or less kind in relative terms. For the same reason, the kind task may also be regarded as neutral.

inducing a good mood increases pro-social choices in a dictator, and an ultimatum game. In contrast to these and similar existing studies, we shed light on how and why letter writing might affect pro-sociality.

To the best of our knowledge, there are only two existing experimental economic studies on the effect of letter writing: Xiao and Houser (2005) and Xiao and Houser (2009). In their 2005 study, the authors find that letting subjects B write a letter to A reduces their willingness to punish A. In their 2009 study, they re-investigate their 2005 data and find that dictators who anticipate receiving a letter are less likely to share unfairly in a dictator game. They conclude, that letter writing works as an efficient and cheaper form of sanctioning. The present paper adds to these studies on the effect of letter writing by exploring its effect on another outcome: the passing on of unkind behavior to an innocent third party.

Our experimental set-up draws its inspiration from the literature on chains of unkindness or kindness, which studies whether affected people pass this behavior on to an independent third party—a phenomenon also referred to as downstream indirect reciprocity (Mujcic and Leibbrandt, 2017) or generalized reciprocity (Herne et al., 2013). Mujcic and Leibbrandt, for example, find that people who have been let through at a car park crossing are more likely to let other people through later and Leimgruber et al. (2014) observe that children who receive a sticker from another child are more likely to give away a sticker to a third child. Others show that subjects behave more non-cooperatively after a non-cooperative experience (Rankin and Taborsky, 2009) or that subjects pass on less money in a dictator game after receiving little in a dictator game themselves (Diekmann, 2004; Strang et al., 2016). While these studies establish that a certain behavior is passed on, we examine how a proposed remedy against such chains operates.

Our paper most directly relates to Strang et al. (2016), who investigate letter writing as a remedy against the chain of unkindness. Probably the most important difference is that, whereas Strang et al. only have very few kindly treated subjects, in our study, half of the subjects experience kind behavior. With this balanced data set, the present paper is the first that allows studying and comparing the effect of letter writing on both, kindly and unkindly treated subjects. This enables us to contribute to existing literature by answering our central question: whether letter writing has a general pro-social effect, that is essentially the same among subjects previously treated kindly and unkindly, or whether it only helps unkindly treated subjects and is thus a remedy specifically aimed at the chain of unkindness.

The remainder of this paper is structured as follows. After an introduction to the experimental design in Section 3.2 and an overview of the implementation and descriptive statistics in Section 3.3, Section 3.4 summarizes the main results. Section 3.5 presents additional, incidental results, followed by a conclusion in Section 3.6.

3.2 Experimental Design

The experiment consists of three treatment variants: the letter writing treatment and two control treatments. In the following we describe the sequence and explain the design of the experiment. Then, building on existing findings, we explain the central hypotheses.

3.2.1 Sequence

The sequence follows naturally from the desire to study how unkind behavior gets passed on. In Part I, a subject is given the choice to assign an unkind and annoying tasks to others. In Part 2, in order to obtain kindly and unkindly treated subjects, some subjects are exposed to this task and others not. In Part 3, we apply the treatment, i.e., whether subjects have the opportunity to write a letter or not. In Part 4, we measure kindness toward the innocent third person.

Part 1: Assignment to roles and tasks

In order to induce unkind behavior to potentially start off a chain of unkindness, we need some participants who decide to be unkind to some, and kind to other participants. Therefore, in each session, one participant is randomly assigned to be the boss (A), and is offered a premium conditional on being unkind to some of the participants. All other participants are by default treated kindly. To make sure that participants hold the boss, and not the experimenter, responsible for their experience, the boss is responsible for deciding which participants to treat unkindly. For the same reason, the boss is also given the option not to treat anyone unkindly, though this would mean forgoing the premium.

Except for the boss, all other participants are randomly assigned into pairs, each pair consisting of one player X and one player Y. All players X and Y are initially assigned to the task of rating movies (the default). Before the start of this task, however, the boss has to decide whether or not to assign either all players X or all players Y to an encryption task instead, which earns him the premium of $10\,€$. He then has to decide which group of participants, X or Y, to assign to this much less pleasant task, thereby being unkind to those who become encrpyters and kind to those who remain raters.

Part 2: Pleasant versus unpleasant tasks

In order for subjects to feel kindly or unkindly treated, they all have to execute the task assigned to them by the boss. The rating task is designed to give the respective subjects a pleasant experience. The task of the raters is to watch short, funny videos. In order to make participants actively engage in this task, they are asked to rate the videos, by checking a box in case they enjoy watching it. To make sure the raters are well entertained throughout the duration of Part 2, a sufficient amount of videos is presented.

In contrast, the encryption task needs to be experienced as unpleasant, so participants perceive the bosses choice to assign them to this task as unkindness. We explain the six ways in which our design ensures that the encryption task is perceived as unkind in comparison to the rater task.

First, as stated by Benndorf et al. (2018), encryption tasks promise minimal learning, which reduces the chance for single participants to enjoy getting better at it. In our case, subjects have to convert sequences of numbers into meaningless combinations of letters. Second, encrypters sit in the same room as raters. They can hear raters laugh at the funny videos which makes them aware of the pleasant alternative experience that they miss due to the boss' decision. Third, the work of the encrypters is interrupted by annoying reminders from the boss, telling them to hurry up. These reminders appear on the encrpyters' screens and have to be clicked away in order to return to the encryption task. After each reminder, the last encryption input is lost and has to be re-entered. Fourth, encrypters are put under time pressure by giving them an incentive to solve as many encryption tasks as possible, as quickly as they can. The incentive is that the experiment only continues once either, every encrpyter has solved an assigned but unknown minimum number of encryption tasks between 3 and 10, or, all participants end up waiting for only a few participants to finish at least their minimum. Fifth, both tasks are only rewarded by a rather symbolic payment of 0.10 €. This payment may feel inadequately low for someone engaged in encrypting. Sixth, to make the relative unfairness directly visible, participants receive 'feedback' on their own and their partners' performance. This includes information on how many videos were liked by the rater, and how often the encrypter was interrupted by the boss.

Part 3: Treatments L and W

The next step differs across treatments. In the *L* treatment, all raters and encrypters are given three minutes to write a letter to the boss. Participants are notified that they have 3 minutes to write a letter to the boss by typing into the text field below the notification. In order to underline the real consequences of their decision in this stage, participants are also informed that letters are printed and handed out to the

boss at the end of the experiment. Whereas participants need to type at least one character into the text box before being able to send the letter and leave the stage early, we are aware that all participants can still freely choose whether and in what way to make use of the letter writing option.

In the W treatments, participants are not given the chance to write a letter. Instead, they wait for 3 minutes in order to ensure that the time that passes between the experience of kind and unkind behavior and the dictator decision is as long as in the L treatment. There are two variants of this treatment: W_0 and W_M. In the W_0 treatment, participants only wait. In the W_M treatment, participants are asked to answer a question while waiting.[3]

Part 4: Dictator game

In the last part of the experiment, we measure whether participants pass on unkind behavior. For making sure that participants actually decide to be unkind, and are not just imitating a choice previously made by someone else, we do not observe unkind action by observing the decision to assign yet another unkind task (mimicking the bosses' decision situation). Instead, unkind action is observed in a dictator game. This comes with the advantage that dictator game results are easy to compare both across treatments, and across raters and encrypters. Hence, we also learn whether kind behavior is passed on. A convenient side effect of designing this last part of the experiment as a dictator game is that it allows rating the participants' actions as more or less kind in the same way as done in previous studies (Strang et al., 2016; Gray et al., 2014; Ben-Ner et al., 2004): Each dictator receives 10 € and is asked to leave between 0 € and 10 € in increments of 1 € to a person C.

To avoid image concerns, we make sure that these persons C will never be in direct contact with the dictators. All participants are informed that person C is neither present in the laboratory nor participating in another session of the same experiment. Instead person C is a participant of a future, unrelated experiment, that is taking place later in the term.[4]

---

[3]They were asked to report how much money ($_M$) should be offered to make up for the kind or unkind experience. The intention was to keep the emotional experience vivid. We do not find this additional question to affect participants' behavior regarding any of the outcomes.

[4]Future participants were invited to different sessions in May, June, and September 2017. As part of these sessions, each future participant received the money left by one dictator of the present experiment.

3.2.2 Elicitation of Happiness

In order to shed light on potential motives behind participants' choices, we elicit participants' happiness at different stages of the experiment. This allows us to measure how participants' happiness changes throughout the experiment, and whether it differs between treatments and assigned tasks. To facilitate comparison, we use the same method to elicit happiness as used by Strang et al. (2016) and first developed by Bradley and Lang (1994).[5] In between the different stages of the experiments, all participants are asked to describe their current mood on a nine-point Likert scale by clicking on a respective symbol (*Self-Assessment Manikin*). This elicitation of happiness, or instantaneous mood, is done four times. First, before participants learn which task they are assigned to, and then repeated immediately after the end of the encryption or rater task respectively. Third, after the treatment, and fourth, after the dictator game.

3.2.3 Hypotheses

Letter writing can help regulating emotions. There are two ways how this may work: through emotional expression or emotional closure. First, Xiao and Houser (2005) suggest that letter writing is an outlet for emotional expression. They observe that people are less inclined to reject unfavorable offers if they can write a letter to the unfair proposer. According to them, unfairly treated people are emotionally charged and have a desire to express their emotion. If they cannot directly do so, they express themselves by rejecting the offer. If they can do so in a letter, they feel less need to reject the offer. Possibly, subjects charged with negative emotion generally seek the next outlet to express their emotion, and do so through an action, if no other outlet is available. Second, Strang et al. (2016) suggest that letter writing leads to emotional closure. By addressing the person that caused the negative emotions, it becomes easier to let go of the negative emotion. Both ways are in line with Damasio (1994), who describes that humans are inherently trying to avoid emotional pain, that may be caused by experiencing unkindness. According to Damasio, humans are pretuned to use any closely following social situation to regulate their emotional pain. In our *L* treatment, the next social situation following the emotional experience is the letter writing stage, where participants can express their emotions by communicating. In the *W* treatment, the next social situation following the emotional experience is the dictator game, where the same need for emotional expression may be satisfied by giving little to *C*.

---

[5]We are aware that despite using the same method of happiness elicitation, there may still be minor procedural differences compared to Strang et al. (2016), and that these may partly explain why happiness measures differ.

Whether it is the emotional closure by Strang and co-authors or the emotional expression by Xiao and Houser, in both cases, negatively charged encrypters are more likely to benefit from letter writing than positively charged raters, which leads to our first hypothesis.

Hypothesis 1. *Letter writing has different effects on participants in a subsequent dictator game: After experiencing unkindness, letter writing leads to higher rates of giving. After experiencing kindness, letter writing does not lead to higher rates of giving.*

When writing a letter, the writer inevitably turns to an audience, the reader. This has two consequences that may affect subsequent decision making. First, as any audience consists of one or several other human beings, writing a letter thus makes the writer aware of the existence of others. Second, to make sure that the writing will be understood by the audience, the writer has to take the reader's perspective.

Neuroscience literature, making use of functional Magnetic Resonance Imaging, shows that under different circumstances, specific regions of the brain become more active. These brain regions then play a greater role in imminent decision making. Several studies describe how both, increased activity in the medial prefrontal cortex, associated with social situations, especially when other humans are involved in ongoing decision making (Krach et al., 2008), and the temporoparietal junction, associated with perspective-taking (Morishima et al., 2012), may enforce pro-sociality. Just to give some examples, Krach et al. find that participants playing the prisoner's dilemma, and Rilling et al. (2004) find that participants playing the ultimatum game, produce more cooperative choices when playing with humans instead of computers. Hare et al. (2010) show that brain activity stimulated through a perspective-taking task leads to increases in donations to charity. When in our experiment the writer addresses the boss as his audience, he inevitable engages in perspective-taking. Therefore, letter writing should activate the same brain regions as other social situations, as well as activate brain regions associated with perspective-taking. Knowing that both of these brain regions are associated with pro-sociality, we expect that dictator game decisions taken after letter writing are more pro-social.

