# Four Economic Experiments on Social Preferences

Inauguraldissertation
zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften
durch die
Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von

**Jana Hofmeier**

aus Haan

Bonn 2019

| | |
|---|---|
| Dekan: | Prof. Dr. Jürgen von Hagen |
| Erstreferent: | Prof. Dr. Lorenz Götte |
| Zweitreferent: | Prof. Dr. Thomas Dohmen |

Tag der mündlichen Prüfung: 24. September 2019

# Acknowledgements

Writing this thesis has been a challenging and yet rewarding experience which would not have been possible without the advice and support of many people whom I want to thank.

First of all, I am very grateful to my supervisors Lorenz Götte, Thomas Dohmen, and Florian Zimmermann for many interesting and motivating discussions, their enthusiasm and for creating a supportive and inspiring environment. Great thanks also go to my co-authors. I want to thank Frederik Schwerter for supporting me from the beginning and for introducing me to many interesting research topics. I also thank Anke Becker, Louis Strang, and Thomas Neuber for the many shared hours of discussions and productive work. Special thanks also go to Holger Gerhardt for his technical support. I also want to thank Silke Kinzig, Britta Altenburg, Andrea Reykers, and Simone Jost-Mussehl for doing a wonderful job in managing the Bonn Graduate School of Economics and the Institute for Applied Micro Economics, always offering support and finding solutions when needed.

I have been very fortunate in being surrounded by so many kind and inspiring fellow graduate students during my PhD. They have never made it feel like a competition but like something that we experience and achieve together. I also want to thank my friends Gašper, Laura, and Svea for always being there, for sharing my complaints and celebrating my successes, for taking my mind off things and for creating wonderful memories that I will never forget.

I want to thank my parents, Anita and Rolf, and my sister Luisa for their unconditional love and the unwavering trust and confidence they have always placed in me. My biggest thanks goes to Thomas whose love, support, and encouragement hast brought me through the most difficult times of the last years. I am very grateful that you are in my life!

# Contents

# List of Figures

# List of Tables

# Introduction

The fact that humans are at their core social beings who need and thrive on the interaction with others has received increasing attention during the last two decades. The neoclassical model that had evolved during the 20th century propagated the concept of a *homo oeconomicus* who is assumed to be rational, selfish, and derives utility only from own consumption. This representation of human nature has been challenged in recent years. Research in behavioral economics has provided compelling evidence, largely based on experiments, for the existence of social preferences. Social preferences entail that an agent's utility does not only depend on his or her own outcomes but also on outcomes and behaviors of others. These social preferences have been essential in explaining various important behaviors that standard preferences had so far not been able to account for, such as helping behavior or the provision of public goods (Becker, 1974; Rabin, 1993; Fehr and Schmidt, 1999).

This thesis presents and discusses four different economic experiments investigating the foundations and formation of social preferences. Chapter 1, Chapter 3, and Chapter 4 all take a closer look at prosociality. Prosociality is one manifestation of a social preference and refers to behavior or intentions that benefit others. Chapter 1 investigates helping behavior in a non-monetary domain and asks to which extent individuals' decisions are guided by the consumption preferences of the receiving party and to which degree by their own. Chapter 3 and Chapter 4 both analyze the dynamics of prosocial behavior and investigate if and under which circumstances initial prosocial behavior, in particular altruistic giving, increases future prosocial behavior and, therefore, how the development of prosociality can be fostered. Chapter 2 takes the notion of social preferences to another domain: risk. We investigate whether an individual prefers risky prospects that lead to the same outcomes for her as for another individual to risky prospects that result in negatively correlated outcomes for the two individuals; we also discuss possible reasons and implications.

Chapter 1[1] investigates how people assess the utility of others when making deci-

[1] This chapter is joint work with Thomas Neuber.

sions about providing costly help to others. Do they apply their own preferences or do they adopt the preferences of the other person? We study this question in a laboratory experiment where subjects in the role of senders can pay money to avoid harm arising to receivers. In a first step, we elicit all subjects' willingness to pay (WTP) for not having to eat food items containing dried insects. We then show senders the WTPs of receivers and repeat the elicitation procedure, but now with receivers having to eat the food items and senders stating their WTPs to spare the receivers from having to eat them. We find that not only receivers' preferences matter for decisions but also senders' own preferences, a phenomenon for which we use the term imperfect empathy. In motivating prosocial transfers, senders' and receivers' WTPs act as complements by reinforcing each other. Conversely, pairs of sender and receiver who are dissimilar generate lower transfers than others. Since transfers usually benefit receivers more than they cost senders, we also find that dissimilarity within pairs reduces welfare. Our results complement the extensive literature on prosocial preferences, which so far abstracts from heterogeneous valuations. The implications might be far-reaching. For example, systematic differences in consumption preferences between net payers and recipients could undermine public support for public welfare systems.

Chapter 2[2] studies if people have preferences over the interpersonal correlation between uncertain outcomes. In our laboratory experiment, 69% of subjects show a preference for positively correlated risky outcomes, whereas only 31% prefer negatively correlated risky outcomes. As we predict, women choose the option with positively correlated outcomes significantly more often than men. This finding might contribute to our understanding of important behavioral differences between men and women, e.g., regarding investment in specialized human capital.

In Chapter 3[3], we present a long-term laboratory experiment that aims to shed light on the development of prosocial behavior. More specifically, we investigate the effects of an environment that stimulates prosocial behavior on subsequent levels of prosociality. We hypothesize a positive effect on later prosocial behavior, which should go along with an increase in reported altruistic preferences. However, we find a negative effect of the intervention on subsequent prosociality, and no adjustment of altruistic preferences. We discuss possible interpretations of the results and implications for future research.

Chapter 4[4] presents a study that is tightly linked to the study described in Chapter 3. It studies the effect of observability on the dynamics of prosocial behavior. We hypothesize a twofold positive effect. First, people should act more prosocially when being observed. Second, this increased level of prosociality should motivate an ongoing ele-

[2] This chapter is joint work with Anke Becker and Frederik Schwerter.
[3] This chapter is joint work with Frederik Schwerter.
[4] This chapter is joint work with Louis Strang.

vated altruistic attitude, in accordance with the concept of altruistic capital formation. We test our predictions running two variants of a laboratory experiment in which subjects make a first donation decision either observed or anonymously. Subsequently, all subjects face a second anonymous donation decision. In general, we observe high rates of altruistic behavior. However, we find only weak positive effects of observability on first-stage prosocial behavior and no effects on second-stage prosocial behavior.

In its four chapters, this thesis sheds some new light on the nature of social preferences and on their development at the level of the individual. I believe that both aspects to which I contribute are crucial for understanding economic decision making and human behavior in general. I am also confident that deepening our understanding of social preferences can be of great help in using their many benefits, for the individual as well as for society at large.

## References

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82 (6): 1063–1093.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–1302.

# Chapter 1

# Motivated by Others' Preferences? An Experiment on Imperfect Empathy[1]

> As we have no immediate experience of what other men feel, we can form no idea of the manner in which they are affected, but by conceiving what we ourselves should feel in the like situation. Though our brother is upon the rack, as long as we ourselves are at our ease, our senses will never inform us of what he suffers. They never did, and never can, carry us beyond our own person, and it is by the imagination only that we can form any conception of what are his sensations.
>
> —ADAM SMITH, *THE THEORY OF MORAL SENTIMENTS*[2]

## 1.1 Introduction

It is widely documented that people consider others when making decisions: they donate to charities, give blood, or volunteer. These behaviors have often been attributed to social preferences such as altruism (Becker, 1974, 1976), warm glow (Andreoni, 1990), inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), or reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006). All of these models have in common that the payoff of others is explicitly incorporated into an agent's utility function. Since we are often interested in the distribution of monetary payoffs, the assumption that people know what is good for others is very plausible. But—given that preferences on goods are not homogeneous—the question arises how

[1] This chapter is joint work with Thomas Neuber.
[2] Smith, 1859, Part I, Section I, Chapter I.

these other persons' hedonic benefits, which are not experienced by the agent herself, are transformed into motives for personal prosocial behavior. It has been claimed that empathy, the ability to feel into others' emotions, is playing a central role. According to the empathy–altruism hypothesis, altruistic motivation arises from empathy felt for a person in need (Batson, 1987) and it has been shown empirically that induced empathy indeed increases prosocial behavior (Coke et al., 1978; Klimecki et al., 2016) and cooperation (Batson and Moran, 1999). However, the ability to sympathize with others' emotions is limited. We find evidence that people behave imperfectly empathic: they judge consequences for others not only by the utility that the other person attaches to it, but also by their own preferences.

In this chapter, we show that in order to make a monetary transfer to help another person receive a specific good, two requirements have to be fulfilled: first, the receiver of the good needs to show a preference for the transferred good and second, the sender needs to have a preference for the good as well. This means that people do not only care about the utility that a prosocial action entails for the other person but also which utility they themselves attach to it. We call this type of behavior imperfect empathy (see also Bisin and Verdier, 2001), since people do not only use the other person's preferences to evaluate their actions' consequences on them (perfectly empathic behavior) but also take into account their own preferences.

We run a laboratory experiment in which participants can make prosocial monetary transfers to help other participants. The aim is to find out to which degree people are guided by their own rather than by the receivers' preferences when acting prosocially. Since our interest lies in the emotional accessibility of others' sensations, we use an experimental setup that cleanly isolates such experiences. Following Ambuehl (2017), we let subjects make choices about eating food items that might provoke feelings of disgust, namely dried insects and worms.[3] These "bads" have several important and useful features. First, people have preferences about the consumption of such items. Second, disgust markedly varies between individuals and across items. Third, rational arguments have no power in arguing what is "more disgusting" among the items, e.g., a cricket or a worm. And fourth, we ask people to eat the items and thereby have tight control over consumption.

In Part 1 of the experiment, participants decide how much money between €0 and €20 they themselves would be willing to spend to avoid having to eat several different dried insects. They can decide between receiving a lower payoff and not eating the insect or receiving a high payoff and eating the insect. In Part 2, participants receive information on how much eight other subjects (*receivers*) would each be willing to pay to avoid

[3]The study was approved by the ethics committee of the University of Bonn (reference number: 174/18).

the insects in the first part. The participants (as *senders*) then decide how much money between €0 and €20 they would be willing to spend on sparing these other subjects from having to eat the dried insects. They can decide between receiving a lower payoff and the receiving subject not having to eat the insect or receiving a high payoff and the receiving subject having to eat the insect. We show that not only the receiver's willingness to pay (WTP) for avoiding an item has a positive effect on the respective transfer but also the sender's own WTP, and—in particular—the interaction of the two. Calculating the distance between the vectors of subjects' WTPs, we can also show that dissimilarity between senders and receivers decreases expected transfers. Defining welfare as the sum of individual utility from personal consumption, we can further show that dissimilarity reduces welfare. In the last part of the experiment, subjects have the option to alter decisions, which others have made for themselves, which gives us a measure of paternalism. We show that imperfect empathy is prevalent among both libertarians and paternalists.

This chapter makes contributions to the literature on the role of empathy in generating prosocial behavior. It confirms the hypothesis that empathy is a driver of prosocial behavior, where a lack of empathy can decrease the extent of prosocial actions and lead to misallocation of help from the receiving party's perspective. Our results also inform models featuring altruism in conjunction with heterogeneous preferences like they are present in the literature on the intergenerational transmission of preferences (Bisin and Verdier, 2001; Doepke and Zilibotti, 2017). Our finding could furthermore be an additional explanation for in-group–out-group bias that might exist, e.g., along the lines of political affiliation (see Fowler and Kam, 2007). While this bias might even exist between groups defined by attributes that are arbitrary (Tajfel et al., 1971), it could also be that people have more similar preferences amongst their in-groups. In this case, imperfect empathy could explain why prosocial behavior is stronger towards members of in-groups than towards members of out-groups. Imperfect empathy is also in line with the literature on the false consensus effect (Ross, 1977), a bias in which people commonly think that their own preferences and choices are relatively more common than other preferences and other choices. A potential implication of imperfect empathy is that heterogeneous preferences reduce the support for redistribution and lower expected welfare. It could therefore be an explanation for the finding that diversity has a negative effect on redistribution and donations (Dahlberg et al., 2012; Andreoni et al., 2016) and is therefore meaningful from a policy perspective in ever more diverse societies.

The remainder of this chapter is structured as follows. Section 1.2 presents a simple theoretical framework and derives our hypotheses. Section 1.3 describes the laboratory experiment. Section 1.4 presents the results on the aggregate level, on the level of individuals, and distinguishing between libertarians and paternalists. Finally, Section 1.5

summarizes and discusses the results.

## 1.2 Theory and Hypotheses

We develop a simple theoretical model in which agents derive utility from own consumption as well as from another person's consumption. When evaluating the other person's consumption, agents use a combination of their own and of the other person's preferences. We use the model to formally derive our hypotheses regarding imperfectly empathic behavior and the consequences arising from dissimilar preferences for the size of transfers and for overall welfare.

Individual $i$ experiences utility from good $x_i$ and disutility from "bad" $y_i$; individual $j$ experiences utility from good $x_j$ and disutility from bad $y_j$. Utilities and disutilities are evaluated by utility functions $u$ and $v$ that are specific to the combinations of individuals and domains. In computing overall utility, consumption value from goods enters additively, while disutility from bad experiences is subtracted. We use money as the numéraire. Therefore, utility from money is simply given by the particular nominal amount of currency.[4] If no consumption takes place, we assume that utility is zero. Individuals receive utility not only from their own consumption but also from the other person's consumption. The total utility of subject $i$ is given by the following expression:

$$U_i(x_i, y_i; x_j, y_j) = u_i(x_i) - v_i(y_i) + \alpha\left(u_j(x_j)^{\beta}\, u_i(x_j)^{1-\beta} - v_j(y_j)^{\beta}\, v_i(y_j)^{1-\beta}\right) \tag{1.1}$$

The first part of overall utility, $u_i(x_i) - v_i(y_i)$, is utility and disutility derived from $i$'s own consumption. The remaining term is the utility that individual $i$ derives from the other individual $j$'s consumption. The general extent to which $i$ cares about $j$ is determined by her level of altruism $\alpha$. When evaluating $j$'s utility in a particular domain, $i$ partially relies on both her own relevant utility function and on $j$'s utility function in the respective domain. The degree of reliance on $j$'s preferences is captured by the empathy parameter $\beta \in [0,1]$. If $\beta$ is zero, $i$ simply projects her own preferences upon $j$. If $\beta$ is one, she fully adopts $j$'s preferences and disregards her own.

The notation can of course be extended to further consumption items. We assume in the model above that subjective valuations are complements in generating vicarious (dis-)utility by modeling them multiplicatively, while other authors have assumed perfect substitutability (see, e.g., Bisin and Verdier, 2001). Our assumption means that, in order to enjoy someone else's consumption, both the sender and the receiver have to attach utility to the consumed good, or—conversely—they both have to attach disutility to

[4] We later test the assumption of linear utility from money in our context.

a particular experience to feel that it is bad. Complementarity of assessments gives rise to additional predictions for our experiment, which we develop below and later also test.

## Transfer Decisions

We now apply the utility function in Equation 1.1 to decisions about prosocial transfers in our experiment. In the experiment, subjects receive money, which corresponds to good $x$ above, and potentially eat food items, corresponding to bad $y$. A sender can decide between making a monetary transfer and a receiver not having to eat a food item, and not making a monetary transfer and a receiver having to eat a food item. Sender $i$ never has to consume any food item herself, i.e., $v_i(y_i) = 0$, and receiver $j$ always gets a monetary payoff of €20, i.e., $u_j(x_j) = u_i(x_j) = 20$. The sender can now decide to make a monetary transfer $t \in [0,20]$ so that the receiver does not have to consume item $k \in K$. If the potential transfer of $t \in [0,20]$ (we abstract from discreteness of choice options) for item $k$ is accepted by the sender, the implied monetary payoff for herself is given by $x_i = 20 - t$ and the receiver does not have to eat, i.e., $v_j(y_j) = 0$ and also $v_i(y_j) = 0$. If she rejects, her payoff is $x_i = 20$ and the receiver has to eat item $k$, i.e., $y_j = k$. For a transfer to be made, it has to hold that the utility for the sender when making the transfer (expression on the left hand side of the equation below) is as least as high as the utility when not making the transfer (expression on the right hand side).

$$20 - t + \alpha\, 20 \geq 20 + \alpha \left(20 - v_j(k)^{\beta}\, v_i(k)^{1-\beta}\right)$$

The highest proposed transfer that a sender still accepts, $t^{\star}$ (later simply *transfer*), is therefore given by

$$t^{\star} = \alpha\, v_j(k)^{\beta}\, v_i(k)^{1-\beta} \tag{1.2}$$

Our key hypothesis about decision making can now be formulated directly in terms of the model parameter $\beta$.

**Hypothesis 1.1.** *People typically exhibit imperfect empathy: transfer decisions depend not only on receivers' preferences but also on senders' own preferences. Formally,* $\beta \in (0,1)$.

The above hypothesis can directly be tested by estimating the parameter $\beta$ on the individual level. Moreover, if the hypothesis was true, the partial derivatives of $t^{\star}$ with respect to *both* agents' valuations would be positive, as would be the cross partial derivative. This prediction thus lends itself to reduced-form testing on the level of the subject population, using OLS. We expect transfers to depend positively on both the respective sender's and the receiver's valuations, and—in particular—on their interaction.

## Welfare

In the next step, we theoretically derive predictions about the effect of dissimilarity in preferences between senders and receivers on the size of transfers and on overall welfare. The welfare criterion that we employ is simply the sum of individual utilities from personal consumption.

$$\text{Welfare} \equiv u_i(x_i) - v_i(y_i) + u_j(x_j) - v_j(y_j) \tag{1.3}$$

We predict dissimilarity to decrease welfare through two channels: the size of transfers and the targeting of transfers. The first channel is based on the premise that transfers are on average too low from a planner's perspective. This simply follows from the fact that the planner weighs individuals' welfare equally, while people usually care more about themselves than about others, i.e., $\alpha$ is smaller than one. As we show below, dissimilarity in preferences further decreases the size of transfers and thereby amplifies the welfare loss.

To understand the effect of dissimilarity on the size of transfers, consider two subjects, $i$ and $j$, behaving in accordance with our model and sharing the same parameter values for $\alpha$ and $\beta$. We denote their respective individual valuations of some item by $v_i(k) \equiv v_i$ and $v_j(k) \equiv v_j$, and we fix the total level of the two subjects' valuations of items such that $v_i + v_j \equiv \bar{v}$. Both subjects are with equal probability of 1/2 either sender or receiver. We further assume that $v_i \geq v_j$. This allows us to express the valuations of subjects in terms of the total valuation of both subjects and the distance between the individual valuations: $v_i = \frac{\bar{v}+|v_i-v_j|}{2}$ and $v_j = \frac{\bar{v}-|v_i-v_j|}{2}$. Plugging into Equation 1.2, we can calculate the expected maximum transfer that this pair of subjects generates.

$$\mathbb{E}[t^\star] = \frac{\alpha}{4}\left[\left(\bar{v}+\left|v_i-v_j\right|\right)^{\beta}\left(\bar{v}-\left|v_i-v_j\right|\right)^{1-\beta} + \left(\bar{v}-\left|v_i-v_j\right|\right)^{\beta}\left(\bar{v}+\left|v_i-v_j\right|\right)^{1-\beta}\right] \tag{1.4}$$

Note that, if we had assumed that $v_i \leq v_j$, Equation 1.4 would be identical. During the derivation, only the order of the two summands would reverse. The assumption about which individual has the higher valuation is thus without loss of generality, as follows from the symmetry of the setup.

The expected maximum transfer given by Equation 1.4 is visualized by Figure 1.1 for $\alpha = 1/2$ and $\bar{v} = 20$. Along the x-axis of the graph, we vary the parameter $\beta$, going from a situation where both people fully project their own preferences ($\beta = 0$) to one where they fully adopt others' preferences ($\beta = 1$). On the z-axis, we vary the difference between both subjects' valuations, holding constant the total of the two. The graph starts at the maximum of 20 and ends at a distance of zero, i.e., a situation where both valuations are the same. On the y-axis, the resulting expected maximum transfer $\mathbb{E}[t^\star]$ is

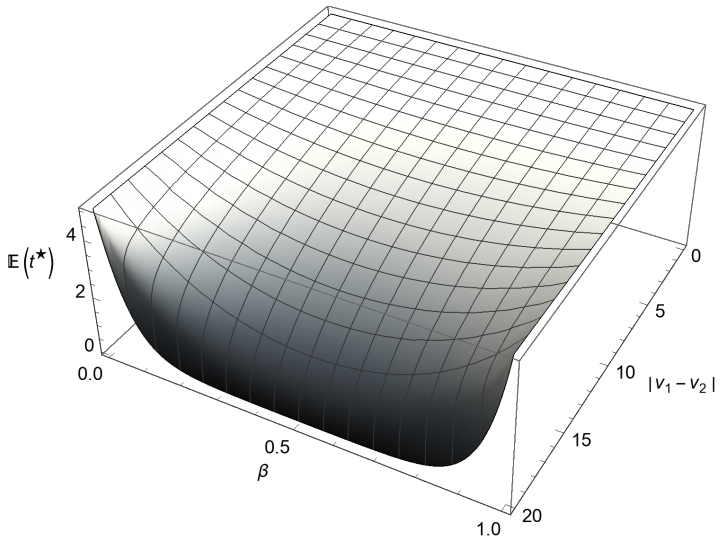



Figure 1.1: Similarity and expected transfers

depicted. If $\beta$ is either zero or one, the expected transfer is always at its maximum value of 5. The same is always the case when the two subjects' valuations coincide. Thus, if we were only talking about money, the degree of empathy would not have any effect on expected transfers, because there would not exist any heterogeneity in valuations. This is, however, only a special case. If, as we expect, $\beta$ typically lies in the interior of the interval from zero to one, dissimilarity in preferences strictly decreases expected transfers, which is our second hypothesis.

**Hypothesis 1.2.** *Transfers decrease with the dissimilarity of preferences within pairs of senders and receivers. Formally,* $\partial \mathbb{E}\left[t^\star\right] / \partial \left|v_i(k) - v_j(k)\right| < 0$.

*Proof.* See Appendix 1.A □

The second channel through which dissimilarity in preferences decreases welfare is saying that—conditional on a given total amount of transfers that a subject is making—senders give for the wrong items. In the extreme case of $\beta = 0$, a sender might be willing to spend a positive amount when the receiver has no problem with eating the respective food item at all, while she gives nothing in case of an item that the receiver finds repulsive. More generally, values of $\beta$ that are smaller than one open up a wedge between how the sender evaluates consequences for the receiver and how the receiver himself—and consequently the social planner—evaluates them. This wedge becomes increasingly relevant as valuations of senders and receivers diverge, leading to misallocation of transfers.[5] We thus arrive at the following hypothesis.

[5] A subtle refinement of the above point would be to consider vicarious experiences—i.e. the sender's

**Hypothesis 1.3.** *The expected net welfare gain from transfer decisions decreases in the dissimilarity of preferences within pairs of senders and receivers. Formally,* $\partial \mathbb{E}\left[\textit{net welfare gain}\right] / \partial \left| v_i(k) - v_j(k) \right| < 0$.

To summarize, we expect that senders base their transfer decisions partially on their own valuations, where the latter and the receiver's own valuation are complements in motivating senders to help. This leads transfers to be on average lower for pairs of senders and receivers who are more dissimilar than others. Reduced transfers and misallocation of existing transfers together lead to welfare losses, which again are larger when subjects are dissimilar.

## 1.3 Experiment

We conducted the experiment at the BonnEconLab in August, September, and December 2018. Subjects were recruited using the software hroot (Bock et al., 2014) and a total of 146 participated. In the invitation email, subjects were asked not to sign up for the experiment in case they were vegetarian, followed a special diet due to health, ethical or religious reasons, or had any food allergy. For details on the composition of our sample of subjects, see the summary statistics in Appendix 1.B. For translated verbal and written instructions, see Appendix 1.C. Before starting the computerized zTree experiment (Fischbacher, 2007), subjects were informed that they might be asked to eat several types of insects during the experiment. They were then shown a tray with all eight different food items (one buffalo worm, five buffalo worms, one meal worm, ten meal worms, one cricket, one grasshopper, three grasshoppers, and one granola bar containing buffalo worms; see Appendix 1.D for pictures of all food items). Furthermore, they received information about the food items' nutritional innocuousness and each participant signed a form of consent.

The experiment consisted of four parts. Subjects received a fixed show-up fee, which was set to be either €5 or €7 for everybody participating in the respective session.[6] In addition, subjects were informed that a single decision among all four parts of the experiment would be randomly chosen for implementation and paid at the end of the experiment. All parts were kept as similar to each other as possible. Always, subjects were endowed with €20 and then used multiple price lists (MPLs) in steps of €1 ranging from

feelings when considering consequences for the receiver—as part of welfare. This would reduce the power of the argument about misallocation but not dismiss it. With imperfect empathy and heterogeneous preferences, there always exists a wedge between vicarious valuations and valuations by people themselves, leading to misallocation.

[6] We varied the show-up fee between sessions in order to test our assumption that utility is linear in money.

€1 to €20 to make payments off this amount. Appendix 1.E includes screenshots of the decision screens of all four parts.