Hypothesis 2. *Introducing letter writing has the same effect on participants in a subsequent dictator game: (a) Letter writing leads to higher rates of giving—irrespective of whether subjects were assigned the role of rater or encrypter. (b) The increase in the rates of giving is equally high among encrypters and raters.*

3.3 Implementation and Descriptive Statistics

The experiment was conducted in May and June 2017 at the Business and Economics Research Laboratory (BaER-Lab) at Paderborn University in Germany. The experiment was programmed using the software z-Tree (Fischbacher, 2007) and participants were recruited with the help of ORSEE (Greiner, 2015). In total, eleven sessions were run, with each session lasting around 45 minutes and earnings averaged at 10.60 €.

3.3.1 Descriptive Statistics

	boss	encrypter	rater	total
W_0	3	40	40	83
W_M	4	59	59	122
L	4	55	55	114
total	11	154	154	319

Table 3.1: Numbers of participants by groups.

As shown in Table 3.1, there were 319 participants in total, 114 in the letter writing treatment L and 205 in the waiting time treatment W. As will become clear in the next section, results from W_0 and W_M do not differ, and will be pooled for most of the analysis. In each session, the boss did choose to assign half of the participants to the encryption task, resulting in a total of 154 observations from encrypters and 154 observations from raters.

	W_0	W_M	L	encrypters	raters
age	23.55	23.09	22.33	22.74	23.10
male	34%	33%	36%	32%	36%
economics major	38%	41%	45%	42%	40%
engineering	8%	7%	7%	7%	8%
cultural science	9%	8%	7%	7%	9%
teaching	40%	41%	36%	40%	39%
initial happiness	5.91	5.58	5.83	5.76	5.75

Table 3.2: Demographic statistics by groups.

Randomization has worked considerably well, as shown in Table 3.2. In particular, the happiness elicited at the very beginning of the experiment is very similar across treatments.
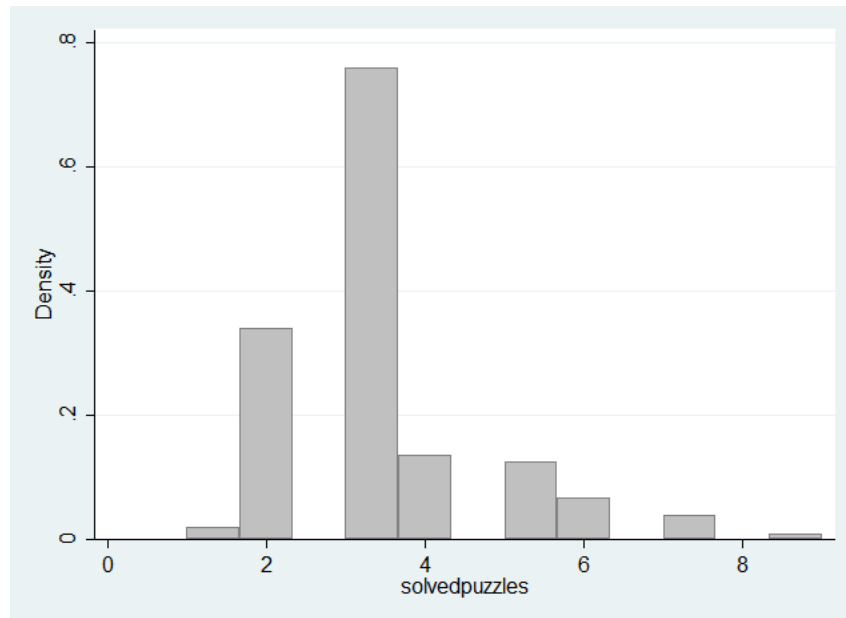

Figure 3.1: The figure shows the distribution of correctly solved encryption tasks by encrypters.

Before actually analyzing the data, there are a couple of design checks. First, we want to see whether all encrypters were actually engaging in their task. Figure 3.1 shows that while most encrypters solved at least three tasks, hardly any of them succeeded in finishing seven or more tasks before the end of the stage.[6] As in each session there was one encrypter assigned to solve 10 encryption tasks, all participants did thus eventually end up waiting for this single person (and a few other, exceptionally slow participants) to reach the assigned target. As a result, every session was terminated by the experimenters after a maximum of three minutes. All treatment phases were thus of equal length.

Second, we want to check whether the manipulation was successful in the sense that it made encrypters relatively more unhappy. Comments from the post-experimental questionnaire suggest that the task was indeed perceived as difficult and annoying, which increased the pressure to finish quickly by solving as many tasks as fast as possible. Answers from the questionnaire also suggest that over

---

[6]The raters did also actively engage in their task. Raters watched between 2 and 8 videos and reported at least one and at most 8 likes.

time, encrypters started to be increasingly frustrated. In all treatments, initial median happiness was rated 6 out of 10. After the treatment, self-reported happiness is 4 points higher among raters compared to encrypters, with a median of 3 and 7 respectively. Figure 3.2 illustrates this difference, and running a Mann-Whitney-U test confirms its significance (p-value 0.00). This effect on reported happiness is, in its size, comparable to the effect found by Strang et al. (2016) when receiving an unfairly low rather than a fair share of money. This similarity supports our assumption that the induced unkindness is strong enough to start off a chain.
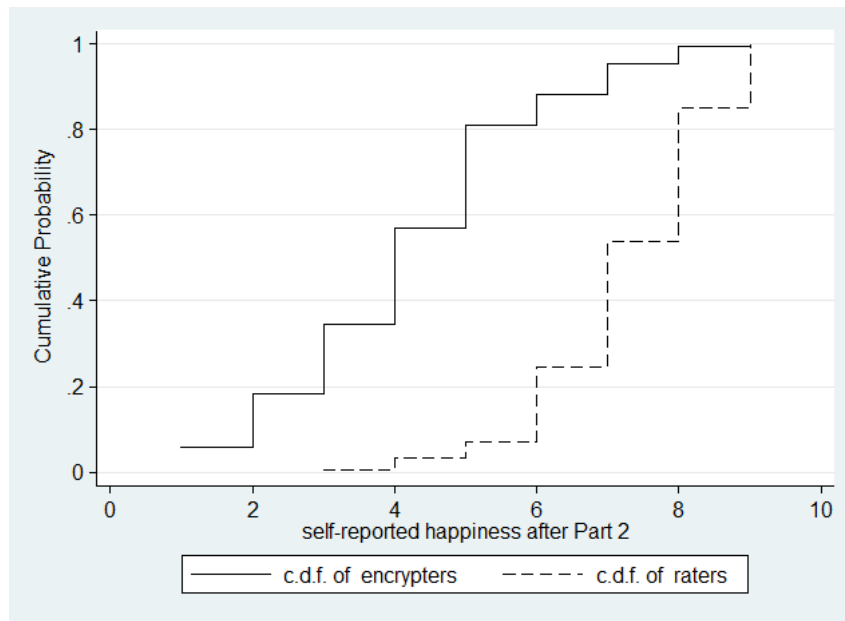

Figure 3.2: The probability of reporting a low happiness is lower for raters than for encrypters.

Finally, we want to know whether encrypters perceived the task assignment as unkindness. Some evidence comes from the content of the letters written by encrypters compared to raters. While participants were free to design the letter as they wanted, most did choose a typical letter format, starting by addressing the recipient, in this case the boss, in formal language as if writing a letter to a stranger. Even informal letters were written as comments or messages directed to the boss. Hence, the overwhelming majority of participants took the letter writing serious.[7] In the typical encrypter letter (see Appendix), the author complains about one or several issues he had while carrying out his task, such as the time pressure, or

---

[7]Only 3 encrypters and 10 raters sent empty letters, meaning that they did write less than one word.

being interrupted by the boss' message to hurry up. Some explicitly blame the boss for his decision and state that they would rather have done the other task. In the typical rater letter (see Appendix), the author thanks the boss, either for having been assigned to the more pleasant task, or for making the task so enjoyable, thereby often commenting on the videos. Overall, the content of the letters supports the assumption that the boss was actually held responsible for the kind or unkind task assignment.

3.4 Main Results

First, we find that in line with both hypotheses, encrypters are more generous in the L treatment. The whole distribution shifts to higher values of giving, as shown in Figure 3.3.[8] On average, the amount passed on to the third person increases by $0.69 \in$ when moving from the W to the L treatment, which is (albeit weakly) statistically significant (p-value of Mann-Whitney-U test 0.058).

Figure 3.3: The willingness to give differs between treatments: Low amounts are less frequent in the L than in the W treatment.

Second, the distribution of rates of giving by raters also shifts to the right, as

---

[8]To create Figures 3.3 and 3.4, the two control conditions were pooled, after running a Mann-Whitney-U test confirmed that there is no significant difference between the amounts given away by encrypters in W_0 and W_M (p-value 0.39) and raters in W_0 and W_M (0.86) respectively.

shown in Figure 3.4. The average amount given increases by 0.64 €, which is again weakly significant (p-value of Mann-Whitney-U test 0.0525). This is more in line with Hypothesis 2 rather than Hypothesis 1.
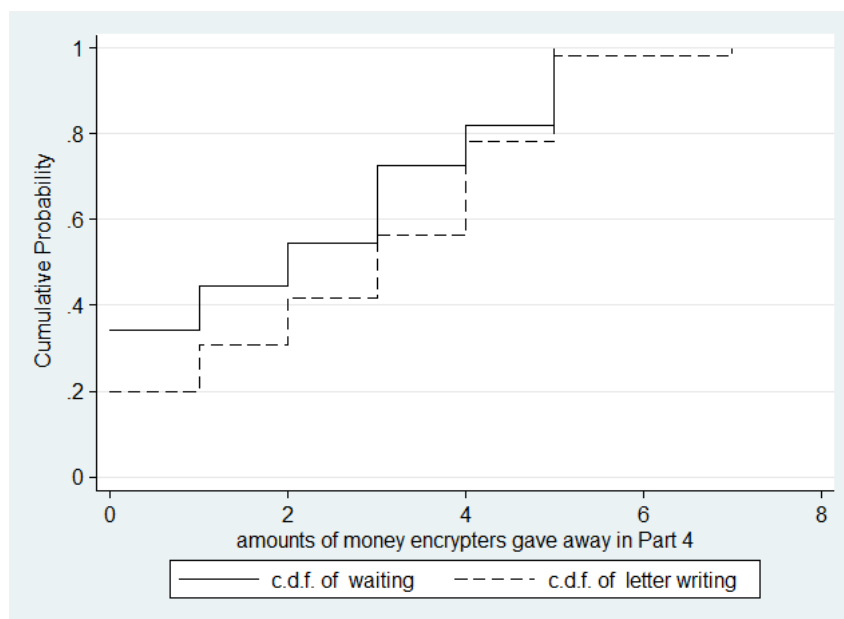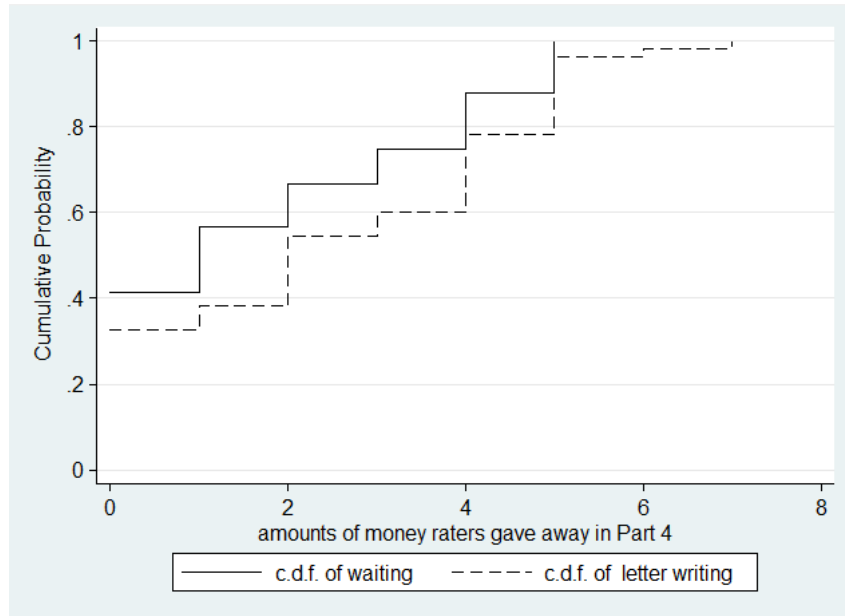

Figure 3.4: The willingness to give differs between treatments: Low amounts are less frequent in the *L* than in the *W* treatment.