In Part 1, we employed separate MPLs to elicit subjects' reservation prices for not having to eat any of the eight food items. Subjects saw one screen per item (see Figure 1.8 for an example of a decision screen). On each screen, subjects saw an informative stimuli picture of the respective item on the left and a list of choices in the middle of the screen. The list of choices was made up of 20 rows; each row containing the choice between a payment (going from €1 up to €20) and eating the food item. Subjects had to indicate their choice for one of the two options for each row; a unique switching point was enforced. The order in which the eight items were shown was randomized between subjects. In case Part 1 was selected for implementation at the end of the experiment, one of the 160 rows (20 rows each for eight items) was randomly drawn for implementation. If the subject had chosen to pay the amount indicated in the specific row, she received her show-up fee as well as €20 minus the amount indicated in the row as payment. She did then not have to eat the item. If the subject had chosen to eat the item, she received the show-up fee as well as €20 as payment. She furthermore had to eat the item. Subjects who indicated that they would eat the item and refused to do so at the end of the experiment only received their show-up fee.

In Part 2, subjects took the role of a *sender* who had the option to pay for a *receiver* not having to eat a food item. The decision screens were kept very similar to the ones in Part 1 and again contained the same respective stimuli pictures on the left hand side of the MPLs (see Figure 1.9 for an example of a decision screen). On the right side of the screen, subjects additionally saw the WTPs for all eight items that the relevant receiver himself had reported in Part 1. Again, each subject saw eight screens—one for each item. The eight decisions were each made for a different receiver. Receivers were sampled from the pool of subjects taking part in the same session and each participant appeared as a receiver at least once to allow for potential implementations of a decision in this part. However, receivers were sampled in such a way that heterogeneity of WTPs between senders and receivers was larger than in the population of subjects.[7] The assignment of receivers to food items was done without any further sophistication. As in Part 1, subjects had to indicate for each row of the choice list if they chose the payment or the insect. In case Part 2 was selected for implementation at the end of the experiment, one of the 160 rows was randomly drawn for implementation. If the sender had chosen to pay the

[7] Receivers were sequentially sampled among subjects in the same experimental session. For each sender, we made eight independent draws pertaining to a specific criterion, and found the remaining subject who came closest to the respective point. During four sessions, the criterion was the Euclidean distance towards the potential sender's vector of WTPs. In five sessions, it was a vector of WTPs. Note that identification with senders fixed effects—or on the level of the individual sender—uses only variation in WTPs among receivers of a given sender. The latter variation is the result of simple random matching with fixed, equal probabilities.

amount indicated in the specific row, she received her show-up fee as well as €20 minus the amount indicated in the row as payment. The receiver of the row did then not have to eat the item and received his show-up fee and €20. If the sender had chosen not to pay, she received the show-up fee as well as €20 as payment. The receiver furthermore had to eat the item and received his show-up fee and €20. Receivers who refused to eat the item even though their senders had indicated that they would not pay only received their show-up fee.

Part 3 elicited subjects' general level of altruism in the domain of money in a way that mimicked the other parts of the experiment as closely as possible (see Figure 1.10 for an example of a decision screen). As a default, receivers got an amount that was less than €20, mirroring a situation where they had to eat a food item for which they have a certain willingness to pay, and senders could decide whether they wanted to pay amounts from €1 and €20 to secure the receiver €20 instead of €15, €10, €5, or €0. The order of amounts was again randomized. Since we do not need data from Part 3 for identification and hence are not using them, we will not go into more detail here.

Finally, in Part 4, subjects were again shown the same eight receivers as in Part 2. This time, however, they did not decide about engaging in helping behavior but had the option to alter receivers' self-regarding choices from Part 1 without any consequences for themselves. Decision screens looked almost the same as the ones from Part 2 and contained the stimuli picture on the left, the MPL in the middle, and the list of the receiver's WTPs from Part 1 on the right (see Figure 1.11 for an example of a decision screen). However, the MPL already contained the choices that the respective receiver had marked for himself in Part 1. In case Part 4 was selected for payment at the end of the experiment, one of the 160 rows was randomly drawn for implementation for the receiver. If the sender had chosen the payment indicated in the specific row, the receiver received his show-up fee as well as €20 minus the amount indicated in the row as payment. The receiver did then not have to eat the item. If the sender had chosen the item, the receiver received the show-up fee as well as €20 as payment. He furthermore had to eat the item. Receivers who refused to eat the item even though the other subject had not chosen the payment only received their show-up fee.

After all subjects had made their decisions, they were ultimately matched to unilateral pairs of senders and receivers for whom a payoff was implemented. For each subject, a single decision was drawn to be paid out. If Part 1 was implemented for the sender, Part 4 was implemented for the receiver. If Parts 2 or 3 were implemented for the sender, the respective part was also implemented for the receiver. After answering a final survey on the Big Five traits (Gerlitz and Schupp, 2005) and the items of the Interpersonal Reactivity Index, which measures empathy (Davis, 1980), subjects—if necessary—ate their food

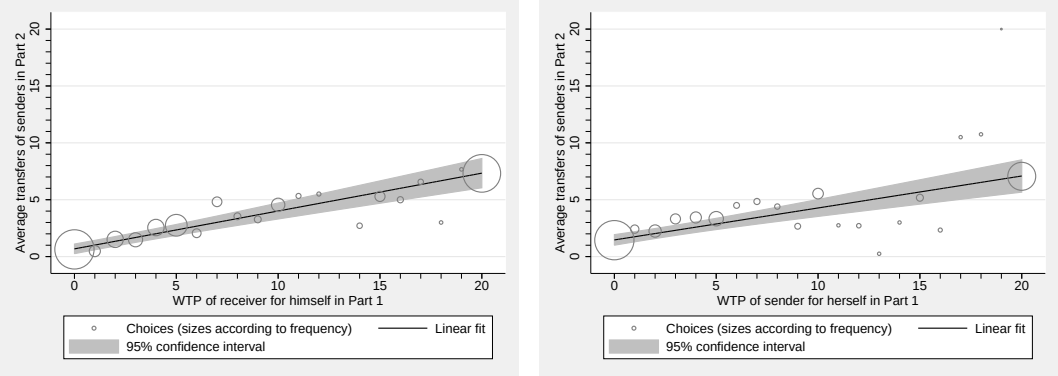
items and then received their payoffs. If subjects did not comply and refused to eat their food items, they were penalized by only receiving the show-up fee.

## 1.4 Results

We start our empirical analysis by estimating the determinants of transfers on the aggregate level by pooling decisions from all individuals. We then proceed by estimating the relationship for each individual separately and recovering individual structural parameters. Next, we turn to the welfare implications by first looking at the effect of dissimilarity on the size of transfers and then directly on net welfare gains. Finally, we show the pervasiveness of imperfect empathy separately among libertarian and paternalist subjects.

### Transfer Decisions

In Part 1, participants spend on average €6.57 per item to avoid eating it; 78% of them have a positive WTP for some item. In Part 2, senders spend on average €3.44 per item to spare receivers from eating it; 75% of senders have a positive WTP for some item–receiver combination. Appendix 1.B shows histograms of WTPs for all items separately for Part 1 and Part 2.




(a) Receiver's WTP (b) Sender's WTP

*Note:* Panel (a) shows the average size of transfers made for a receiver in Part 2 for every possible WTP of the *receiving* subject in Part 1. Panel (b) again shows the average size of transfers made for a receiver in Part 2, this time for every possible WTP of the *sending* subject in Part 1. The lines show fits from OLS estimations, and shaded areas correspond to 95% confidence intervals for standard errors that are clustered at the subject level. Both positive relationships are significant at the 1% level.

Figure 1.2: Individual willingness to pay (WTP) and average transfers

Figure 1.2 visualizes how transfers towards receivers in Part 2 depend on the WTPs of receivers and senders from Part 1. Figure 1.2a shows the average size of transfers made for a receiver in Part 2 for every possible WTP of the receiving subject from Part 1. The

higher the receiver's WTP, the higher is the average transfer made towards him. The positive relationship is statistically significant at the 1% level. Figure 1.2b shows the average size of transfers made for a receiver in Part 2 for every possible WTP of the sending subject from Part 1. The higher the sender's WTP, the higher is the average transfer made towards the receiver. Again, the positive relationship is significant at the 1% level.

To test Hypothesis 1.1, we regress the maximum transfer accepted in Part 2, $t^\star$, on the receiver's willingness to pay $\text{WTP}_{\text{receiver}}$ elicited in Part 1 as well as on the sender's willingness to pay $\text{WTP}_{\text{sender}}$, also elicited in Part 1. Column 1 Table 1.1 shows the results

| | *Dependent variable:* Transfer | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Receiver's WTP | $0.308^{****}$ | $0.309^{****}$ | $0.163^{****}$ | $0.160^{****}$ | $0.100^{***}$ |
| | (0.0311) | (0.0364) | (0.0300) | (0.0354) | (0.0359) |
| Sender's WTP | $0.252^{****}$ | $0.176^{****}$ | 0.0627 | -0.0156 | -0.0578 |
| | (0.0345) | (0.0466) | (0.0394) | (0.0582) | (0.0576) |
| $\sqrt{\text{Sender's} \times \text{receiver's WTP}}$ | | | $0.381^{****}$ | $0.369^{****}$ | $0.364^{****}$ |
| | | | (0.0646) | (0.0693) | (0.0702) |
| Sender fixed effects | No | Yes | No | Yes | Yes |
| Receiver fixed effects | No | Yes | No | Yes | Yes |
| Item fixed effects | No | No | No | No | Yes |
| Observations | 1168 | 1168 | 1168 | 1168 | 1168 |
| Clusters | 146 | 146 | 146 | 146 | 146 |
| (Within-)$R^2$ | 0.362 | 0.197 | 0.417 | 0.285 | 0.171 |

*Note:* OLS regression; standard errors are clustered for senders; $^{*}$ $p<0.1$; $^{**}$ $p<0.05$; $^{***}$ $p<0.01$; $^{****}$ $p<0.001$; robust standard errors in parentheses.

Table 1.1: Aggregate analysis of transfers

without any fixed effects. We see that both WTPs, of the receiver and the sender, enter with large and highly significant coefficients, which in fact are not so different in size. An increase of €1 in the WTP of the receiver increases the transfer on average by €0.31, while the same increase in the sender's WTP increases the average transfer by €0.25. A potential concern could be that senders might differ in their general levels of altruism, and that this variation is systematically related to their own WTPs. Moreover, particular patterns in receivers' WTPs could trigger responses of senders, irrespective of the particular item in question. To rule out such problems, Column 2 adds sender and receiver fixed effects. Due to sender fixed effects, identification only comes from differences in WTPs between receivers of a given sender and from variation in this sender's WTPs across items. Receiver fixed effects allow to account, e.g., for some receivers having generally low WTPs

and receiving low transfers, irrespective of the particular item and the respective sender. The coefficient for the receiver remains almost unchanged, while the coefficient referring to the WTP of the sender somewhat decreases. The latter points to some degree of "spillovers" in empathy: e.g., a sender who feels strong disgust for worms might also better understand why somebody would strongly dislike eating a grasshopper, even if the grasshopper itself does not seem repulsive for the sender. Despite the proximity of preference domains that we use, variation in preferences within individuals is sufficient to show that there is a strong and significant effect of senders' WTPs on transfers. In Column 3, the square root of the product of the sender's and the receiver's WTP is added to the regression without fixed effects.[8] The coefficient of the receiver's WTP drops by half but is still highly significant, whereas the coefficient of the sender's WTP is not significantly different from zero anymore. However, the interaction term enters with a large and highly significant coefficient. This confirms that WTPs of receivers and senders act as complements in generating transfers. Column 4 again adds sender and receiver fixed effects. In Column 5, we additionally add fixed effects for the eight different food items, accounting for differences in the general levels of transfers. In both Columns 4 and 5, coefficients stay similar and the qualitative results remain unchanged.

We show below in Section 1.4 that the above qualitative results also hold within subsamples of the subject populations where senders are restricted to only libertarians or paternalists, respectively. Our empirical results are also insensitive to the size of the show-up fee (see Appendix 1.F), and the assumption of utility from money being linear in the relevant range thus seems innocent. Overall, we find clear support for Hypothesis 1.1. We interpret this as evidence that imperfect empathy is a pervasive phenomenon among our subject population.

## Individual-Level Analysis

In the next step, we analyze behavior at the level of individuals and recover estimates for the model parameters. To do so, we first linearize Equation 1.2 for the size of transfers by taking the logarithm on both sides.

$$\ln\big(t^\star\big) = \ln(\alpha) + \beta \ln\big[v_j(k)\big] + \big(1-\beta\big)\ln[v_i(k)] \tag{1.5}$$

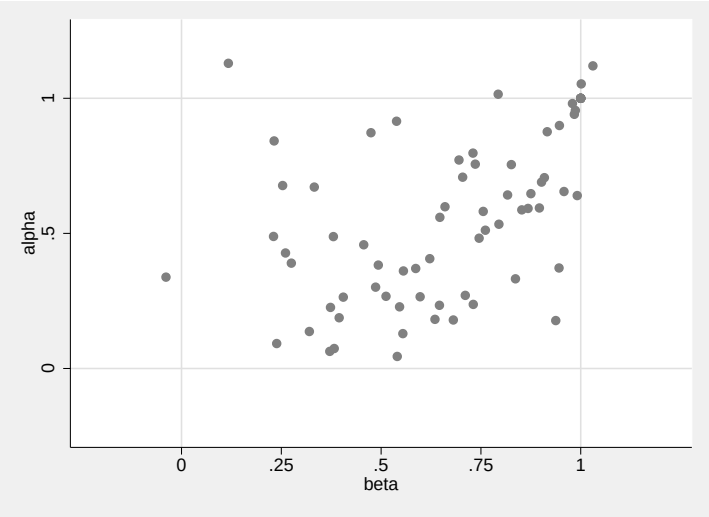
We estimate Equation 1.5 separately for each individual subject. Note that all quantities except the parameters in the equation are directly observed in our experimental data. $\ln\big(t^\star\big)$ is the logarithm of the maximum transfer accepted in Part 2, $\ln\big[v_j(k)\big]$ equals the logarithm of the receiver's willingness to pay $\text{WTP}_{\text{receiver}}$ elicited in Part 1, and $\ln[v_i(k)]$

[8] This corresponds to a proposed value of $\beta = 1/2$

is the logarithm of the sender's willingness to pay $\text{WTP}_{\text{sender}}$ elicited in Part 1.[9] We thus estimate the following linear equation.

$$\ln\left(t^{\star}\right) = \gamma_0 + \gamma_1 \ln(\text{WTP}_{\text{receiver}}) + \left(1-\gamma_1\right)\ln\left(\text{WTP}_{\text{sender}}\right) + \epsilon \tag{1.6}$$

The estimates for the general level of altruism are given by $\hat{\alpha} = \exp\left(\hat{\gamma}_0\right)$ and those for empathy by $\hat{\beta} = \hat{\gamma}_1$.




*Note:* The figure shows estimated parameter values for $\alpha$ and $\beta$. Only those subjects entered the analysis who made varying choices within Part 1 and Part 2. In addition, six subjects were excluded due to implausible parameter estimates and three further subjects were excluded due to large standard deviations of the parameter estimates, leaving 71 observations.

Figure 1.3: Estimates for individual parameters

Figure 1.3 shows parameter estimates for $\beta$ on the horizontal axis and $\alpha$ on the vertical axis. Variation in senders' WTPs and transfers in principle allows us to identify parameters for 80 subjects, of whom we get reasonable estimates for 71. Among the latter, the vast majority of subjects is assigned estimates that lie inside the ranges of expected values from zero to one. We see large heterogeneity in parameter estimates, and the variation in estimates for $\beta$ indicates that the effects that we find in the analysis on the aggregate level are not only driven by a small number of subjects. Moreover, the figure shows that, for any given level of general altruism, there exists marked heterogeneity in the empathy parameter. The two thus appear to be distinct characteristics of the individuals.

[9] To avoid missing values at zero, we add 0.1 to all WTPs and transfers.

## Welfare

We now turn to study the welfare implications of the decisions that were observed in the experiment. To test Hypothesis 1.2, we regress transfers on two different measures of dissimilarity between sender and receiver. We define partial dissimilarity as the absolute difference between sender $i$'s and receiver $j$'s WTP regarding the relevant item $k$, divided by its maximum of 20.

$$\text{Partial dissimilarity}_{ijk} = \frac{\left|\text{WTP}_{ik} - \text{WTP}_{jk}\right|}{20}$$

Total dissimilarity is the Euclidean distance between the full vectors of sender $i$'s and receiver $j$'s WTPs for all items $k$, again divided by its potential maximum value.

$$\text{Total dissimilarity}_{ij} = \frac{\sqrt{\sum_{k=1}^{8}\left(\text{WTP}_{ik} - \text{WTP}_{jk}\right)^2}}{20\sqrt{8}}$$

| | *Dependent variable:* Transfer | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Partial dissimilarity | -4.156$^{****}$ | | -4.046$^{****}$ | | -3.724$^{****}$ |
| | (0.718) | | (0.728) | | (0.718) |
| Total dissimilarity | | -4.244$^{****}$ | | -4.393$^{****}$ | -0.652 |
| | | (0.970) | | (0.973) | (0.789) |
| Receiver's WTP | 0.378$^{****}$ | | 0.305$^{****}$ | | 0.301$^{****}$ |
| | (0.0341) | | (0.0428) | | (0.0418) |
| Sender's WTP | 0.285$^{****}$ | | 0.145$^{***}$ | | 0.143$^{***}$ |
| | (0.0357) | | (0.0442) | | (0.0447) |
| Receiver's average WTP | | 0.393$^{****}$ | | | |
| | | (0.0419) | | | |
| Sender's average WTP | | 0.310$^{****}$ | | | |
| | | (0.0470) | | | |
| Sender fixed effects | No | No | Yes | Yes | Yes |
| Receiver fixed effects | No | No | Yes | Yes | Yes |
| Item fixed effects | No | No | Yes | Yes | Yes |
| Observations | 1168 | 1168 | 1168 | 1168 | 1168 |
| Clusters | 146 | 146 | 146 | 146 | 146 |
| (Within-)$R^2$ | 0.429 | 0.305 | 0.208 | 0.0773 | 0.208 |

*Note:* OLS regression; standard errors are clustered for senders; $^{*}$ $p<0.1$; $^{**}$ $p<0.05$; $^{***}$ $p<0.01$; $^{****}$ $p<0.001$; robust standard errors in parentheses.

Table 1.2: Similarity and transfers

Table 1.2 shows the results. Column 1 and Column 3 use partial dissimilarity, while

Column 2 and Column 4 use total dissimilarity. Columns 1 and 2 present the baseline results without any fixed effects. Column 1 uses the receiver's and the sender's WTP to control for level effects. The effect of partial dissimilarity is thus conditional on both parties' own valuations, and it shows that dissimilarity decreases the size of transfers. In Column 2, level effects are correspondingly controlled for by using the receiver's and the sender's average WTP, because total similarity also refers to all items. Total dissimilarity enters negatively and with a similar effect size as partial dissimilarity in Column 1. Columns 3 and 4 replicate the previous two with fixed effects for senders, receivers, and items, making controls for average WTPs redundant. The effects of dissimilarity remain almost unchanged. In Column 5, regressors from the previous two columns are combined. Total dissimilarity has no significant effect beyond the effect through partial dissimilarity, which is line with Hypothesis 1.2. Interestingly, this means that senders descriptively discriminate against receivers whose preferences are different but only because of imperfect empathy and not because they generally dislike them.

To test Hypothesis 1.3, we first need to derive the welfare consequences of decisions over any proposed transfer level $t$. The net welfare gain from no transfer is—by definition—zero. If the proposed transfer (i.e. the row on the decision screen) is accepted, the net welfare gain can be calculated according to Equation 1.3. If a transfer of $t$ is made, welfare is given by $20 - t + 20$. If the transfer is not made, welfare is given by $20 + 20 - v_j(k)$, where the latter valuation corresponds to the WTP of the receiver. For any proposed transfer (a row in the MPL), the welfare impact can thus be calculated as follows.

$$\text{Net welfare gain} \equiv \begin{cases} \text{WTP}_{\text{receiver}} - t & \text{if transfer of } t \text{ is made} \\ 0 & \text{if transfer of } t \text{ is not made} \end{cases}$$

By the design of the MPLs used in the experiment, the probability of a transfer being made for a given item and receiver is the maximum size of the transfer, $t^\star$, divided by the number of rows, which is 20. If a transfer is made, the receiver experiences a welfare gain equivalent to his corresponding WTP. The sender loses the transfer amount of the respective row. We can thus calculate the expected net welfare gain of any transfer decision made by senders as follows.

$$\mathbb{E}\big[\text{net welfare gain}\big] = \underbrace{\frac{t^\star}{20}}_{\mathbb{P}\,(\text{transfer made})} \Big(\text{WTP}_{\text{receiver}} - \underbrace{\frac{t^\star + 1}{2}}_{\mathbb{E}\,[\text{transfer}\,|\,\text{transfer made}]}\Big)$$

Table 1.3 shows the results of regressing the expected net welfare gain on partial dissimilarity or total dissimilarity. Columns correspond to the ones in Table 1.2. Columns 1 and 2 present the baseline results without any fixed effects. Column 1 uses the receiver's and

| | *Dependent variable:* $\mathbb{E}$[net welfare gain] | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Partial dissimilarity | -2.648$^{****}$ | | -2.460$^{****}$ | | -2.398$^{****}$ |
| | (0.373) | | (0.370) | | (0.378) |
| Total dissimilarity | | -2.490$^{****}$ | | -2.484$^{****}$ | -0.126 |
| | | (0.481) | | (0.504) | (0.352) |
| Receiver's WTP | 0.270$^{****}$ | | 0.260$^{****}$ | | 0.260$^{****}$ |
| | (0.0183) | | (0.0224) | | (0.0224) |
| Sender's WTP | 0.0878$^{****}$ | | 0.0805$^{****}$ | | 0.0802$^{****}$ |
| | (0.0141) | | (0.0183) | | (0.0186) |
| Receiver's average WTP | | 0.267$^{****}$ | | | |
| | | (0.0214) | | | |
| Sender's average WTP | | 0.0915$^{****}$ | | | |
| | | (0.0204) | | | |
| Sender fixed effects | No | No | Yes | Yes | Yes |
| Receiver fixed effects | No | No | Yes | Yes | Yes |
| Item fixed effects | No | No | Yes | Yes | Yes |
| Observations | 1168 | 1168 | 1168 | 1168 | 1168 |
| Clusters | 146 | 146 | 146 | 146 | 146 |
| (Within-)$R^2$ | 0.536 | 0.331 | 0.397 | 0.0665 | 0.397 |

*Note:* OLS regression; standard errors are clustered for senders; $^{*}$ $p < 0.1$; $^{**}$ $p < 0.05$; $^{***}$ $p < 0.01$; $^{****}$ $p < 0.001$; robust standard errors in parentheses.

Table 1.3: Similarity and welfare

the sender's WTPs to control for level effects. We find that partial dissimilarity decreases the expected net welfare gain. In Column 2, level effects are controlled for by using the receiver's and the sender's average WTP. Total dissimilarity also enters negatively, with a magnitude that is comparable to that of partial dissimilarity. Columns 3 and 4 replicate the previous two columns with fixed effects for senders, receivers, and items. The estimated effects remain stable. Column 5 combines the regressors from Columns 3 and 4, showing that the effect of total dissimilarity is entirely driven by the effect of partial dissimilarity regarding the relevant item. Thus, Hypothesis 1.3 is confirmed. Moreover, we again find no evidence for taste-based discrimination against receivers with different preferences.

## Libertarians vs. Paternalists

Since we have shown that people partially rely on their own preferences in choosing the level to which they provide others with help, it is natural to ask whether and how this might be related to paternalistic behavior: if people are not willing to support choices that seem strange to them, they might also want to change them. Nevertheless, imperfect

empathy and paternalism are different concepts. First, imperfect empathy pertains to a certain kind of preferences, whereas paternalism is a certain kind of behavior. Second, the ranges of relevant applications of both phenomena might overlap (see, e.g., Jacobsson et al., 2007) but are not identical: imperfect empathy is relevant in many situations where restricting others' freedom is not even an option; and paternalism occurs in many contexts where empathy is not relevant but is often driven by, e.g., asymmetric information. Third, it is not clear whether people who make helping behavior depend on their own valuations regard the latter as normatively warranted or would rather—if they were aware of it—object to such behavior and therefore also not want to restrict others' freedom.

To study the relationship between imperfect empathy and paternalism empirically, we use subjects' choices from Part 4 to classify them as paternalists or libertarians. A subject is only classified as a libertarian if she abstained from altering any other subjects' decisions. All subjects that altered any decision are classified as paternalists. According to this definition, we end up with 74 libertarian subjects and 72 paternalists.

| | *Dependent variable:* Transfer | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Libertarians | | | | Paternalists | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Receiver's WTP | 0.412$^{****}$ | 0.398$^{****}$ | 0.210$^{****}$ | 0.219$^{****}$ | 0.207$^{****}$ | 0.179$^{****}$ | 0.134$^{****}$ | 0.105$^{**}$ |
| | (0.0482) | (0.0528) | (0.0554) | (0.0583) | (0.0356) | (0.0386) | (0.0315) | (0.0428) |
| Sender's WTP | 0.238$^{****}$ | 0.156$^{***}$ | 0.0283 | -0.0434 | 0.251$^{****}$ | 0.209$^{***}$ | 0.127$^{*}$ | 0.0862 |
| | (0.0451) | (0.0574) | (0.0470) | (0.0792) | (0.0550) | (0.0657) | (0.0669) | (0.0808) |
| $\sqrt{\text{Sender's} \times \text{receiver's WTP}}$ | | | 0.442$^{****}$ | 0.387$^{****}$ | | | 0.236$^{***}$ | 0.224$^{**}$ |
| | | | (0.0894) | (0.0979) | | | (0.0888) | (0.0964) |
| Sender fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Receiver fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 592 | 569 | 592 | 569 | 576 | 558 | 576 | 558 |
| Clusters | 74 | 74 | 74 | 74 | 72 | 72 | 72 | 72 |
| (Within-) $R^2$ | 0.451 | 0.269 | 0.516 | 0.357 | 0.267 | 0.142 | 0.291 | 0.177 |

*Note:* OLS regression; standard errors are clustered for senders; $^{*}$ $p<0.1$; $^{**}$ $p<0.05$; $^{***}$ $p<0.01$; $^{****}$ $p<0.001$; robust standard errors in parentheses. Columns with fixed effects include fewer observations than others because some receivers were only matched to a single paternalist or libertarian sender, respectively.