Comparing Figure 3.3 and 3.4 shows that the shift in giving among raters is very similar to that of encrypters. To formally test our hypotheses, we compare these differences in the rates of giving across raters and encrypters. As illustrated in Figure 3.5, the magnitudes of the two effects are very similar, with average increases of 0.69 € versus 0.64 € respectively.

We calculate and compare Cohen's d effect size measures for both effects. Cohen's d for the effect among raters is -0.34 standard deviations and for encrypters is -0.35 standard deviations. This difference between the effect sizes of 0.009 standard deviations, is statistically not significantly different from zero (p-value 0.97).
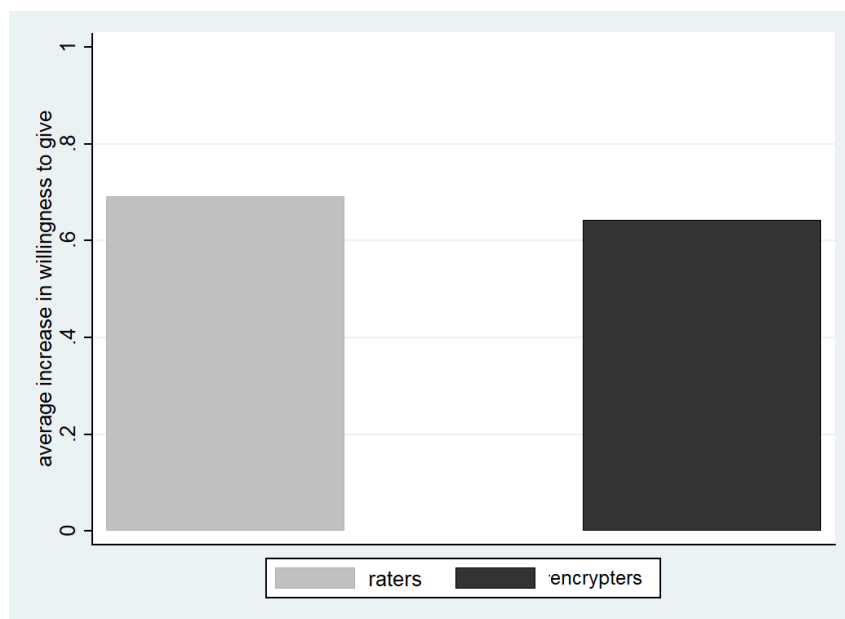
Figure 3.5: The average increase in amounts given due to letter writing is similar for raters and encrypters.

	Tobit1	Tobit2	Tobit3
letter writing	0.945***	0.942***	0.855*
	(0.337)	(0.335)	(0.470)
rater		-0.548*	-0.613
		(0.323)	(0.407)
letter writing x rater			0.177
			(0.674)
male	-0.809**	-0.784**	-0.794**
	(0.347)	(0.345)	(0.347)
_cons	1.517***	1.783***	1.818***
	(0.236)	(0.280)	(0.310)
sigma _cons	2.715***	2.702***	2.701***
	(0.147)	(0.146)	(0.146)
n=308		* p<0.1, ** p<0.05, *** p<0.01	

Table 3.3: Estimates from Tobit regressions with giving in the dictator game as dependent variable, double-censored and standard errors displayed in parenthesis.

A regression analysis provides further evidence. The first line of Table 3.3 reveals a highly significant and positive baseline effect of the letter writing treatment on the

amount given away by dictators of around 0.90 €. The second line shows the baseline effect of the rating task, compared to the encrpytion task. In one specification, this effect is marginally significant (p=0.091), which is an interesting side observation but not related to our hypotheses, that are concerned with whether letter writing affects raters and encrypters differently. In order to address these hypotheses, we turn to the interaction between letter writing treatment and being a rater in the third line. The interaction effect is close to zero and not significant at any conventional level. Hence, we cannot support that the letter writing effect is different for raters compared to encrypters. Summarizing, we reject Hypothesis 2 in favor of Hypothesis 1:

Result 1. *Raters and encrypters are affected similarly by letter writing: they both increase giving to an innocent third person.*

This similarity allows us to pool the data obtained from raters and encrypters, in order to more precisely estimate the general increase in giving caused by letter writing. This general effect of letter writing is an average increase in giving of 0.67 € (or 35%), which is highly significant (p-value of Mann-Whitney-U test below 0.01). Figure 3.6 shows how the overall distribution of giving shifts to the right.
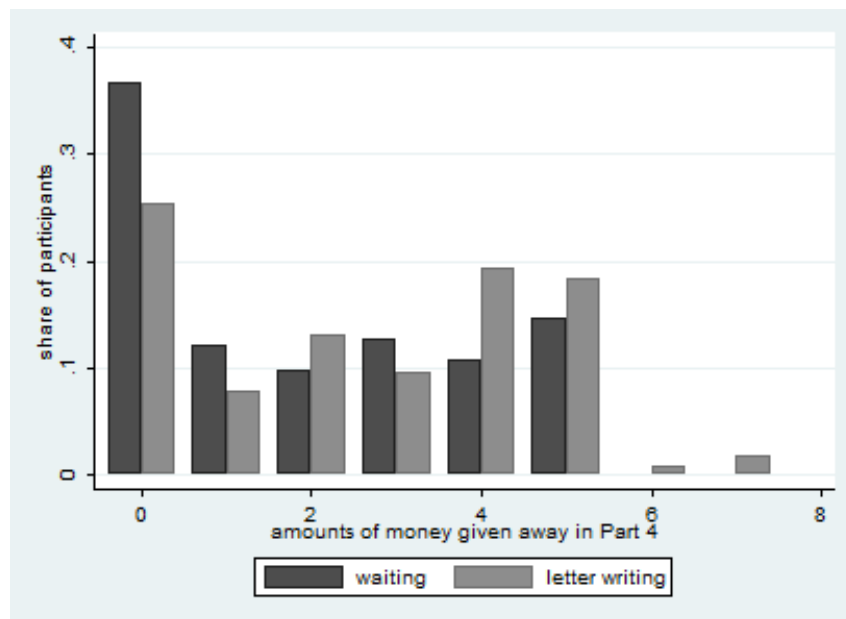

Figure 3.6: Subjects in the *L*, rather than *W* treatment give more to an innocent third person.

After having had the opportunity to write a letter, the distribution of amounts given by raters and encrypters is not only increased, but it also looks strikingly similar,

as becomes evident when comparing the dashed lines in Figures 3.3 and 3.4. Indeed, there is no significant difference between these two distributions (Kolmogoroff-Smirnov test p-value: 0.98). At first glance, this may suggest that letter writing successfully interrupts a chain of unkindness. For this interpretation, however, we have to establish that there has been a chain of unkindness in the first place. We will turn to this issue in the following section.

3.5 Incidental Results

After having established the main result, this section describes additional incidental findings related to our topic of research and based on the experimental data we collected. First, we describe the effect of gender on our outcome variable. Second, we revisit the question whether our data support the existence of a chain of unkindness that arose at the end of Section 3.4.

Gender effect

In our regression, being male has a significant negative effect on the willingness to give, as shown in the fourth line of Table 3.3. This is in accordance with previous literature on dictator games (Engel, 2011). Across all specifications, we find that males give away between 0.784 € and 0.809 € less than females. While this is an interesting finding for itself, it also stresses the importance of controlling for gender when estimating effect sizes, and underlines the robustness of our results.

Chains of unkindness

For evidence of the existence of a chain of unkindness, we turn to the *W* treatment, where subject's choices to give are not affected by the opportunity to write a letter. In this treatment, given that chains of unkindness exist, encrypters should give less than raters, because they have been treated unkindly. The cumulative distributions of amounts given by raters and encrypters, however, seem very similar—compare solid lines in Figures 3.3 and 3.4. Testing for a difference using a Kolmogoroff-Smirnoff test reveals that these two distributions are also statistically not different from each other (p-value 0.99). If anything, raters, rather than encrypters, are less willing to give away money. This negative effect of being a rater is confirmed in our regression analysis. The second column of Table 3.3 reveals a negative, though only marginally significant effect of 0.548 € of being a rater on the willingness to give.

77

Result 2. *Raters are not passing on more to an innocent third person than encrypters.*

Recall that both groups do significantly differ in their happiness. Subjects seem to be capable to mentally separate what they have been exposed to by the boss from what they are giving to the third person. There is no chain of unkindness in our data and subjects do not need to be stopped from paying it forward. As our results do not provide evidence in support of the chain of unkindness, we turn to differences between our and other studies that might explain this. Compared to Strang et al. (2016), our setup differs firstly because their study only uses observations from female participants and secondly because unkindness is induced by a dictator game.
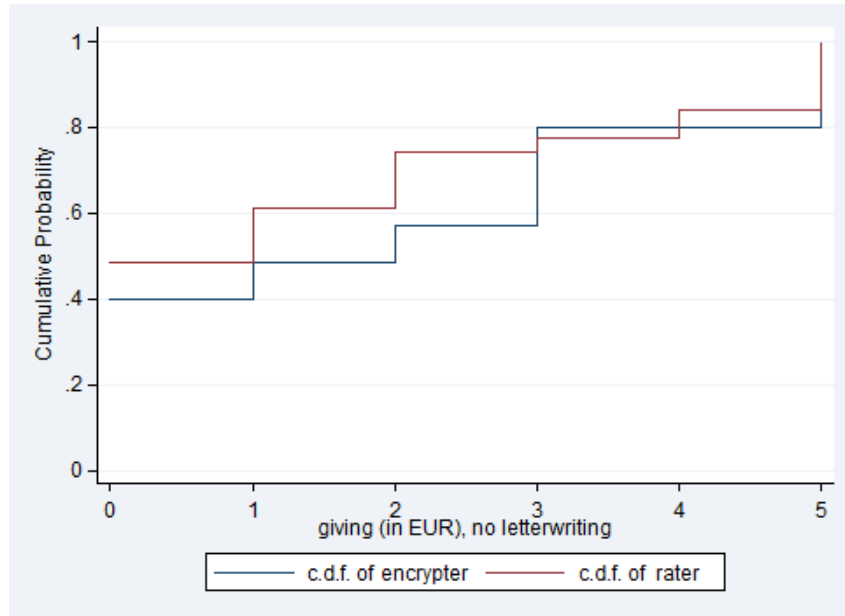

Figure 3.7: The willingness to give does not depend on which task preceded the dictator game: There is no significant difference between the amounts given by raters and encrypters.

First, we restrict our sample to only female subjects to make it more comparable. Still, we do not find evidence that encrypters give less than raters (Kolmogoroff-Smirnov test p-value: 0.89), as illustrated in Figure 3.7.

Second, we revisit the way we induce unkindness. A necessary condition for a chain of unkindness is that participants *B* are actually affected unkindly by going through the encryption task to start off the chain. A possible limitation of our study may be that the induction of unkindness is not strong enough. However, the induction of unkindness does significantly decrease elicited happiness, as described in Section 3.3 and illustrated in Figure 3.2. Median happiness among encrypters drops from 6 to 4 (p-value <0.001 Mann-Whitney-U test), and is thus dropping by as

much as in the experiment by Strang and co-authors, where the unkind experience consisted in receiving 5 € out of 25 €, instead of 12.50 € out of 25 € from a previous dictator.