Table 1.4: Libertarians vs. paternalists

Table 1.4 replicates our main results on transfer decisions from the first four columns of Table 1.1 separately for libertarians in Columns 1–4 and paternalists in Columns 5–8. Comparing Columns 1 and 2 with Columns 5 and 6, respectively, we see that the role of senders' WTPs is slightly weaker among libertarians as compared to paternalists, although the differences in coefficients are not statistically significant. This is plausible since paternalists are subjects that are less willing to accept others' choices and it is therefore not surprising that they react less to receivers' WTPs. More importantly, the effect

of senders' WTPs enters with considerable magnitude and high statistical significance within both subsamples. Proceeding towards the comparison of Columns 3 and 4 with Columns 7 and 8, it turns out that the effect of the interactions between senders' and receivers' WTPs in fact tends to be stronger among libertarians than among paternalists. The latter finding also alleviates concerns about measurement error driving our results, i.e., about senders trying to "correct" receivers' choices. While the interpretation of WTPs as noisy signals about true valuations would not be able to explain the asymmetry of our results in the first place—i.e., the complementarity of valuations—it would also be incompatible with senders not intervening in other subjects' own decisions: senders care about receivers—they make transfers—and if they thought others were making mistakes, they should save them from doing so. In contrast, in our experiment, even people who put faith in others' personal judgments and who do not show any signs of paternalism exhibit imperfect empathy.

## 1.5 Conclusion

In this chapter, we show that people behave imperfectly empathic when acting prosocially. They assess consequences arising to others based on a combination of their own and the other persons' preferences. In particular, own and others' valuations act as complements in bringing about helping behavior. We show that this property of imperfect empathy leads to the effect that dissimilar preferences lower the size of transfers as well as overall welfare.

We hereby confirm the hypothesis that empathy plays a role in generating prosocial behavior and furthermore show that a lack of it can lead to lower—according to a basic libertarian welfare criterion—and poorly aimed helping behavior. Our results also inform models featuring altruism in conjunction with heterogeneous preferences. However, the mechanism of imperfect empathy is not only relevant for individual behaviors such as charitable giving or volunteering. It also allows for an alternative perspective on the phenomenon of in-group bias. We observe that transfers are lower if other people have overall different preferences. Within our experiment, however, this effect is entirely driven by imperfect empathy and not by a dislike against subjects who are different. Imperfect empathy might also have implications on the aggregate level for the working of welfare states. If people cannot relate to the consumption choices made by recipients of welfare benefits, this could decrease the willingness to finance such redistributive policies. An implication for policy might be that exposure to individuals with different sets of preferences, e.g. due to cultural or religious backgrounds, could be central for the political sustainability of welfare states in increasingly diverse societies.

# References


Ambuehl, Sandro. 2017. *An Offer You Can't Refuse? Incentives Change How We Inform Ourselves and What We Believe*. CESifo Working Paper 6296. Munich: Center for Economic Studies & Ifo Institute.

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

Andreoni, James, A. Abigail Payne, Justin Smith, and David Karp. 2016. "Diversity and Donations: The Effect of Religious and Ethnic Diversity on Charitable Giving." *Journal of Economic Behavior & Organization* 128:47–58.

Batson, C. Daniel. 1987. "Prosocial Motivation: Is it ever Truly Altruistic?" *Advances in Experimental Social Psychology* 20:65–122.

Batson, C. Daniel, and Tecia Moran. 1999. "Empathy-Induced Altruism in a Prisoner's Dilemma." *European Journal of Social Psychology* 29 (7): 909–924.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82 (6): 1063–1093.

— . 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14 (3): 817–826.

Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review* 71:117–120.

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

Coke, Jay S., C. Daniel Batson, and Katherine McDavis. 1978. "Empathic Mediation of Helping: A Two-Stage Model." *Journal of Personality and Social Psychology* 36 (7): 752–766.

Dahlberg, Matz, Karin Edmark, and Heléne Lundqvist. 2012. "Ethnic Diversity and Preferences for Redistribution." *Journal of Political Economy* 120 (1): 41–76.

Davis, M. H. 1980. "A Multidimensional Approach to Individual Differences in Empathy." *JSAS Catalog of Selected Documents in Psychology* 10:85.

Doepke, Matthias, and Fabrizio Zilibotti. 2017. "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission." *Econometrica* 85 (5): 1331–1371.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2): 268–298.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2): 293–315.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

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

Fowler, James H., and Cindy D. Kam. 2007. "Beyond the Self: Social Identity, Altruism, and Political Participation." *Journal of Politics* 69 (3): 813–827.

Gerlitz, Jean-Yves, and Jürgen Schupp. 2005. "Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP." *Research Notes 4*: 1–44.

Jacobsson, Fredric, Magnus Johannesson, and Lars Borgquist. 2007. "Is Altruism Paternalistic?" *Economic Journal* 117 (520): 761–781.

Klimecki, Olga M., Sarah V. Mayer, Aiste Jusyte, Jonathan Scheeff, and Michael Schönenberg. 2016. "Empathy Promotes Altruistic Behavior in Economic Interactions." *Scientific Reports* 6:31961.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–1302.

Ross, Lee. 1977. "The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process." *Advances in Experimental Social Psychology* 10:173–220.

Smith, Adam. 1859. *The Theory of Moral Sentiments*. Glasgow.

Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1 (2): 149–178.

## 1.A Proof of Hypothesis 1.2

The partial derivative of expected transfers given by Equation 1.4 with respect the distance between subjects' valuations is negative whenever $\beta$ lies in the open interval from zero to one and the distance between individual valuations is larger than one. Valuations are denoted by $v_i, v_j > 0$ and $\bar{v} = v_i + v_j$ denotes the total of both valuations.

*Proof.*

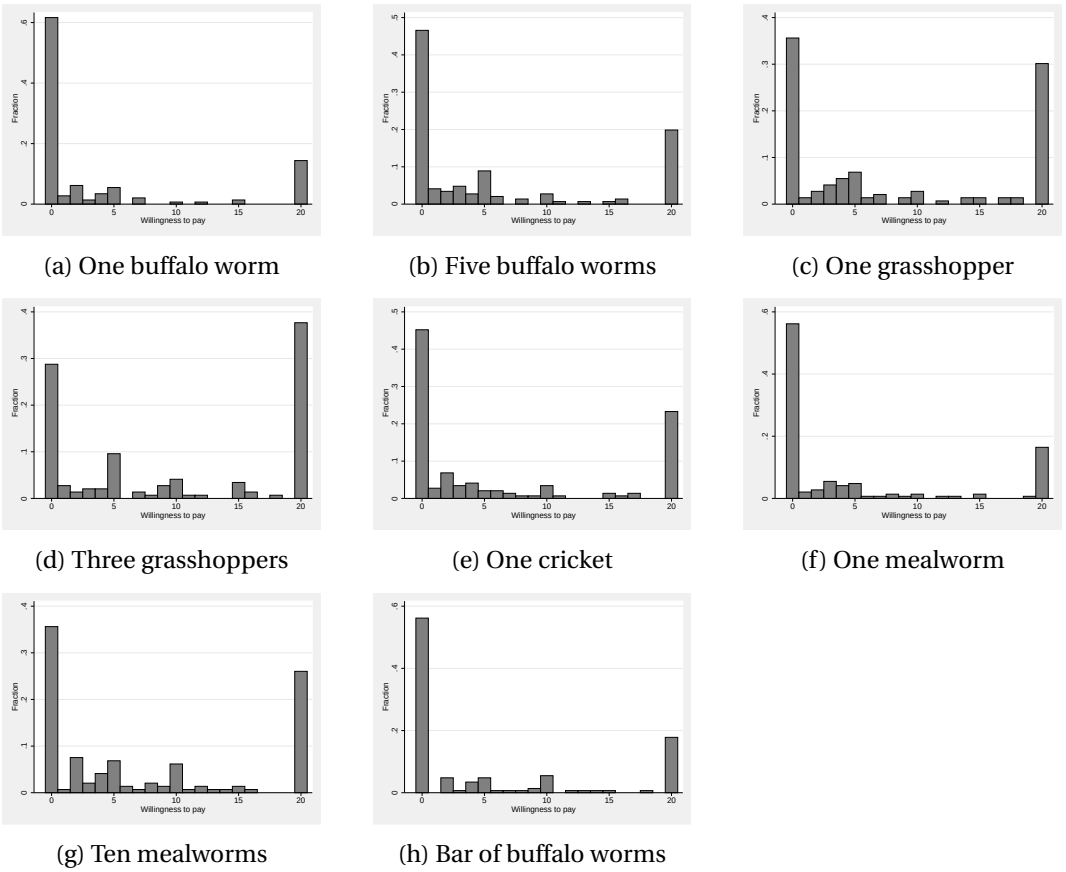
$$\begin{aligned}
\frac{\partial \mathbb{E}[t^\star]}{\partial |v_i - v_j|} &= \frac{\alpha}{4}\left[\beta\left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)^{1-\beta} - (1-\beta)\left(\frac{\bar{v} + |v_i - v_j|}{\bar{v} - |v_i - v_j|}\right)^{\beta}\right. \\
&\quad \left. - \beta\left(\frac{\bar{v} + |v_i - v_j|}{\bar{v} - |v_i - v_j|}\right)^{1-\beta} + (1-\beta)\left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)^{\beta}\right] \\
&= \frac{\alpha}{4}\left\{\beta\left[\underbrace{\left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)}_{\in(0,1]}^{1-\beta} - \left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)^{-(1-\beta)}\right]\right. \\
&\quad \left. + (1-\beta)\left[\left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)^{\beta} - \left(\frac{\bar{v} - |v_i - v_j|}{\bar{v} + |v_i - v_j|}\right)^{-\beta}\right]\right\} \\
&\begin{cases} < 0 & \text{if } \beta \in (0,1) \wedge |v_i - v_j| > 0 \\ = 0 & \text{if } \beta \in \{0,1\} \vee |v_i - v_j| = 0 \end{cases}
\end{aligned}$$

□

## 1.B Descriptive Statistics

| Variable | Mean | Minimum | Maximum | Standard deviation | Observations |
|---|---|---|---|---|---|
| Female | 0.500 | 0 | 1 | 0.502 | 146 |
| Age | 25.630 | 18 | 69 | 7.741 | 146 |
| Partial distance | 0.416 | 0 | 1 | 0.386 | 1168 |
| Total distance | 0.493 | 0 | 1 | 0.281 | 1168 |

Table 1.5: Summary statistics



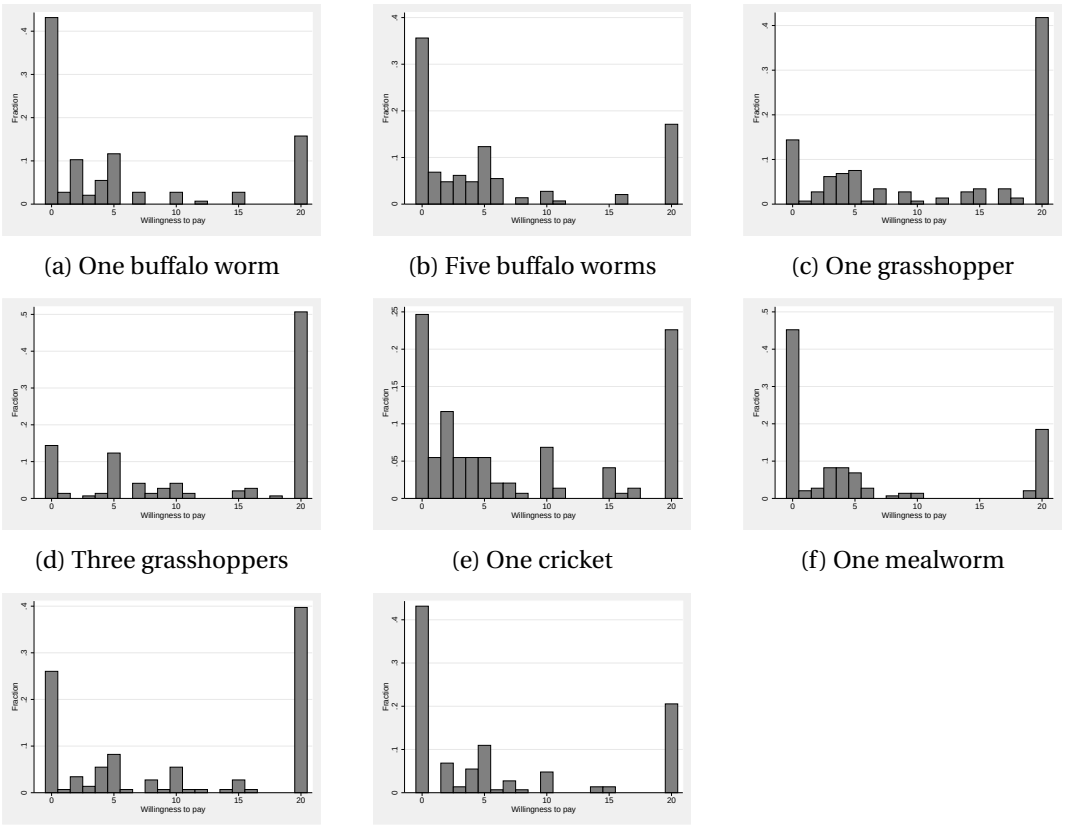

(a) One buffalo worm (b) Five buffalo worms (c) One grasshopper

(d) Three grasshoppers (e) One cricket (f) One mealworm

(g) Ten mealworms (h) Bar of buffalo worms

*Note:* The figure shows the distribution of the WTPs for the eight food items of all subjects who acted as senders in Part 2. Shown are the decisions they made for themselves in Part 1.