While it is clear from the present results that we do not find a chain of unkindness, it remains unclear whether there actually is a chain of unkindness in the experimental setting by Strang et al. (2016). Their study does not have enough kindly treated subjects for a statistical analysis in order to see whether these behave differently from unkindly treated subjects.

3.6 Conclusion

Analyzing data from a laboratory experiment, we investigate whether writing a letter helps with regulating emotions or whether it has a general pro-social effect. We find that letter writing leads to increased rates of giving irrespective of whether subjects experienced an annoying or enjoyable task beforehand. This result is especially remarkable since people do report to be in very different moods before writing the letter, which is also reflected in the content of the letters. Even the magnitudes by which giving increases, are very similar for the subjects who previously experienced unkindness and kindness. This suggests, that letter writing may activate more social modes of thinking in the sense of stimulating caring for innocent others in general. For further confirmation of the explanation behind this hypothesis, future research could make use of functional Magnetic Resonance Imaging. This would allow to see whether letter writing activates regions of the brain that are typically associated with social activities. Our result suggests that letter writing can promote selfless behavior.

As a side-result, we cannot support the presence of a 'chain of unkindness' with our data. It may be that chains do not continue when changing the decision situation. While experiencing kind and unkind behavior has opposing effects on participants' self-reported happiness, they do not act upon it. Instead, they give equal amounts to a third party in a subsequent dictator game.

3.7   Appendix

To give the reader an impression of how the letters written by raters and encrypters in the *L* treatment look like, two representative letters are displayed below.

Letter from encrypter

Sehr geehrter Herr/ Frau

hiermit möchte ich sie bitten, den Übersetzern doch bitte mehr Zeit bei ihrer Arbeit zu geben, da dies unter ihren Bedingungen kein effizientes Arbeiten darstellt. Zudem sind die anderen Arbeiter mit ihrer Nebenbeschäftigung, dem Videos gucken, ziemlich laut, sodass dies das Übersetzen noch schwieriger gestaltet.

Mit freundlichen Grüßen

Letter from rater

Sehr geehrter Herr Boss,

ich möchte die Möglichkeit nutzen, Ihnen perönlich für diese angenehme Arbeitsatmosphäre zudanken.DIe Videos sind überwiegend lustig.

Ihr zufriedener Mitarbeiter

Chapter 4

Do children stop cheating when someone else reaps the benefits?

Julia Kramer
Paderborn University

Silvia Lübbecke
Paderborn University

Nina Lucia Stephan
Paderborn University

4.1 Introduction

Honesty is an essential ingredient of economic activity. Therefore, economists have long been interested in studying cases of and motives behind dishonest behavior in many areas, such as tax compliance (Andreoni et al., 1998), job applications (Stigler, 1962), honesty at the work place (Tyler and Blader, 2005; Nonis and Swift, 2001), in education (Jacob and Levitt, 2003; Whitley, 1998), in research (McCabe et al., 2001; McCabe and Trevino, 1997) or in politics (Nye, 1967). The present paper contributes to existing literature concerned with the factors driving individuals to behave dishonestly.

Factors driving dishonesty do not only include individual profits and potentially the risk of being detected, but also describe the existence of psychological costs of lying (Abeler et al., 2014; Fischbacher and Föllmi-Heusi, 2013; Thielmann and Hilbig, 2018). It is widely accepted that individuals are motivated to behave honestly to preserve a positive image, not only towards the outside world but also towards themselves (e.g. Barkan et al., 2015; Mazar et al., 2008; Pittarello et al., 2015). Hence, individuals look for ways to excuse their dishonest behavior. The availability of a justification is thus another driving factor underlying dishonest behavior.

One possible justification is to describe own dishonesty as a white lie (Erat and Gneezy, 2012), that is a lie that (also) benefits someone else, more generally referred to as self-serving altruism (Irlenbusch and Villeval, 2015; Shalvi et al., 2015; Klein et al., 2017). For example, Gneezy (2005) or Conrads et al. (2013) show that dishonesty increases when a lie also benefits someone else. Other studies find that even in the absence of individual profits from cheating, individuals engage in cheating to benefit others (e.g. Gino et al., 2013; Wiltermuth, 2011). Klein et al. (2017) compare how the willingness to cheat depends on whether the profit is received by the decision maker, partly received by someone else, or solely received by someone else. While they find that benefits through self-profit outweigh psychological costs of social justifiability, they also find that when the profit is received solely by someone else, cheating rates are not significantly different from those when the profit is received by the decision maker alone.

Even though there is a large literature in economics as well as other social sciences on both, the reasons for and impacts of dishonesty, little is known about how dishonesty develops in childhood and adolescence. We address this by studying decisions made by young children and adolescents. More precisely, we investigate whether the motive to altruistically cheat, despite the absence of individual profits from cheating, already exists among children, and thus very early in life, just as Klein et al. (2017) show it to be the case among adults. In order to do so, we test whether there is more or less cheating, when individual profits from cheating are taken away, and instead the benefits from cheating go to another child.

In the present paper, we describe a lab in the field experiment inspired by Bucciol and Piovesan (2011) who study cheating behavior of children aged 5 to 15 years in a lab in the field experiment. Just as in their study, we ask children to flip a coin with a 50% chance of winning a prize. The coin is flipped in private and the outcome is only observed by the children themselves, hence they can either obtain the prize by winning or by falsely reporting the winning outcome. Bucciol and Piovesan find that a significant share of children does cheat, especially children above the age of eight. Just as in the present study, they investigate how choices depend on age. We take the investigation a step further by not only studying cheating out of selfish motives. Instead, our central question is how decisions change when cheating does not benefit the cheater himself, but benefits someone else instead, and may thus happen out of other motives. Moreover, we provide 256 independent observations from 521 children aged 3 to 16, including observations from siblings, friends, as well as strangers. This rich dataset enables us to describe how the answer to our central question depends on both, age, and the relationship to the other child.

Our central result is that while a significant share of young children cheat in order to obtain an individual profit for themselves, there is significantly less beneficial cheating when the profit is obtained by someone else instead. Cheating to benefit someone else is relatively larger when the other person is known. Our central result is not only dependent on the relationship to the other child, but also dependent on the age. The result of beneficial cheating being reduced when the profit is obtained by somebody else is strongest among very young children. This finding suggests that whereas selfish motives may be strongest in this subgroup, altruistic motives may not yet have developed among the very young. In contrast, older children make less of a difference between obtaining a profit for themselves, versus obtaining it for someone else, which is in line with literature on adult cheating behavior (Klein et al., 2017).

The remainder of the paper is organized as follows: In Section 4.2, we formulate expectations regarding our results, based on existing related literature. Section 4.3 describes the experiment in more detail. Section 4.4 describes the implementation and participants. Section 4.5 continues with a summary of the results and their analysis. Section 4.6 concludes.

4.2 Theoretical Considerations

Currently, dishonesty is a broadly discussed topic. A recent overview is provided by Rosenbaum et al. (2014), who review 63 economic and psychological experiments on cheating. While most experimental economic literature investigates adults' lying behavior (e.g. Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2014; Conrads and Lotz, 2015; Houser et al., 2012; Jiang, 2013), very little research has been done to investigate the existence of dishonesty among children, and how dishonest

behavior develops with age. Chytilova and Korbel (2014) find that in when deciding in groups, children cheat more than when deciding individually. Glätzle-Rützler and Lergetporer (2015) and Maggian and Villeval (2016) find that older children lie less than younger children, as they are more aware of the harm caused by lies.

As early as with the age of 3, children start understanding what is socially considered right and wrong (Buchmann et al., 2009; Killen and Smetana, 2008). However, according to Nunner-Winkler and Sodian (1988), while they start realizing moral concepts at very young ages of 3 or 4 years, learning to translate this knowledge into actions takes more life experience. In our experiment, the simple decision to cheat in order to gain a profit for oneself becomes more complex in the treatment group, where someone else receives the benefit from cheating. Very young children may not yet have developed motives for cheating other than egoism. This motive disappears, when cheating only benefits a partner. Older children, however, understand that the decision is now also about being altruistic. Hence, additionally to investigating the overall treatment effect, we are also interested in describing in which way the answer depends on the age of the children. We expect to find that very young children do cheat for themselves, but rather not when cheating solely benefits another child.

Lately, some studies on adults' cheating behavior describe cheating that increases others' payoff. For example, just as in the present paper, Houser et al. (2016) are interested in the question whether cheating changes when it benefits another person, rather than the decision maker himself. They observe that parents are more likely to cheat for their children's benefit than for their own benefit. The present study extends on this by investigating whether this preference for altruistic children is already present during childhood, or whether it develops as people age.

Fehr et al. (2013) study children's and adolescents' behavior, but focus on other-regarding preferences, rather than cheating to benefit others. They find that while choices motivated by spite and egalitarianism decrease with age, choices motivated by altruism increase. In line with this, we expect older children to be more motivated to obtain a profit that only benefits someone else, rather than themselves.

Moreover, several existing studies provide evidence for some individuals engaging in spiteful behavior (e.g. Kimbrough and Reiss, 2012). In other words, some individuals have a preference for making choices that harm, rather than benefit others. Hence, when analyzing cheating in favor of others, we should consider the possibility that some individuals cheat in order to avoid producing a profit for others.

Further existing literature on dishonesty suggests that the decision to benefit someone else is likely also dependent on the relationship of the decision maker to this other person (Belot and Van de Ven, 2011; Chen and Li, 2009). Neuroscientific research offers the explanation that providing to a friend, rather than a stranger, produces more feelings of joy (Fareri et al., 2012). Therefore, we will further specify

the results by considering whether the partner being a stranger versus a known child, or even a sibling, changes the results. We expect that when the two children are related in some way, it is easier for the children to understand the positive benefit of cheating in favor of this other child. Hence, they are more likely to cheat in favor of the other child, or in other words, to behave altruistically.

While gender effects are not in the focus of this study, evidence from previous studies shows that female adults (Eckel and Grossman, 1998; Andreoni and Vesterlund, 2001), female adolescents (Dreber et al., 2014), and female children (Houser and Schunk, 2009) tend to behave more generously. From this, we expect that female subjects in our experiment may also behave more generously towards others, and hence to be more likely to cheat in favor of others.

4.3   Experimental Design

In order to answer our central research question, as well as whether the results depend on age, relationship to the other child, or gender, it is necessary to attract a heterogeneous group of participants. In order to obtain this variation in characteristics, the present experiment is conducted as a lab in the field experiment. In this section, we describe the design of our experiment. First, we describe the setup in the surrounding environment. Second, we explain the sequence of the experiment itself.

4.3.1   Surrounding Field Environment

We conduct the experiment during the open days of a large museum. Specifically, the experimental area is located within the large entrance hall of the museum. The open day is visited by school classes, kindergarten groups and families, and thus provides a heterogeneous pool of potential participants.

We take several measures to advertise our experiment and attract participants outside the experimental area. Doing so outside the experimental area also allows children to get to know the experimenters, before deciding to actually follow them into the experimental area. First, we attract children by asking for their 'help' testing a die. We invite children to throw the die and mark the outcome on a big flip chart, thereby collecting statistics to show the frequency of the different outcomes (Appendix Figures 4.4 and 4.5). Second, we install a poster wall with pictures and information on random devices for the passer-by (Appendix Figure 4.6). Both attractions are also meant as an entertainment for the siblings, parents, or teachers, who are waiting for their children to finish participating in the experiment.

The experiment is framed such that participating in it appears to the children as playing a game of fortune, as it is all about throwing dice or coins and winning something. Moreover, the experiment is just one out of many stands that are

exclusively set up for the occasion of the open day. This ensures that children behave rather normally, and do not make choices different from how they would usually make them.

In order to avoid participants from being distracted or feeling observed while participating in the actual experiment, the experimental area is surrounded by solid walls (e.g. Appendix Figure 4.5). There is one single entry and exit to the experimental area, so children cannot get lost. Additional walls inside the experimental area are used to guide the children through the different steps of the experiment, as explained in Section 4.3.2.

Children who decide to participate are asked to wait in a single line in front of the entrance to the experimental area. Upon arrival at the entrance to the experimental area, every other pair is either sent to one side, to participate in the control group, or to the other side, in order to participate in the treatment group. Hence, this is a between-subject design. In order to increase comparability, both control and treatment group sessions take place parallel to each other throughout the experiment. When entire school classes or kindergarten groups come to participate, they are thus split up into both control and treatment group, thereby avoiding randomization issues. Hence, inside the experimental area, two pairs of children can participate in the experiment at the same time. Within the experimental area, the control and treatment group areas are separated from each other with the help of additional walls.

4.3.2 Sequence

Children participate in pairs of one 'active' and one 'passive' child. We assign roles randomly by letting one child role a two-colored die, and letting the other child bet on the winning outcome, as illustrated in Appendix Figure 4.7. The winner becomes the active child. After assigning the roles, both children accompany the experimenter into the experimental area to Position 1, as shown in Figure 4.1.
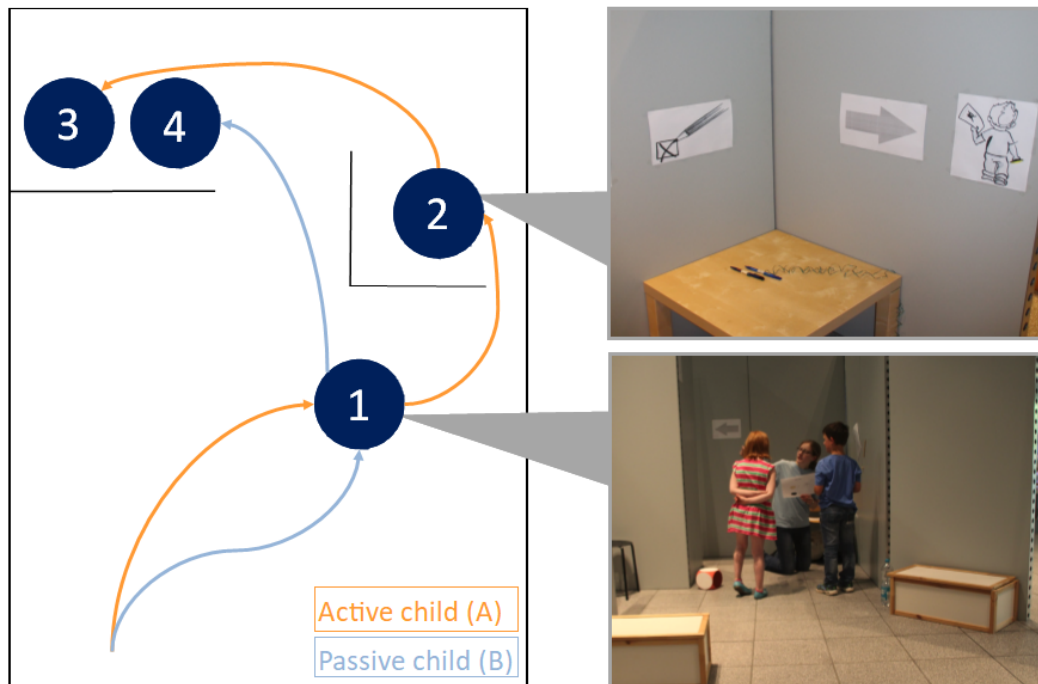
Figure 4.1: Setup of the experiment with positions of participating children

At Position 1, the experimenter explains the entire set of instructions to both children, carefully ensuring that both understand them fully. Pictures on the walls of the experimental area are used to support the instructions (Figure 4.1).

In order to avoid asymmetric information between the participants joining the experiment at different times of the day, due to early participants showing off their payoffs to future participants, we take another simple measure. All participants are shown the potential payoffs upfront, as part of being instructed to participate in the experiment. The payoff consists of a fixed payoff in form of a small die. Additionally, one child from each pair of children can win a prize. The prize is a rainbow color pencil, displayed in Appendix Figure 4.8.

The active child is asked to enter a private corner (Position 2), where no one can observe her actions. The experimenter makes sure that she knows she will not be observed by anyone while carrying out her task in the private corner. She is given a fair yellow-and-black coin, along with instructions to toss the coin in private. She is asked to report the outcome by a mark on the report sheet (Figure 4.2) before leaving the private corner. Hence, the active child has the possibility to report any color, also the one different from the actual outcome, without any risk of being detected. Next, she hands the report sheet over to a second experimenter, waiting at the exit of the

private corner (Position 4).[1]

Winning the prize has different consequences in the control compared to the treatment group: In case of a win in the control group, the prize is obtained by the active child. In the treatment group, the prize is obtained by the passive child instead. Hence, in the treatment group the active child cannot influence her own payoff, but only that of the other child.
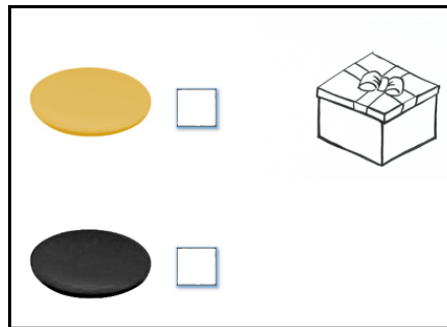

Figure 4.2: Report sheet, on which the active child is asked to mark yellow or black in the private corner

While the active child is carrying out the main task, the passive child receives a simple assistant task. This is important in order to ensure that both children take the experiment serious, and are not bored or confused. Thus, we use this assistant task to integrate and involve the passive child without impacting or disturbing the active child's behavior. This assistant task consists in walking around the private corner and bringing an empty form to the second experimenter, waiting at Position 4. Hence, both children meet again, once the active child finishes her task (moving from Position 3 to 4).[2]

Next, the second experimenter, after receiving the report sheet and with both children present, evaluates the outcome reported on the sheet. In case of a yellow report in the control group, only the active child is then awarded the rainbow colored pencil. In case of a yellow report in the treatment group, only the passive child is then awarded the rainbow color pencil. In case of a black report, in both the control and treatment group, no one receives the prize. The possible choices and resulting payoffs are displayed in form of a game tree in Appendix Figure 4.11.

Finally, the experimenter asks a set of control questions and demographic questions printed on the form brought by the passive child. The latter enables

---

[1]Within the private corner, additional graphical instructions, displayed in Appendix Figures 4.9 and 4.10, remind the active child of what to do.

[2]In most cases, the active child's task did not take longer than a minute, such that there was no long waiting time for the passive child.

analyzing gender- age- and relationship specific treatment effects. The experimenter marks the answers on the form. He awards both children with a small die, as well as a stamp on their hand. Giving the children a stamp mark comes with the advantage that the experimenters outside the experimental area can make sure that no child participates in the experiment more than once.

4.4   Implementation and Participants

The experiment was carried out in Paderborn, Germany, in June 2015. A total of 512 children participated in the experiment over the course of three consecutive days. As children participated in pairs, the dataset contains 256 independent observations. Two children left the experiment without reporting a coin flip outcome and before answering the questionnaire, which is why we exclude these two observations from the dataset for any of the following calculations. Demographic characteristics of the remaining sample of 254 active children are displayed in Table 4.1. We will refer to these 254 participants, out of the total of 512 participants, as our 'subjects'.

	Control Group	Treatment Group	
Observations (254 subjects)	120	134	
Age (min 3, max 15)	7.31	8.10	**
Female (%)	51.26	42.54	
Paired with female (%)	59.66	48.87	*
Partner known (%)	67.22	74.63	
Paired with sibling (%)	24.58	14.93	*

Table 4.1: differences between control and treatment group (t-test); level of significance denoted * p<0.1, ** p<0.05, *** p<0.01

Running the experiment in the field allowed us to gather observations not only from pairs of strangers, but also from pairs of friends and even siblings. Comparing demographic characteristics of the subjects using a t-test reveals that subjects in the treatment group are significantly older (p=0.01), and weakly significantly less often paired with a sibling (p=0.05) or with a female partner (p=0.09). When evaluating the data, we will thus control for these three variables to make sure that potential differences between the control and treatment group are not only driven by these differences. Regarding the gender of the active children, or the share of children who were paired with a child already known to them, there are no notable differences

between the treatment and control group. Here, randomization was successful. Despite a few imperfections in the randomization, we still obtained a rich dataset with the desired heterogeneity in participant characteristics. There are sufficient observations of each type of participant in both the control and the treatment group to allow testing for type-specific effects.

4.5 Results and Discussion

First, we analyze whether we find evidence for cheating in the overall sample within the control and treatment group, and compare the two. Next, we go into more detail and describe how results depend on the relationship between the children, as well as on age.

4.5.1 Central Results

Out of the 134 subjects in the treatment group, 64 subjects, a share of 47.76%, reported the prize winning outcome "yellow". In contrast to that, 81 out of 120 subjects in the control group, a share of 67.50%, reported the prize winning outcome. According to a two-sided binomial test, the share of reported prize winning outcomes in the treatment group is not significantly different from 50%. In contrast, within the control group it is significantly different from 50% ($p<0.05$), which means that a significant share of subjects in the control group is engaging in cheating in order to obtain the prize.

Comparing the "yellow" reports of the treatment group to those of the control group using a Chi-Square test shows that yellow reports are significantly more likely in the control group ($p<0.01$). This difference is illustrated in Figure 4.3, and leads to our first central result:

Result 1. *Subjects are more likely to report the price winning outcome of a coin flip when the prize benefits themselves, compared to when the prize benefits another child instead.*

We interpret this result with great care, when it comes to drawing conclusions regarding cheating behavior. Cheating, in this case, can go in two directions. First, subjects can falsely report "yellow" in order to win a prize, either for themselves or for someone else. Second, subjects can falsely report "black" in order to avoid winning a prize themselves, or to avoid that someone else receives the prize. In existing literature, the second type of cheating is referred to as downward lying
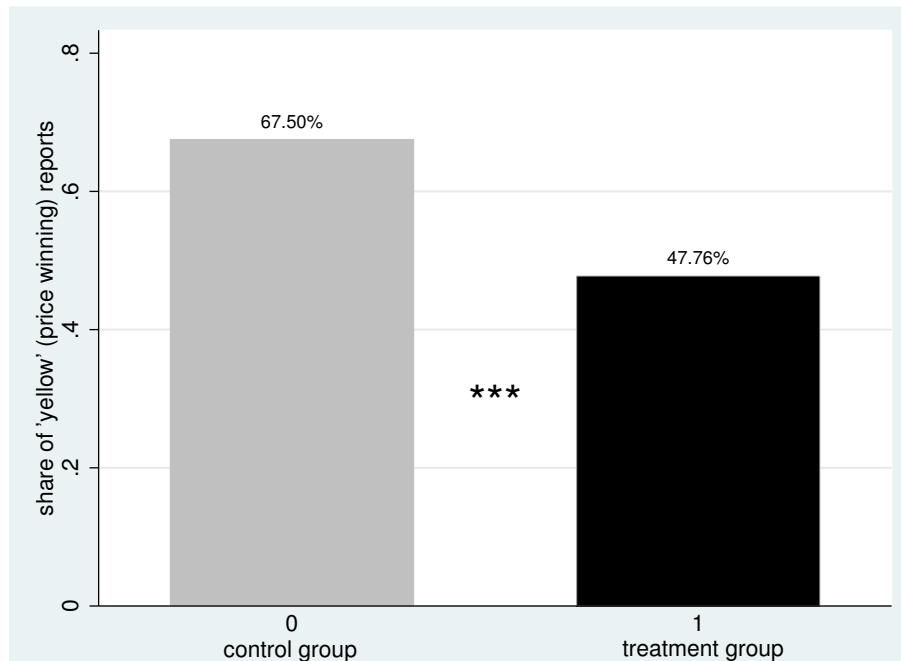
Figure 4.3: Treatment effect: significantly different shares of subjects report the prize winning outcome of a coin flip (*** p<0.01)

or downward cheating (e.g. Houser et al., 2012; Abeler et al., 2014; Utikal and Fischbacher, 2013). When both types of cheating behavior are present, upward and downward cheating cancel each other out, either fully or partly. Therefore, we cannot conclude that subjects in the treatment group are not cheating, even though the prize winning outcome is not reported in more than 50% of the cases. However, we can conclude that there are relatively more cases of upward cheating, compared to downward cheating, in the control group, while this is not the case in the treatment group. Hence, comparing types of cheating across groups, the share of upward to downward cheating is higher in the control, rather than the treatment group.