Figure 1.4: Senders' willingness to pay (Part 1)



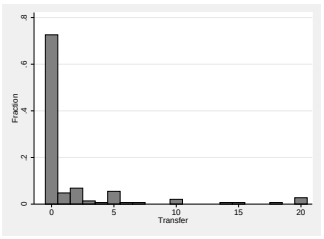

(a) One buffalo worm (b) Five buffalo worms (c) One grasshopper

(d) Three grasshoppers (e) One cricket (f) One mealworm

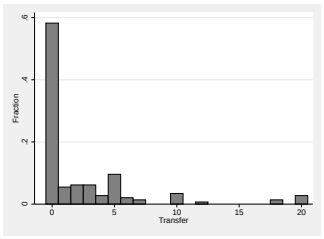
(g) Ten mealworms (h) Bar of buffalo worms

*Note:* The figure shows the distribution of the WTPs for the eight food items of all subjects who acted as receivers in Part 2. Shown are the decisions they made for themselves in Part 1.

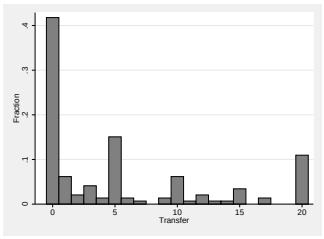
Figure 1.5: Receivers' willingness to pay (sampled from Part 1)



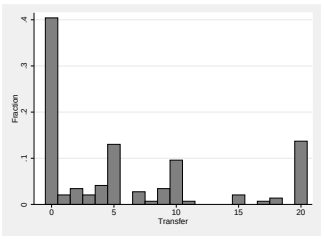

(a) One buffalo worm



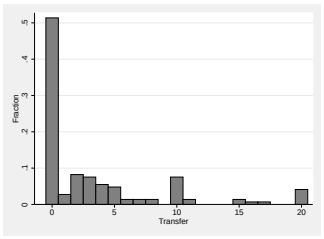

(b) Five buffalo worms



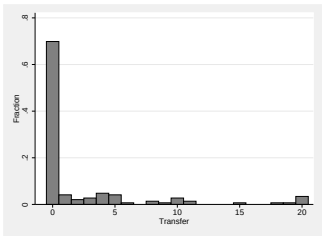

(c) One grasshopper



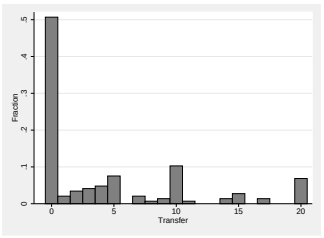

(d) Three grasshoppers



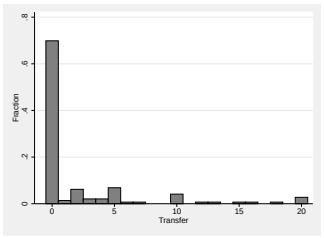

(e) One cricket




(f) One mealworm




(g) Ten mealworms




(h) Bar of buffalo worms

Figure 1.6: Transfers (Part 2)

## 1.C Instructions

**Verbal Instructions**

"Welcome to today's experiment! Before we start, we would like to provide you with some information.

In this experiment, you will, under specific circumstances, eat dried and non-living insects (or food containing dried insects) that were farmed in the European Union for the consumption of humans. We will give you information on the specific insects and we will also show them to you. We will now come to your cubicles and show you a tray with the insects and the food containing insects. We will furthermore distribute written information on the nutritional values of the insects and forms of consent. Please read everything carefully and sign the forms of consent; we will afterwards collect the written information and one of the forms. You can keep the other form of consent as a copy. As soon as you have read and signed everything, please hold up your hand to let us know you are ready.

The insects that you are seeing now are crickets, grasshoppers, meal worms, and buffalo worms. You are furthermore seeing a cereal bar containing buffalo worms. You can also touch the insects and the bar carefully. All of these insects were farmed in the European Union for human consumption. They have been certified as food and are completely innocuous for your health; on the contrary, insects are typically very healthy.

Please close your curtains now; we will start the experiment."

**2. Welcome and Introduction**

Welcome!

For participating in this experiment, you and all other present participants will receive a payment of €5/€7. All further payments, which will depend on your decisions during this experiment, will be added.

Today's study is made up of four parts. You will make decisions that can influence your payment or the payment of another participant in all four parts. At the same time, other participants will make decisions that can in turn influence your payment.

At the end of the experiment, with a probability of 50%, one of your own decisions will be implemented. All single decisions that you make during the course of the experiment will be paid out with equal probability.

With a probability of 50%, a decision another person has made for you will be implemented. All single decisions that a participant took for somebody else will be implemented with equal probability.

Please raise your hand if you have questions at any time. One of the experimenters will then come to you.

Please click "Continue" to start the experiment.

**3. Instructions Part 1**

For Part 1 of this study, you receive a payment of €20. In the following, we would like to know which amount of these €20 you would be willing to pay in order to not eat a specific insect. For every insect or food item containting insects that were shown to you at the beginning of the study you will in the following see a list of decisions of the following form:

Option **Payment**: you pay €x and do not have to eat the insect – Option **Insect**: you eat the insect

For each of these decisions, you have to decide between the option **Payment** and the option **Insect**.

In case one of the decisions is implemented at the end of the experiment and you chose option **Payment**, you receive €20 minus the amount that was indicated in the description and you do not have to eat. In case you chose option **Insect** for the chosen decision, you will have to eat the insect before payment and receive €20 without deduction.

**Important: in case you choose option Insect and–contrary to your decision–refuse to eat the insect, you will receive a penalty of €20 and receive €0.**

If you understood everything, please click "Continue" in order to continue with the experiment.

**4. Decision Screen Part 1 (see Screenshot in Figure 1.8)**

**5. Instructions Part 2**

Thank you for finishing Part 1. You are now starting Part 2.

For Part 2 of this study, you receive a payment of €20. In the following we would like to know which amount of these €20 you would be willing to pay in order for **another participant of this experiment** not having to eat a specific insect. For every insect or

food containing insects that were shown to you at the beginning of the study you will in the following see a list of decisions of the following form:

Option **Payment**: you pay €x and the other person does not have to eat the insect – Option **Insect**: the other person has to eat the insect

You are making the decision for **eight other participants that are all taking part in this study at this moment**. At the same time, other participants who are also taking part in this study at this moment are making the same decision for you. Before you make your decision for the other participant, you will receive information on how much money this participant was maximally willing to pay for not having to eat the insect in Part 1.

For every decision, you have to decide between option **Payment** and option **Insect**.

In case that at the end of the experiment one of your decisions for somebody else is being implemented, the following will happen: if you chose option **Payment** for the chosen decision, you receive €20 minus the amount that was indicated in the description and the other participant does not have to eat. If you chose option **Insect** for the decision, the other participant has to eat the insect before payment and you receive €20 without deduction.

**Important: in case you choose option Insect and–contrary to your decision–the other person refuses to eat the insect, the other person will receive a penalty of €20 and receive €0. You will receive €20.**

In case that at the end of the experiment one of the decisions another participant has made for you is being implemented, the following will happen: if the other participant chose option **Payment**, you will receive €20 and you will not have to eat the insect. If the other participant chose option **Insect**, you will have to eat the insect before payment and you will receive €20.

**Important: in case the other participant chose option Insect and–contrary to his or her indication–you refuse to eat the insect, you will receive a penalty of €20 and receive €0.**

If you understood everything, please click "Continue" in order to continue with the experiment.

**6. Decision Screen Part 2 (see Screenshot in Figure 1.9)**

**7. Instructions Part 3**

Thank you for finishing Part 2. You are now starting Part 3.

For Part 3 of this study, you receive a payment of €20. In the following, we would like to know which amount of these €20 you would be willing to pay in order for **another participant of this experiment** to receive a higher payoff. In the following, you will see a list of decisions of the following form:

Option **Payment**: you pay €x and the other person receives €20 – Option **No Payment**: the other person receives €0/€5/€10/€15

For every decision, you have to decide between option **Payment** and option **No Payment**.

In case that at the end of the experiment one of your decisions for somebody else is being implemented, the following will happen: if you chose option **Payment** for the chosen decision, you receive €20 minus the amount that was indicated in the description and the other participant receives €20. If you chose option **No Payment** for the decision, the other participant receives the respective lower payment and you receive €20 without deduction.

In case that at the end of the experiment one of the decisions another participant has made for you is being implemented, the following will happen: if the other participant chose option **Payment**, you will receive €20. If the other participant chose option **No Payment**, you will receive the respective lower amount.

If you understood everything, please click "Continue" in order to continue with the experiment.

**8. Decision Screen Part 3 (see Screenshot Figure 1.10)**

**10. Instructions Part 4**

Thank you for finishing Part 3. You are now starting Part 4.

In the following we would like to know if you would change the decisions from **Part 1 of another participant**.

As a reminder: in Part 1, every participant decided how much he or she would be willing to pay maximally in order not to eat a specific insect. Every participant saw one list of decisions per insect of the following form and had to decide between option **Payment** and option **Insect**.

Option **Payment**: you pay €x and do not have to eat the insect – Option **Insect**: you eat the insect

You will now see the lists of **eight participants that are all participating in this study at**

**this moment** and their decisions. You can change the decisions of the participants as you want.

In case that at the end of the experiment one of your decisions for somebody else is being selected, this decision will be implemented for the other person. If option **Payment** was chosen, the other participant receives €20 minus the amount that was indicated in the description and does not have to eat. If option **Insect** was chosen, the other participant has to eat the insect before payment and receives €20 without deduction.

**Important: in case option Insect was chosen and–contrary to the decision–the other person refuses to eat the insect, the other person will receive a penalty of €20 and will receive €0.**

In case that at the end of the experiment one of the decisions another participant has made for you is being implemented, the following will happen: if option **Payment** was chosen, you receive €20 minus the amount that was indicated in the description and you do not have to eat the insect. If option **Insect** was chosen, you have to eat the insect before payment and you receive €20.

**Important: in case option Insect was chosen and–contrary to the decision–you refuse to eat the insect, you will receive a penalty of €20 and receive €0.**

If you understood everything, please click "Continue" in order to continue with the experiment.

**9. Decision Screen Part 4 (see Screenshot in Figure 1.11)**

**11. Questionnaire**

Thank you for finishing Part 1 to Part 4.

In the following we are asking you to answer some questions.

After you have answered all questions, the experiment is over.

How old are you?

What is your gender?

How high is your monthly income (after taxes and before all expenses)?

In the following, we are interested in how much you are willing to take on risks. Please state your evaluation on a scale from 0 to 10. 0 means "not at all willing to take on risks"

and 10 means “very willing to take on risks”. You can grade your evaluation with the values in between.

0 1 2 3 4 5 6 7 8 9 10

We are now asking you for your willingness to behave a certain way. Please state your evaluation on a scale from 0 to 10. 0 means “not at all willing to do this” and 10 means “very willing to do this”. You can grade your evaluation with the values in between.

How much are you willing to forego something that carries utility for you in order to benefit from it in the future?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to punish someone who has treated you unfairly even though this has negative consequences for you?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to punish someone who has treated somebody else unfairly even though this has negative consequences for you?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to donate to a good cause without expecting something in return?

0 1 2 3 4 5 6 7 8 9 10

Please think about how you would act in the following situation. You are in an unknown area und notice that you got lost. You are asking a stranger for the way. The stranger offers to accompany you to your destination.

Helping you costs the stranger approximately €20. However, the stranger says that he does not want money from you. You have six gifts with you. The cheapest gift costs €5, the most expensive gift costs €30. Would you offer the stranger one of the gifts as a thank you?

Yes/No

Which gift would you give to the stranger?

The gift worth €5

The gift worth €10
The gift worth €15
The gift worth €20
The gift worth €25
The gift worth €30
I don't know

Imagine the following situation: today, you received an unexpected €1.000.