4.5.2 Further Results

To obtain more detailed results and test for type-specific effects, we run Probit regressions,[3] with the prize winning outcome as the dependent variable. Marginal effects are reported in Table 4.2.

Overall, Probit estimates confirm Result 1. Reporting the prize winning outcome is significantly (p<0.01) less likely in the treatment group, where the prize benefits

---

[3]The number of observations used for the different regressions varies due to a few missing questionnaire entries from some children who were not patient enough to answer all questions.

the partners, rather than the subjects themselves. While the size of this effect varies between 19.4 percentage points without controls and 17.6 percentage points with controls, it is robust to controlling for differences between the control and treatment group described in Section 4.4, regarding age, being paired with a sibling, and being paired with a female (see column 2).

Next, we include interaction terms to test whether the effect differs, depending on whether the subject knows his or her partner, as well as depending on the subject's age, and consider estimates in Table 4.2 Columns 3 and 4. Testing for gender-specific treatment effects did not yield significant results.[4]

	(1)	(2)	(3)	(4)
treatment	-0.194***	-0.176***	-0.348***	-0.559***
	(0.06)	(0.06)	(0.11)	(0.19)
age_subject		-0.011	-0.010	-0.041**
		(0.01)	(0.01)	(0.02)
female_partner		0.097	0.105*	0.101*
		(0.06)	(0.06)	(0.06)
siblings		-0.042		-0.043
		(0.08)		(0.08)
known			-0.064	
			(0.10)	
known_treat			0.249*	
			(0.13)	
age_treat				0.050**
				(0.02)
N	254	250	251	250

Table 4.2: Probit marginal effects with reporting "yellow" as dependent variable. Control variables include the subject's age, the partner's gender and controls for whether the subject and partner are siblings or known to each other. We also control for relationship-specific and age-specific treatment effects by including interactions of the variables known and age with the treatment dummy. * $p<0.1$, ** $p<0.05$, *** $p<0.01$, (se)

First, we consider regression estimates in Column 3. Intuitively, the effect of the partner being a sibling is intuitively stronger than the partner only being someone known. However, there are relatively few siblings in the sample. Moreover, in real live, interactions with known persons or strangers are both more frequent than with siblings and hence, we drop the control variable siblings in favor of the control

---

[4]Further regression results, including interaction terms to test for gender-specific effects, are reported in the Appendix Table 4.3.

variable known. In line with the previous Result 1, we find a highly significant (p<0.01) baseline treatment effect, also when testing for relationship-specific effects. Moreover, it is shown that in the control group whether the passive child is known to the subject does not seem relevant for the decision to report the prize winning outcome. Intuitively, this makes sense since in the control group, partners are not affected by the outcome of the coin flip. In contrast, whether the partner is known or not becomes relevant in the treatment group, where winning benefits the partner: The positive estimate of the interaction between treatment and known (0.249, p<0.1) shows that the treatment effect is reduced when children know each other. In other words, the strongly negative treatment effect is driven by pairs of children who are strangers. The effect is smaller when the children know each other. Thus, subjects are more likely to report the prize winning outcome when the prize is received by somebody they know, rather than a stranger.

Result 2. *The size of the treatment effect depends on the relationship between the subject and her partner. When they know each other, the effect is reduced.*

Second, regression estimates in Column 4 show that in the control group, where winning benefits the participants themselves, older children are significantly (p<0.05) less likely to report the prize winning outcome. This reflects an unintended drawback of the design: Anecdotal evidence from comments by participants suggest that the pencil with rainbow colors, the winning prize, is less attractive to older children. At the same time, estimates again show a highly significant (p<0.01) and negative treatment effect, in line with the previous Results 1 and 2. However, the coefficient of the interaction term shows that there is an age-specific positive treatment effect, hence working in the opposite direction. This means that older subjects are less reluctant to also report the prize winning outcome to benefit somebody else, instead of themselves.

Assuming that the prize to win is less attractive for older subjects, this makes intuitive sense: On the one hand, younger subjects, who are keen on winning the prize themselves, are less wiling to report the prize winning outcome to benefit somebody else. On the other hand, older subjects, who are not so keen on winning the prize themselves, may rather see the opportunity to do somebody else a favor. This leads to our next result:

Result 3. *The size and direction of the treatment effect depend on the age of the subject. The treatment effect is largest among very young subjects and decreases in age. Among the oldest subjects, the effect turns positive: The prize winning outcome is more likely reported when the receiver is another child.*

4.6 Conclusion

In this study we investigated whether children adapt their cheating behavior, dependent on who would benefit from cheating. By changing the recipient of the benefit, we ruled out that cheating happens out of selfish motives. With our results, we contribute to a growing literature on dishonest behavior, shedding light on motives for cheating that may originate during childhood.

We ran a lab in the field experiment with pairs of children as participants. Participants could benefit themselves through cheating (control group), or benefit the partnering child instead (treatment group). We found that there is relatively more upward compared to downward cheating when the participants can benefit themselves, or when they can benefit somebody who is not a stranger. With this finding we contribute evidence, that while most children are either honest or motivated to cheat out of selfishness, some children do cheat out of other motives: At least when the other party was known to them, a significant number of children was altruistically motivated and cheated in favor of someone else.

Moreover, our results contribute evidence in line with existing literature claiming that very young children are self-centered and only start understanding moral concepts such as fairness or equality as they age. This result is shown by the size of the treatment effect decreasing in age. A few observations from older children, who seemed less attracted by the prize, suggest that some may even have engaged in downward cheating to avoid receiving the benefit themselves. Disentangling shares of upward and downward cheating remains an open challenge and limits the explanatory power of our results. More generally, future research is necessary to closer examine the motives of downward cheating, and under which circumstances such behavior, present among adults, is picked up during childhood.

4.7   Appendix

Figure 4.4: entertaining a child in front of the experimental area

Figure 4.5: outside the experimental area

Figure 4.6: wall with pictures and information on random devices

Figure 4.7: two-color die to randomly determine roles in the experiment

Figure 4.8: rainbow-colored pencil

Figure 4.9: graphical instructions within the private corner 1

Figure 4.10: graphical instructions within the private corner 2

	Additional Probit Estimates
treatment	-0.108
	(0.08)
age_subject	-0.012
	(0.01)
female_partner	0.086
	(0.07)
siblings	-0.031
	(0.08)
female_subject	0.112
	(0.09)
female_treat	-0.137
	(0.12)
N	250

Table 4.3: Probit marginal effects with reporting "yellow" as dependent variable. Control variables include the subject's age, the partner's gender and controls for whether the subject and partner are siblings. Other than in Table 2, here we also include a control variable for the subject's gender, and an interaction between treatment and subject gender (female_treat). * p<0.1, ** p<0.05, *** p<0.01, (se)

Figure 4.11: game tree with payouts for participating subjects and passive partners

Summary of the results and outlook for future research

This dissertation is about the role that interpersonal factors play in individual decision-making. In four experimental studies, it presents four different decision situations, in each of which particular types of interpersonal factors make a significant difference. In this section, I summarize the main insights from the empirical results of these studies, showing that each study makes a distinct contribution to existing research. Additionally, some of the results lead to the disclosure of new questions. These questions, providing viable possibilities for future research on related topics, are pointed out as well.

In the first chapter, I investigate whether receiving an endowment in a socially framed situation affects the decision to embezzle. The experiment is designed in such a way that results can be drawn dependent on a specific height of endowment. The social framing in this experiment consists in participants receiving an amount of money from a dictator of a previous, unrelated experiment as their initial endowment, instead of receiving the endowment in a neutrally framed situation.

The results show that it is necessary to distinguish between different cases in order to describe the effect of the social framing: only when receiving a less-than-half share from a dictator, in comparison to receiving the equivalent of less than 5 € in the neutral framing treatment, is embezzlement significantly more likely. This central result is in line with the explanation that self-deception is a driver of the likelihood of embezzlement: observing the dictator's choice to give away less than an equal share, and perceiving this as immoral, may serve some participants as an excuse for their own misconduct. This way, they are able to engage in embezzlement without perceiving their own behavior as immoral. The self-deception explanation is supported by evidence from an additional control variable, namely the participants' guess on the share of peers abstaining from embezzlement.

The most surprising result may be that within the neutral framing treatment, I find the likelihood of embezzlement to be significantly higher the larger the initial endowment. This puzzling incidental finding is producing interesting questions to be investigated in future research: Under which circumstances do higher endowments lead to more immoral behavior? Does the evidence support either the explanation that relatively high wealth produces a feeling of entitlement, or the explanation that possessing more produces greed? Even without answering these questions, in the

end this study still provides a potential explanation for all results taken together: in all cases, participants tend to make their decisions in line with what they perceive to be the behavioral norm — though this norm seems to differ depending on the social framing and height of endowment.

The next two chapters are still concerned with the impact of interpersonal experiences but in a very different decision situation. First, chapter two reviews a remedy against the forwarding of unkind behavior, which we refer to as chain of unkindness. The question at the core of this chapter is whether the remedy, namely writing a letter to the person who first started the unkind action, is actually working because it helps participants to 'close the case' by ruling out an alternative explanation: instead of facilitating closure, letter writing may instead serve as a signal that imitating the predecessor is not wanted. This is done by designing the experiment such that unkindness can not be passed on within one decision domain, but only across domains, which rules out the possibility for direct imitation. Next, chapter three takes a closer look at the effect of letter writing. Here the aim is to find out whether letter writing truly has a curative effect, overriding the chain that was started by the interpersonal experience, or whether instead it has a general pro-social effect. For this purpose, we take into account the effect that letter writing has on the passing on of kindness, and compare the results to the effect on the passing on of unkindness.

In both chapters, the results draw a clear picture. The findings in chapter two confirm that letter writing leads to a highly significant increase in rates of giving in a subsequent dictator game. Thus, it seems to be effective when unkind behavior may be passed on across domains, and not only help against direct imitation of unkind decisions. However, the results of chapter three take our understanding of the effects of letter writing yet a step further: letter writing does not only have a favorable effect following a negative interpersonal experience, but also following a positive interpersonal experience. More generally speaking, our results show that letter writing has a general pro-social effect. This socially desirable effect is not only highly statistically significant, but it even has a similar magnitude: in both cases, after a positive or negative interpersonal experience, it increases rates of giving by about 35%. We suggest that the pro-social effect of letter writing may stem from the activation of more social modes of thinking while writing. In future research, experiments making use of functional Magnetic Resonance Imaging could provide further evidence for this theory. More specifically, it could be used to test whether letter writing activates the same brain regions that are also active when people behave pro-socially.

Whereas it was not an aim of our studies to find further evidence for the existence of chains of unkindness, the data gathered in our experiment can be used to find out whether in general, without any remedy in place, unkind behavior is passed on.

Interestingly, we do not find evidence for unkind behavior to be passed on in our setting, where unkindness can only be payed forward in a different decision situation. Therefore, we suggest that future experiments test whether chains of unkindness may in general not continue when the decision situation is altered down the chain. More precisely, previous experimental studies on the chain of unkindness could be replicated, including a variation of the treatment where unkindness can only be passed on in a different decision situation.

Just as the previous chapters, the fourth chapter is also concerned with the impact of interpersonal factors, and whether they rather enhance socially desired or socially undesired behavior, in other words immoral behavior. More specifically, we examine the circumstances under which children decide to engage in cheating. We vary the receiver from the prize that can be obtained through cheating in a coin toss game. This allows us to investigate whether the motive to altruistically cheat, despite the absence of individual profits, already exists among children. This study is a lab-in-the-field experiment, which comes with the advantage of a heterogeneous subject sample, including children of different ages as well as relationships to each other. This diversity enables us to also describe in which way the results depend on different characteristics of the children.

The central result of this fourth chapter is that subjects are significantly more likely to report the price winning outcome of a coin flip when the prize benefits themselves in comparison to when the prize benefits another child. We estimate that the average treatment effect lies somewhere between 19.4 and 17.6 percentage points in the likelihood of reporting the prize winning outcome. Taking into account the different characteristics of the children, we are able to describe our result in even more detail. We find that the size of the treatment effect depends on the relationship between the child and her partner. When the children know each other, they seem to make less of a difference between cheating for their own and cheating for the other child's benefit. Last, we find that both the size and the direction of the effect depend on age. While the youngest children seem most selfish and are a lot less likely to cheat for the benefit of someone else, this effect fades out with age. According to our estimates, altruistic cheating may even be more popular than selfish cheating among the oldest participants. Partly, this may be due to the prize not being equally attractive for children of different ages, which could be improved in future experiments.

The biggest open question that remains is how to design future experiments in a way that allows estimating actual rates of cheating. In our study, both cheating upward and cheating downward is possible. Consequently, some may cheat in order to obtain the prize while others cheat in order to avoid the prize, and we cannot reveal downward cheating on the aggregate level. Future research is thus necessary to test our suggestion that very young children may be cheating downwards, in order to avoid other children to receive a prize.

Bibliography

Abeler, J., A. Becker, and A. Falk
  2014. Representative evidence on lying costs. *Journal of Public Economics*, 113:96–104.

Agnew, R., N. L. Piquero, and F. T. Cullen
  2009. General strain theory and white-collar crime. In *The Criminology of White-collar Crime*, Pp. 35–60. Springer.

Akerlof, G. A. and R. E. Kranton
  2000. Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Alos-Ferrer, C. and K. H. Schlag
  2009. Imitation and learning. In *The handbook of rational and social choice : an overview of new foundations and applications*, P. Anand, ed., Pp. 271–297. Oxford: Oxford University Press.

Andreoni, J.
  1995. Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics*, 110(1):1–21.

Andreoni, J. and B. D. Bernheim
  2009. Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, 77(5):1607–1636.

Andreoni, J., B. Erard, and J. Feinstein
  1998. Tax compliance. *Journal of Economic Literature*, 36(2):818–860.

Andreoni, J. and L. Vesterlund
  2001. Which is the fair sex? gender differences in altruism. *The Quarterly Journal of Economics*, 116(1):293–312.

Apesteguia, J., S. Huck, and J. Oechssler
  2007. Imitation theory and experimental evidence. *Journal of Economic Theory*, 136(1):217–235.

Attanasi, G., C. Rimbaud, and M. C. Villeval
2018. Embezzlement and guilt aversion. *SSRN, available at: https://ssrn.com/abstract=3169219 or http://dx.doi.org/10.2139/ssrn.3169219.*

Bahr, G. and T. Requate
2014. Reciprocity and giving in a consecutive three-person dictator game with social interaction. *German Economic Review*, 15(3):374–392.

Bardsley, N.
2008. Dictator game giving: Altruism or artefact? *Experimental Economics*, 11(2):122–133.

Barkan, R., S. Ayal, and D. Ariely
2015. Ethical dissonance, justifications, and moral behavior. *Current Opinion in Psychology*, 6(DEC):157–161.

Becker, G. S.
1993. Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, 101(3):385–409.

Belot, M. and J. Van de Ven
2011. Friendships and favouritism on the schoolground–a framed field experiment. *The Economic Journal*, 121(557):1228–1251.

Ben-Ner, A., L. Putterman, F. Kong, and D. Magan
2004. Reciprocity in a two-part dictator game. *Journal of Economic Behavior & Organization*, 53(3):333–352.

Benndorf, V., H. A. Rau, and C. Solch
2018. Minimizing learning behavior in experiments with repeated real-effort tasks. *Available at SSRN: https://ssrn.com/abstract=2503029.*

Bertenthal, B. I., M. R. Longo, and A. Kosobud
2006. Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2):210.

Bien, N., A. Roebroeck, R. Goebel, and A. T. Sack
2009. The brain's intention to imitate: The neurobiology of intentional versus automatic imitation. *Cerebral Cortex*, 19(10):2338–2351.

Blanken, I., N. van de Ven, and M. Zeelenberg
2015. A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin*, 41(4):540–558.

Bradley, M. M., M. Codispoti, D. Sabatinelli, and P. J. Lang
2001. Emotion and motivation ii: Sex differences in picture processing. *Emotion*, 1(3):300.

Bradley, M. M. and P. J. Lang
1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59.

Bressler, M. S.
2009. The impact of crime on business: A model for prevention, detection & remedy. *Journal of Management and Marketing Research*, 2(1):12–20.

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

Buchmann, M., M. Gummerum, M. Keller, and T. Malti
2009. Children's moral motivation, sympathy, and prosocial behavior. *Child Development*, 80(2):442–460.

Byrne, R. W. and A. E. Russon
1998. Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21(5):667–684.

Camerer, C. F.
2003. *Behavioral Game Theory*. Princeton: Princeton University Press.

Cappelen, A. W., T. Halvorsen, E. Ø. Sørensen, and B. Tungodden
2017. Face-saving or fair-minded: What motivates moral behavior? *Journal of the European Economic Association*, 15(3):540–557.

Capra, C. M.
2004. Mood-driven behavior in strategic interactions. *The American Economic Review*, 94(2):367–372.

Capraro, V. and A. Marcelletti
2014. Do good actions inspire good actions in others? *Scientific Reports*, 4:7470.

Carver, C. S. and M. F. Scheier
1998. *On the Self-Regulation of Behavior*. New York: Cambridge University Press.

Chen, Y. and S. X. Li
2009. Group identity and social preferences. *American Economic Review*, 99(1):431–57.

Chlaß, N., L. Gangadharan, and K. Jones
2015. Charitable giving and intermediation. Technical report, Jena Economic Research Papers.

Chytilova, J. and V. Korbel
2014. Individual and group cheating behavior: A field experiment with adolescents. Technical report, IES Working Paper.

Cialdini, R. B., P. K. Petrova, and N. J. Goldstein
2004. The hidden costs of organizational dishonesty: companies that engage in unethical practices face consequences far more harmful than is traditionally recognized. the resulting damage can easily outweigh the short-term gains. *MIT Sloan Management Review*, 45(3):67–74.

Cojoc, D. and A. Stoian
2014. Dishonesty and charitable behavior. *Experimental Economics*, 17(4):717–732.

Conlisk, J.
1980. Costly optimizers versus cheap imitators. *Journal of Economic Behavior & Organization*, 1(3):275–293.

Conrads, J., B. Irlenbusch, R. M. Rilke, and G. Walkowitz
2013. Lying and team incentives. *Journal of Economic Psychology*, 34:1–7.

Conrads, J. and S. Lotz
2015. The effect of communication channels on dishonest behavior. *Journal of Behavioral and Experimental Economics*, 58:88–93.

Croson, R., F. Handy, and J. Shang
2009. Keeping up with the joneses: The relationship of perceived descriptive social norms, social information, and charitable giving. *Nonprofit Management and Leadership*, 19(4):467–489.

Croson, R. and J. Y. Shang
2008. The impact of downward social information on contribution decisions. *Experimental Economics*, 11(3):221–233.

Crossan, M., D. Mazutis, and G. Seijts
2013. In search of virtue: The role of virtues, values and character strengths in ethical decision making. *Journal of Business Ethics*, 113(4):567–581.

Crumpler, H. and P. J. Grossman
2008. An experimental test of warm glow giving. *Journal of Public Economics*, 92(5-6):1011–1021.

d'Adda, G., V. Capraro, and M. Tavoni
2017. Push, don't nudge: Behavioral spillovers and policy instruments. *Economics Letters*, 154:92–95.

Damasio, A. R.
1994. *Descartes error: Emotion, rationality and the human brain*. New York (Grosset/Putnam).

Day, R. H.
1984. Disequilibrium economic dynamics: A post-schumpeterian contribution. *Journal of Economic Behavior & Organization*, 5(1):57–76.

Delfino, A., L. Marengo, and M. Ploner
2016. I did it your way. an experimental investigation of peer effects in investment choices. *Journal of Economic Psychology*, 54:113–123.

Diekmann, A.
2004. The power of reciprocity: Fairness, reciprocity, and stakes in variants of the dictator game. *Journal of Conflict Resolution*, 48(4):487–505.

Doris, J. M.
2002. *Lack of character: Personality and moral behavior*. Cambridge University Press.

Dreber, A., E. von Essen, and E. Ranehill
2014. Gender and competition in adolescence: Task matters. *Experimental Economics*, 17(1):154–172.

Dreher, A. and T. Herzfeld
2005. The economic costs of corruption: A survey and new evidence.

Dufwenberg, M. and A. Muren
2006. Generosity, anonymity, gender. *Journal of Economic Behavior & Organization*, 61(1):42–49.

Duval, S. and R. A. Wicklund
1972. *A theory of objective self awareness*. Academic Press.

Eckel, C. C. and P. J. Grossman
1996. Altruism in anonymous dictator games. *Games and Economic Behavior*, 16(2):181–191.

Eckel, C. C. and P. J. Grossman
  1998. Are women less selfish than men? evidence from dictator experiments. *Economic Journal*, 108(448):726–35.

Engel, C.
  2011. Dictator games: A meta study. *Experimental Economics*, 14(4):583–610.

Erat, S. and U. Gneezy
  2012. White lies. *Management Science*, 58(4):723–733.

Falco, S. D., B. Magdalou, D. Masclet, M. C. Villeval, and M. Willinger
  2016. Can transparency of information reduce embezzlement? experimental evidence from tanzania. *Working Paper 1618 Lyon Saint-Etienne*.

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

Fan, C. S., C. Lin, and D. Treisman
  2010. Embezzlement versus bribery. Technical report, National Bureau of Economic Research.

Fareri, D. S., M. A. Niznikiewicz, V. K. Lee, and M. R. Delgado
  2012. Social network modulation of reward-related signals. *Journal of Neuroscience*, 32(26):9045–9052.

Fehr, E. and A. Falk
  2002. Psychological foundations of incentives. *European Economic Review*, 46:687–724.

Fehr, E. and U. Fischbacher
  2004. Third-party punishment and social norms. *Evolution and Human Behavior*, 25(2):63–87.

Fehr, E., D. Glätzle-Rützler, and M. Sutter
  2013. The development of egalitarianism, altruism, spite and parochialism in childhood and adolescence. *European Economic Review*, 64:369–383.

Fischbacher, U.
  2007. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fischbacher, U. and F. Föllmi-Heusi
  2013. Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547.

Fong, C. M.
2007. Evidence from an experiment on charity to welfare recipients: Reciprocity, altruism and the empathic responsiveness hypothesis. *The Economic Journal*, 117(522):1008–1024.

Fukui, H. and K. Toyoshima
2014. Chill-inducing music enhances altruism in humans. *Frontiers in Psychology*, 5:1215.

Gibson, R., C. Tanner, and A. F. Wagner
2013. Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review*, 103(1):532–48.

Gino, F., S. Ayal, and D. Ariely
2013. Self-serving altruism? the lure of unethical actions that benefit others. *Journal of Economic Behavior & Organization*, 93:285–292.

Gino, F., M. E. Schweitzer, N. L. Mead, and D. Ariely
2011. Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, 115(2):191–203.

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

Gneezy, U.
2005. Deception: The role of consequences. *American Economic Review*, 95(1):384–394.

Gneezy, U., A. Imas, and K. Madarasz
2014. Conscience accounting: Emotion dynamics and social behavior. *Management Science*, 60(11):2645–2658.