How much of the money would you donate to a good cause? Donation:

Here are different characteristics a person can have. Probably, some of the characteristics will apply to you personally, whereas others do not. For some, you may be undecided.

Please use the following scale to answer: value 1 means "does not apply at all" and value 7 means "applies very much". With the values between 1 and 7, you can grade your evaluation.

I see myself as someone, who...

- is a reliable worker ..... 1 2 3 4 5 6 7
- is talkative ..... 1 2 3 4 5 6 7
- is sometimes rude to others ..... 1 2 3 4 5 6 7
- is original, comes up with new ideas ..... 1 2 3 4 5 6 7
- worries a lot ..... 1 2 3 4 5 6 7
- has a forgiving nature ..... 1 2 3 4 5 6 7
- tends to be lazy ..... 1 2 3 4 5 6 7
- is outgoing, sociable ..... 1 2 3 4 5 6 7
- values artistic, aesthetic experiences ..... 1 2 3 4 5 6 7
- gets nervous easily ..... 1 2 3 4 5 6 7
- does things efficiently ..... 1 2 3 4 5 6 7
- is reserved ..... 1 2 3 4 5 6 7
- is considerate and kind to almost everyone ..... 1 2 3 4 5 6 7
- has an active imagination ..... 1 2 3 4 5 6 7
- is relaxed, handles stress well ..... 1 2 3 4 5 6 7

Please indicate for each of the following statements to which extent it applies to you personally. Please state your evaluation on a scale from 1 to 5. A 1 means "describes me very well" and a 5 means "does not describe me well". You can grade your evaluations with the values in between.

- I daydream and fantasize, with some regularity, about things that might happen to me.
- I often have tender, concerned feelings for people less fortunate than me.
- I sometimes find it difficult to see things from the "other guy's" point of view.
- Sometimes I don't feel very sorry for other people when they are having problems.
- I really get involved with the feelings of the characters in a novel.
- In emergency situations, I feel apprehensive and ill-at-ease.
- I am usually objective when I watch a movie or play, and I don't often get completely caught up in it.
- I try to look at everybody's side of disagreement before I make a decision.
- When I see someone being taken advantage of, I feel kind of protective towards them.
- I sometimes feel helpless when I am in the middle of a very emotional situation.
- I sometimes try to understand my friends better by imagining how things look from their perspective.
- Becoming extremely involved in a good book or movie is somewhat rare for me.
- When I see someone get hurt, I tend to remain calm.
- Other people's misfortunes do not usually disturb me a great deal.
- If I'm sure I'm right about something, I don't waste much time listening to other people's arguments.
- After seeing a play or movie, I have felt as though I were one of the characters.
- Being in a tense emotional situation scares me.
- When I see someone being treated unfairly, I sometimes don't feel very much pity for them.
- I am usually pretty effective in dealing with emergencies.
- I am often quite touched by things that I see happen.
- I believe that there are two sides to every question and try to look at them both.

- I would describe myself as a pretty soft-hearted person.
- When I watch a good movie, I can very easily put myself in the place of a leading character.
- I tend to lose control during emergencies.
- When I'm upset at someone, I usually try to "put myself in his shoes" for a while.
- When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me.
- When I see someone who badly needs help in an emergency, I go to pieces.
- Before criticizing somebody, I try to imagine how I would feel if I were in their place.

How did you decide how much you would be willing to pay for not having to eat an insect?

Based on which criteria did you decide how much money to pay for the other participants?

Do you have any further comments?

Thank you for your participation!

We will begin the payment shortly. Please wait on your seat until your cubicle number is called and then come to the adjoining room to receive your payment.

## 1.D Stimuli Pictures

(a) One buffalo worm

(b) Five buffalo worms

(c) One mealworm

(d) Ten mealworms

(e) One grasshopper

(f) Three grasshoppers

(g) One cricket

(h) Bar of buffalo worms

Figure 1.7: Stimuli pictures of insects

# 1.E Screenshots


Figure 1.8: Screenshot of Part 1


Figure 1.9: Screenshot of Part 2


Figure 1.10: Screenshot of Part 3


Figure 1.11: Screenshot of Part 4

## 1.F Robustness Regarding Income Levels

In Section 1.2, we have made the assumption that utility from money is linear, which we have used throughout this chapter. We believe that this assumption is innocent since we are concerned with monetary amounts in a range of €0 to €27. However, as a simple robustness exercise, we varied the fixed show-up fee that subjects received between sessions. In four sessions, subjects received €7 and in five sessions, they received €5. If the level of earnings during the experiment mattered for subjects' decision making, this should voice itself in results that differ between sessions depending on the size of the show-up fee.

| | *Dependent variable:* Transfer | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Show-up fee = €7 | | | | Show-up fee = €5 | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Receiver's WTP | 0.292$^{****}$ | 0.285$^{****}$ | 0.135$^{***}$ | 0.120$^{**}$ | 0.321$^{****}$ | 0.328$^{****}$ | 0.179$^{****}$ | 0.188$^{****}$ |
| | (0.0491) | (0.0605) | (0.0428) | (0.0515) | (0.0403) | (0.0450) | (0.0412) | (0.0488) |
| Sender's WTP | 0.259$^{****}$ | 0.202$^{**}$ | 0.0808 | 0.0121 | 0.250$^{****}$ | 0.156$^{***}$ | 0.0433 | -0.0407 |
| | (0.0590) | (0.0878) | (0.0645) | (0.109) | (0.0386) | (0.0506) | (0.0483) | (0.0604) |
| $\sqrt{\text{Sender's} \times \text{receiver's WTP}}$ | | | 0.363$^{****}$ | 0.359$^{***}$ | | | 0.415$^{****}$ | 0.384$^{****}$ |
| | | | (0.0933) | (0.113) | | | (0.0859) | (0.0824) |
| Sender fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Receiver fixed effects | No | Yes | No | Yes | No | Yes | No | Yes |
| Observations | 528 | 528 | 528 | 528 | 640 | 640 | 640 | 640 |
| Clusters | 66 | 66 | 66 | 66 | 80 | 80 | 80 | 80 |

*Note:* OLS regression, standard errors are clustered for senders; $^{*}$ $p < 0.1$; $^{**}$ $p < 0.05$; $^{***}$ $p < 0.01$; $^{****}$ $p < 0.001$; robust standard errors in parentheses.

Table 1.6: High show-up fee vs. low show-up fee

Table 1.6 shows the results corresponding to the ones in Table 1.1 split according to the size of the show-up fee. Qualitative results are robust within both subsamples; all differences in coefficients are insignificant. Differences in the income level during the experiment therefore do not seem important for our results.

# Chapter 2

# Preferences Over the Interpersonal Correlation Between Uncertain Outcomes[1]

## 2.1 Introduction

Uncertainty is a pervasive facet of everyday life and very prevalent in economic decision contexts. Therefore, risk preferences play a part in many economic models and have been subject to extensive empirical investigation.

However, most models and risk measures have in common that they consider preferences over uncertain individual states in isolation, disregarding any potential uncertain states of others. This is surprising, given that social preferences are well established in the economic literature. In fact, situations in which people make decisions under uncertainty are often connected to decisions about the correlation between individuals' own and others' risky outcomes (e.g. financial investment decisions). Nevertheless, research on preferences over the interpersonal correlation between uncertain outcomes is almost non-existent, and—to our knowledge—even suggestive empirical evidence in which direction preferences tend to go is scarce.

We seek to find out if people have a preference for positive or for negative interpersonal correlation between uncertain outcomes. Furthermore, we investigate if women have stronger preferences for positively correlated uncertain outcomes than men. In order to do so, we run a simple laboratory experiment to elicit if people prefer risky payoff to be positively or negatively correlated with the risky payoff of another person. In the experiment, a subject decides between two different lotteries that both entail equal

[1] This chapter is joint work with Anke Becker and Frederik Schwerter.

chances of winning and losing: a 50%–50%-lottery in which she and another subject win at the same time or lose at the same time (positively correlated risky payoff) and a 50%–50%-lottery in which she wins if the other subject loses and vice versa (negatively correlated risky payoff).

We find that around 69% of subjects in our laboratory experiment choose the positively correlated uncertain payoffs, whereas around 31% choose the negatively correlated uncertain payoffs. Furthermore, women choose the positive correlation with a significantly higher probability than men.

There are several real world situations in which preferences on the interpersonal correlation between uncertain outcomes might play a role. One example concerns financial investment decisions. If people have a preference for positively correlated uncertain outcomes, they might invest too much in things in which others have invested already, a phenomenon that is often observed and which is discussed in the literature about herding (Devenow and Welch, 1996; Sias, 2004).[2] A particular interesting application might be the purchasing of private homes, since houses not only constitute the main financial asset of many households but also determine their relevant peer group: neighbors. When a house gains or loses in value, the same is usually also the case for surrounding houses, and thus risky financial outcomes amongst neighbors are positively correlated. If people value this fact, the latter might contribute to explaining the high home ownership rates in many countries (Oswald, 1996; Chambers et al., 2009).

Another situation in which preferences on the interpersonal correlation between uncertain outcomes might play a role is to what extent people choose jobs that require high or low degrees of specialization. Acquiring specialized human capital can lead to higher wages and more job-related success (Tam, 1997). However, apart from specialization being a riskier choice, at the same time, it also means that a person's risky labor market outcomes correlate less with the average risky labor market outcomes of others. Let us assume that the economy consists of a set of industries that each are subject to industry-specific shocks as well as to shared general shocks. Agents can now either be generalists, who can work in every industry, or they can be specialists, in which case they are restricted to a particular industry but also earn higher wages. Generalists are only subject to general shocks; their risky labor market outcomes correlate positively with the average outcome of the market. Specialists are additionally subject to the respective industry-specific shocks; their risky labor market outcomes correlate less with the average outcome of the market. We therefore expect that people with a preference for negatively

[2] In its most general form, herding can be defined as behavioral patterns that are correlated across individuals. More narrowly, it is defined as a group of investors following each other into (or out of) the same securities over some period of time. Herding can lead to systematically erroneous decision-making by populations and is closely linked to phenomena such as imperfect expectations and bubbles.

correlated uncertain outcomes are more likely to specialize and focus on certain areas when acquiring human capital. Furthermore, it is commonly observed that women seek less specialization than men. This finding is, amongst others, used to explain the gender wage gap (Tam, 1997).

In our experiment, we additionally elicit school grades in order to compute a tractable though potentially simplistic measure for specialization. We analyze if the gender difference in preferences over the correlation between uncertain outcomes can explain why women specialize less than men. Indeed, we find that women in our sample specialize slightly less than men. However, this effect is not explained by the difference in preferences over uncertain outcomes, and, overall, we do not find a robust correlation between preferences over the interpersonal correlation between uncertain outcomes and specialization.

This chapter proceeds as follows. In Section 2.2, we summarize the relevant literature and present theoretical considerations as well as the resulting hypotheses. In Section 2.3, we describe the experimental design. In Section 2.4, we present the results before arriving at our discussion along with concluding remarks in Section 2.5.

## 2.2 Theoretical Background and Hypotheses

Decisions under uncertainty and risk preferences are widely studied areas in economics. Starting out with the expected utility theory of von Neumann and Morgenstern (1944), the framework has been extended and modified by introducing subjective probabilities (Savage, 1954; Anscombe and Aumann, 1963), risk aversion (Pratt, 1964; Arrow, 1974) and loss aversion (Kahneman and Tversky, 1979), to name just a few of the most influential contributions. Apart from these theoretical models, extensive empirical research has been conducted and various methods to measure risk preferences and estimate utility functions have been developed (see Charness et al. (2013) for an overview).

As mentioned before, all of these models have in common that they only consider individual risk, disregarding the risk that other people face at the same time. However, there exists a theoretical branch of literature on the evolutionary development of preferences that analyzes the evolutionary advantages and disadvantages of correlated vs. uncorrelated risky outcomes. The conclusion that is commonly drawn based on various models is that, given that risks are the same from an individual point of view, positively correlated uncertain outcomes are a greater danger for the survival of a population than uncorrelated uncertain outcomes (e.g. Cooper and Kaplan, 1982; Robson, 1996; Curry, 2001). Thus, groups of individuals choosing negative correlations between individual outcomes tend to have higher levels of evolutionary fitness than groups who do the op-

posite, and one could expect the former preference to have prevailed. Indeed, it has been shown that different correlations between outcomes can shape economic preferences with persistent consequences (Galor and Savitskiy, 2018).

However, there also exist several motives for choosing the positively correlated uncertain outcomes. These reasons include, amongst others, inequity aversion and social support. Positively correlated risky outcomes typically mean that payoffs for people are the same or at least similar, whereas negatively correlated risky outcomes mean different payoffs. Inequity aversion, or fairness preferences (Fehr and Schmidt, 1999), could therefore lead people to choose the positively correlated outcomes. Furthermore, it has been shown that people have a desire to share their emotions with others during difficult life events, but also during happy times as is described in the social support literature (see for example Kaplan et al. (1977)). A motive for avoiding negatively correlated risky payoffs could therefore be that people prefer to have social support and someone with whom to share the experience of the gain or the loss.