Gouldner, A. W.
1960. The norm of reciprocity: A preliminary statement. *American Sociological Review*, Pp. 161–178.

Gray, K., A. F. Ward, and M. I. Norton
2014. Paying it forward: generalized reciprocity and the limits of generosity. *Journal of Experimental Psychology: General*, 143(1):247.

Greenberg, J.
2002. Who stole the money, and when? individual and situational determinants

of employee theft. *Organizational Behavior and Human Decision Processes*, 89(1):985–1003.

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

Grossman, M. and W. Wood
1993. Sex differences in intensity of emotional experience: A social role interpretation. *Journal of Personality and Social Psychology*, 65(5):1010.

Hare, T. A., C. F. Camerer, D. T. Knoepfle, J. P. O'Doherty, and A. Rangel
2010. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *Journal of Neuroscience*, 30(2):583–590.

Hassan, S. and F. A. Jan
2017. Psychological capital intervention towards perceived white collar crime and employee's job related attitude in pakistan. *Journal of Managerial Sciences*, 11(2).

Herne, K., O. Lappalainen, and E. Kestilä-Kekkonen
2013. Experimental comparison of direct, general, and indirect reciprocity. *The Journal of Socio-Economics*, 45:38–46.

Heyes, C.
2011. Automatic imitation. *Psychological Bulletin*, 137(3):463.

Hochstetler, A., K. R. Kerley, and K. Mason
2002. Structural predictors of embezzlement: A preliminary analysis. *Journal of Crime and Justice*, 25(1):1–22.

Houser, D., J. A. List, M. Piovesan, A. Samek, and J. Winter
2016. Dishonesty: From parents to children. *European Economic Review*, 82:242–254.

Houser, D. and D. Schunk
2009. Social environments with competitive pressure: Gender effects in the decisions of german schoolchildren. *Journal of Economic Psychology*, 30(4):634–641.

Houser, D., S. Vetter, and J. Winter
2012. Fairness and cheating. *European Economic Review*, 56(8):1645–1655.

Irlenbusch, B. and M. C. Villeval
  2015. Behavioral ethics: How psychology influenced economics and how economics might inform psychology. *Current Opinion in Psychology*, 6:87–92.

Jacob, B. A. and S. D. Levitt
  2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3):843–877.

Jiang, T.
  2013. Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior and Organization*, 93:328–336.

Keizer, K., S. Lindenberg, and L. Steg
  2008. The spreading of disorder. *Science*, 322(5908):1681–1685.

Killen, M. and J. Smetana
  2008. Moral judgment and moral neuroscience: Intersections, definitions, and issues. *Child Development Perspectives*, 2(1):1–6.

Kimbrough, E. O. and J. P. Reiss
  2012. Measuring the distribution of spitefulness. *PloS one*, 7(8):e41812.

Kirchler, M., J. Huber, M. Stefan, and M. Sutter
  2015. Market design and moral behavior. *Management Science*, 62(9):2615–2625.

Kirchsteiger, G., L. Rigotti, and A. Rustichini
  2006. Your morals might be your moods. *Journal of Economic Behavior & Organization*, 59(2):155–172.

Klein, S. A., I. Thielmann, B. E. Hilbig, and I. Zettler
  2017. Between me and we: The importance of self-profit versus social justifiability for ethical decision making. *Judgment & Decision Making*, 12(6).

Krach, S., F. Hegel, B. Wrede, G. Sagerer, F. Binkofski, and T. Kircher
  2008. Can machines think? interaction and perspective taking with robots investigated via fmri. *PloS one*, 3(7):e2597.

Leighton, J., G. Bird, C. Orsini, and C. Heyes
  2010. Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology*, 46(6):905–910.

Leimgruber, K. L., A. F. Ward, J. Widness, M. I. Norton, K. R. Olson, K. Gray, and L. R. Santos
  2014. Give what you get: Capuchin monkeys (cebus apella) and 4-year-old

children pay forward positive and negative outcomes to conspecifics. *PloS one*, 9(1).

Lindebaum, D., D. Geddes, and Y. Gabriel
2017. Moral emotions and ethics in organisations: Introduction to the special issue. *Journal of Business Ethics*, 141(4):645–656.

List, J. A.
2007. On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493.

Luccasen, A. and P. J. Grossman
2017. Warm-glow giving: Earned money and the option to take. *Economic Inquiry*, 55(2):996–1006.

Maggian, V. and M. C. Villeval
2016. Social preferences and lying aversion in children. *Experimental Economics*, 19(3):663–685.

Makowsky, M. D. and S. Wang
2018. Embezzlement, whistleblowing, and organizational architecture: An experimental investigation. *Journal of Economic Behavior & Organization*, 147:58–75.

Marx, K.
1976. *Preface and Introduction to a Contribution to the Critique of Political Economy*. Foreign Languages Press.

Matthey, A.
2010. Imitation with intention and memory: An experiment. *The Journal of Socio-Economics*, 39(5):585–594.

Mazar, N., O. Amir, and D. Ariely
2008. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6):633–644.

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

McCabe, D. L. and L. K. Trevino
1997. Individual and contextual influences on academic dishonesty: A multicampus investigation. *Research in Higher Education*, 38(3):379–396.

McCabe, D. L., L. K. Trevino, and K. D. Butterfield
   2001. Cheating in academic institutions: A decade of research. *Ethics & Behavior*, 11(3):219–232.

Morishima, Y., D. Schunk, A. Bruhin, C. C. Ruff, and E. Fehr
   2012. Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, 75(1):73–79.

Motro, D., L. D. Ordóñez, A. Pittarello, and D. T. Welsh
   2016. Investigating the effects of anger and guilt on unethical behavior: a dual-process approach. *Journal of Business Ethics*, Pp. 1–16.

Mujcic, R. and A. Leibbrandt
   2017. Indirect reciprocity and prosocial behaviour: Evidence from a natural field experiment. *The Economic Journal*.

Nelissen, R. M. and M. Zeelenberg
   2009. Moral emotions as determinants of third-party punishment: Anger, guilt and the functions of altruistic sanctions. *Judgment and Decision making*, 4(7):543.

Nicholson, L. G.
   1999. *Security Investigations: A Professional's Guide*. Butterworth-Heinemann.

Nonis, S. and C. O. Swift
   2001. An examination of the relationship between academic dishonesty and workplace dishonesty: A multicampus investigation. *Journal of Education for Business*, 77(2):69–77.

Nowak, M. A. and K. Sigmund
   2005. Evolution of indirect reciprocity. *Nature*, 437(7063):1291–1298.

Nunner-Winkler, G. and B. Sodian
   1988. Children's understanding of moral emotions. *Child Development*, Pp. 1323–1338.

Nye, J. S.
   1967. Corruption and political development: A cost-benefit analysis. *American Political Science Review*, 61(2):417–427.

Oechssler, J.
   2010. Searching beyond the lamppost: Let's focus on economically relevant questions. *Journal of Economic Behavior & Organization*, 73(1):65–67.

Oechssler, J., A. Roider, and P. W. Schmitz
  2015. Cooling off in negotiations: Does it work? *Journal of Institutional and Theoretical Economics*, 171(4):565–588.

Pingle, M.
  1995. Imitation versus rationality: An experimental perspective on decision making. *The Journal of Socio-Economics*, 24(2):281–315.

Pittarello, A., M. Leib, T. Gordon-Hecker, and S. Shalvi
  2015. Justifications shape ethical blind spots. *Psychological Science*, 26(6):794–804.

Pruckner, G. J. and R. Sausgruber
  2013. Honesty on the streets: A field study on newspaper purchasing. *Journal of the European Economic Association*, 11(3):661–679.

Rand, D. G. and M. A. Nowak
  2013. Human cooperation. *Trends in Cognitive Sciences*, 17(8):413–425.

Rankin, D. J. and M. Taborsky
  2009. Assortment and the evolution of generalized reciprocity. *Evolution*, 63(7):1913–1922.

Reinstein, D. and G. Riener
  2012. Decomposing desert and tangibility effects in a charitable giving experiment. *Experimental Economics*, 15(1):229–240.

Rilling, J. K., D. A. Gutman, T. R. Zeh, G. Pagnoni, G. S. Berns, and C. D. Kilts
  2002. A neural basis for social cooperation. *Neuron*, 35(2):395–405.

Rilling, J. K. and A. G. Sanfey
  2011. The neuroscience of social decision-making. *Annual Review of Psychology*, 62:23–48.

Rilling, J. K., A. G. Sanfey, J. A. Aronson, L. E. Nystrom, and J. D. Cohen
  2004. Opposing bold responses to reciprocated and unreciprocated altruism in putative reward pathways. *Neuroreport*, 15(16):2539–2243.

Robert, I. and M. Arnab
  2013. Is dishonesty contagious? *Economic Inquiry*, 51(1):722–734.

Rosenbaum, S. M., S. Billinger, and N. Stieglitz
  2014. Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology*, 45:181–196.

Schachter, S. and B. Latane
  1964. Crime, cognition, and the autonomic nervous system. In *Nebraska Symposium on Motivation*. University of Nebraska Press.

Schnatterly, K.
  2003. Increasing firm value through detection and prevention of white-collar crime. *Strategic Management Journal*, 24(7):587–614.

Schnedler, W. and N. L. Stephan
  2019. Revisiting a remedy against the chain of unkindness. Working Papers Dissertations 45, Paderborn University, Faculty of Business Administration and Economics.

Schuessler, K. F. and D. R. Cressey
  1950. Personality characteristics of criminals. *American Journal of Sociology*, 55(5):476–484.

Schurr, A. and I. Ritov
  2016. Winning a competition predicts dishonest behavior. *Proceedings of the National Academy of Sciences*, 113(7):1754–1759.

Schweitzer, M. E. and D. E. Gibson
  2008. Fairness, feelings, and ethical decision-making: Consequences of violating community standards of fairness. *Journal of Business Ethics*, 77(3):287–301.

Shalvi, S., F. Gino, R. Barkan, and S. Ayal
  2015. Self-serving justifications: Doing wrong and feeling moral. *Current Directions in Psychological Science*, 24(2):125–130.

Stanca, L.
  2009. Measuring indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*, 30(2):190–202.

Stigler, G. J.
  1962. Information in the labor market. *Journal of Political Economy*, 70(5, Part 2):94–105.

Strang, S., X. Grote, K. Kuss, S. Q. Park, and B. Weber
  2016. Generalized negative reciprocity in the dictator game–how to interrupt the chain of unfairness. *Scientific Reports*, 6:22316.

Tan, H. B. and J. P. Forgas
  2010. When happiness makes us selfish, but sadness makes us fair: Affective influences on interpersonal strategies in the dictator game. *Journal of Experimental Social Psychology*, 46(3):571–576.

Thielmann, I. and B. E. Hilbig
2018. No gain without pain: The psychological costs of dishonesty. *Journal of Economic Psychology*.

Toca Torres, C. E. and J. C. Rodríguez
2013. Donating and volunteering as social capital in bogota. *Bulletin of Latin American Research*, 32(2):133–148.

Tyler, T. R. and S. L. Blader
2005. Can businesses effectively regulate employee conduct? the antecedents of rule following in work settings. *Academy of Management Journal*, 48(6):1143–1158.

Utikal, V. and U. Fischbacher
2013. Disadvantageous lies in individual decisions. *Journal of Economic Behavior & Organization*, 85:108–111.

van Apeldoorn, J. and A. Schram
2016. Indirect reciprocity; a field experiment. *PloS one*, 11(4):e0152076.

Whitley, B. E.
1998. Factors associated with cheating among college students: A review. *Research in Higher Education*, 39(3):235–274.

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

Xiao, E. and D. Houser
2005. Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20):7398–7401.

Xiao, E. and D. Houser
2009. Avoiding the sharp tongue: Anticipated written messages promote fair economic exchange. *Journal of Economic Psychology*, 30(3):393–404.