Since theoretical economic arguments do not, in sum, deliver a clear prediction for whether people should be expected to choose one option or the other in our experimental setup, the main part of our experiment is of an exploratory nature. Our main aim is to find out which preferences people have over the interpersonal correlation between uncertain outcomes.

**Research Question.** *Which proportions of people prefer a positive or a negative interpersonal correlation between uncertain outcomes?*

Conditional on finding some heterogeneity in people's preferences, meaning that motives differ in their strength across people, we would expect a gender difference. From past research, we know that women have more egalitarian preferences than men (Fehr et al., 2006). Furthermore, women are more likely to look for social support from others and also provide it more often (Flaherty and Richman, 1989; Belle, 1991). Since we consider these behaviors to be indicators for preferences for positively correlated risky outcomes, we expect women to choose the positively correlated risky outcomes more often than men, which is formulated in Hypothesis 2.1.

**Hypothesis 2.1.** *Women choose the positively correlated uncertain payoffs with a higher probability than men.*

Hypothesis 2.2 concerns the relationship between preferences over the interpersonal correlation between uncertain outcomes and the decision to acquire specialized human capital as opposed to becoming a generalist. As elaborated in detail before, specialization means that a person's risky labor market outcomes correlate less with others' risky labor market outcomes. We therefore hypothesize that people with a preference for negatively

correlated uncertain outcomes are more likely to specialize and to focus on certain areas when acquiring human capital. Women commonly seek less specialized jobs than men (Tam, 1997). We therefore expect women to be less specialized than men and hypothesize that the weaker specialization of women can be partially explained by a lower preference for negatively correlated uncertain outcomes.

**Hypothesis 2.2.** *Subjects who choose the positively correlated uncertain outcomes are on average less specialized. Women are less specialized than men. This gender difference in specialization can be partially explained by women having stronger preferences for positively correlated uncertain outcomes.*

## 2.3 Experimental Design and Procedure

We conducted a laboratory experiment in December 2018 at the Cologne Laboratory for Economic Research using the softwares Qualtrics and ORSEE (Greiner, 2015). For every session, 32 participants were invited. Each session lasted 15 minutes and consisted of two parts. Participants received a fixed payoff of €5 and had the chance to earn an additional €10. The payoff-relevant decision was made in Part 1. Instructions translated from German can be found in Appendix 2.B.

In Part 1, subjects were told that two participants would be randomly selected at the end of the experiment and that for these two participants, the choice from Part 1 would be implemented. One of the two selected subjects would be the active participant and one of them would be the passive participant. The additional payoff of the two participants would be determined by a coin toss. The coin would show "heads" with a probability of 50% and "tails" with a probability of 50%. The winning side of the passive participant would be randomly determined by the computer before the coin toss. If the passive participant's winning side showed up, he would receive an additional payoff of €10; if the losing side showed up, he would receive an additional payoff of €0. The active participant's winning side would be determined by herself. She could choose in Part 1 if—in case she was randomly determined to be the active participant at the end of the experiment—she would like to have the same winning side as the passive subject or if she would like to have the other winning side. If the active participant chose the same winning side as the passive participant, and if this was the side that showed up during the random coin toss, the active and the passive participant both received an additional €10. If the other side showed up, both received an additional €0. If the active participant chose the other winning side and this side showed up, the active participant received an additional €10 and the passive participant €0. If the active participant's losing side showed up, the active participant received an additional €0 and the passive participant

€10. As can be seen, the individual chance of winning of the active participant did not change with the side she chose as winning side; neither changed the chance of winning of the passive participant. Furthermore, the total expected payoff for both participants together stayed the same (€10), irrespective of the chosen winning side. However, if the same side was chosen, the payoffs of the participants were positively correlated. If the other side was chosen, then the payoffs of the participants were negatively correlated. Subjects received two examples to further explain the setup and had to answer two control questions. In Part 1, all participants chose if they would want the same or the other winning side in case they were later chosen to be the active participant.

In Part 2, subjects answered a questionnaire. The questionnaire included demographic questions (gender, age, field of study, income). Furthermore, we asked them for 17 common school subjects (e.g. math or German) if they had been enrolled in this subject during their last years of high school and which had been the last grade they had received. We use the standard deviation between grade averages in different fields of subjects as a measure of specialization, as we will later discuss in more detail.

## 2.4 Results

The 267 subjects who participated in the experiment are of an average age of 25.4 years; 58.8% of them are women. For summary histograms on all elicited variables, see Appendix 2.A.

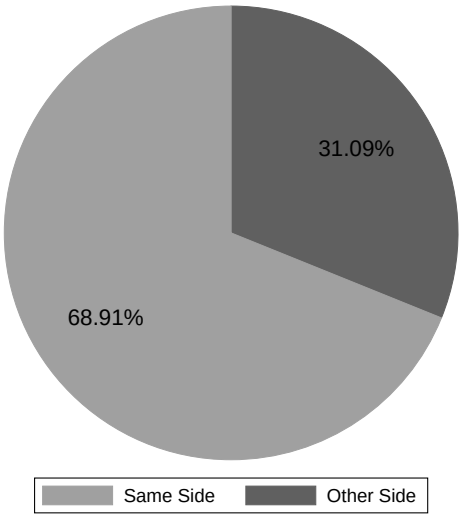



Figure 2.1: Shares of subjects choosing the same and the other side of the coin, respectively

Figure 2.1 shows the fraction of participants choosing the same and the other coin side, respectively. We find that 68.91% of all subjects choose the same winning side (positively correlated uncertain outcomes), which is significantly different from 50%, which would be expected if people simply randomized (two sided binomial test, p-value $< 0.001$). Furthermore, also when analyzed separately, men as well as women both choose the positively correlated uncertain outcomes significantly more often than the negatively correlated uncertain outcomes.

As predicted by Hypothesis 2.1, there is a gender difference of 10.5 percentage points in choosing the positively correlated uncertain outcomes. 73.3% of women choose the same coin side, whereas only 62.7% of men do so. Table 2.1 shows regression results concerning this gender difference. In Column 1, the dummy variable "Same" (1 if the same coin side was chosen and 0 if the other coin side was chosen) is regressed on the dummy variable "Female" (1 if subject was female and 0 if subject was male). The difference of 10.5 percentage points is significant at a significance level of 10%.

| | *Dep. variable:* Same | |
|---|---|---|
| | (1) | (2) |
| Female | $0.105^{*}$ | $0.117^{**}$ |
| | (0.0583) | (0.0580) |
| Age | | $0.00781^{****}$ |
| | | (0.00205) |
| Constant | $0.627^{****}$ | $0.422^{****}$ |
| | (0.0463) | (0.0777) |
| Observations | 267 | 267 |
| $R^2$ | 0.0125 | 0.0295 |

*Note:* OLS regression; $^{*}$ $p < 0.1$; $^{**}$ $p < 0.05$; $^{***}$ $p < 0.01$; $^{****}$ $p < 0.001$; robust standard errors in parentheses.

Table 2.1: Effect of gender and age on probability of choosing the same coin side

Since women in our sample are on average 1.5 years younger than men and since choosing the same side is positively correlated with age (correlation coefficient of 0.119), we additionally control for age (and thereby for the cohort) in years in Column 2. Age enters positively and highly significantly. The coefficient of Female increases slightly and is now significant at a significance level of 5%. The strong effect of age on choosing the positively correlated payoffs is driven entirely by participants that are older than 35 (15 subjects in total). Of these 15 subjects, seven are women and eight are men; 14 of them choose the positively correlated risk. It therefore seems that preferences for a positive correlation of uncertain outcomes increase as people get older. However, since we are running our experiment with a student sample, we do not have enough variation in our

data to draw robust conclusions. It would be interesting to investigate the development of the preferences on the correlation between risky payoffs over the life cycle.

We now turn to Hypothesis 2.2, which claims that men are more specialized than women and that this gender difference can partially be explained by women having stronger preferences for positively correlated uncertain outcomes than men. We compute a proxy for a particular, potentially important decision about specialization in subjects' past: effort invested in school subjects. In Part 2 of the experiment, we ask participants to report the last grades they received in 17 different school subjects during their last years of high school on the German scale for school grades from 1 to 6 (whereas 1 is the best grade and 6 is the worst grade). We group all 17 school subjects into one of four categories.[3] For each participant, we then calculate the grade average for each of these categories and use the standard deviation between those grade averages as our measure of specialization. We decided to first group the school subjects into categories to reduce measurement error and account for the fact that people were enrolled in different numbers of school subjects. Moreover, fields correspond to broad university study areas, whereas for many academic subjects there do not exist corresponding individual school subjects (e.g. economics is usually not taught at German high schools). People, who are very good in one field and very bad in others (hence are specialized in one area), score high on the variable "Specialization". In contrast, people who are similarly good in all fields (and hence generalists), score very low on the variable. Analyzing the degree of specialization, we exclude all participants reporting grades on a different scale than the one that they were asked to use (1 to 6). Doing so, we end up with 202 observations.

Our measure of specialization is possibly problematic since it is not clear if a higher standard deviation in school grades actually translates into occupations that require more specialized human capital. We later discuss measures that might be more suitable. However, we argue that school grades and the decision on which ones to focus is an early form of specialization. Students have limited resources (such as time, intelligence, etc.) and have to decide in which school subjects to invest. Given that they are not bounded at the top (their resources are insufficient to receive the top grade in all subjects; we further assume that everybody is capable of achieving grades that at least in some subjects are better than the bottom grade), they can decide to distribute their effort evenly across fields, or they can decide to focus all their resources on only a few fields. Distributing resources evenly leads to similar grades and therefore a low standard deviation. Focusing on certain areas leads to good grades in these areas and to bad ones in the neglected areas

[3] Mathematics, Computer Science, Physics, Biology, and Chemistry are grouped into category "Natural Science". German, English, French, Latin, Greek, and Spanish are grouped into category "Languages". Geography, History, Religion, and Politics are grouped into category "Social Science". Sports and Art are grouped into category "Others".

and therefore to a high standard deviation. This early form of specialization can translate into specialization in later life since school grades—and also the acquired knowledge behind them—pave the way to apprenticeships and college degrees: when the time comes to choose a future career, students with low specialization (generalists) still have all their options open but are suboptimally educated for any area, whereas students with high specialization (specialists) are already restricted to certain areas in which they are, however, better educated.

Since students with a grade average of 1 (seven observations in our sample) do not seem to face a binding budget constraint regarding their (efficiency units of) effort and, by construction, have a standard deviation of 0, we drop them from our sample.[4] Our remaining sample therefore consists of 195 observations.

| | *Dep. variable:* Specialization | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Female | -0.146** | -0.114** | -0.112** |
| | (0.0580) | (0.0527) | (0.0532) |
| Grade Average | | 0.509*** | 0.512*** |
| | | (0.179) | (0.180) |
| Grade Average Squared | | -0.0620* | -0.0630* |
| | | (0.0345) | (0.0348) |
| Age | | | 0.00127 |
| | | | (0.00284) |
| Constant | 0.869**** | -0.00935 | -0.0431 |
| | (0.0488) | (0.225) | (0.233) |
| Observations | 195 | 195 | 195 |
| $R^2$ | 0.0350 | 0.168 | 0.169 |

*Note:* OLS regression; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

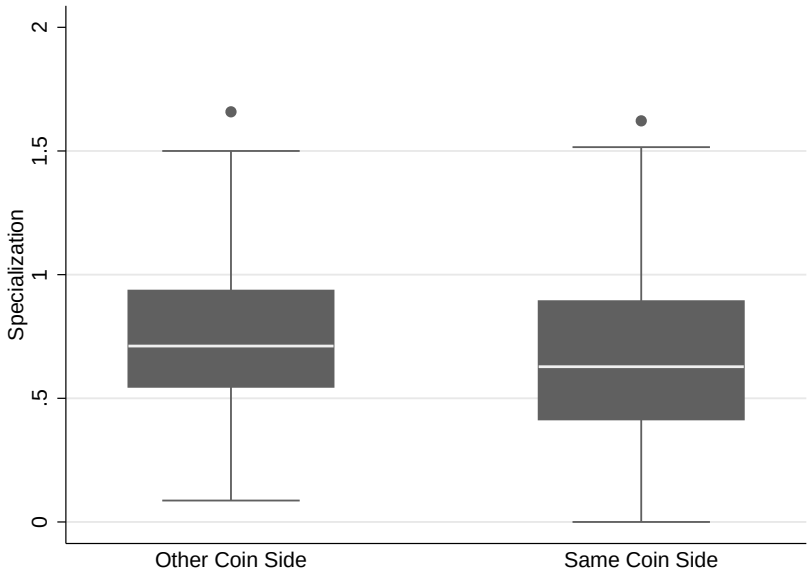
Table 2.2: Effect of gender, grade average, squared grade average, and age on specialization

Table 2.2 shows regression results analyzing if there is a gender difference in specialization in our sample. In Column 1, the variable Specialization is regressed on the dummy variable Female. As expected, the coefficient is negative and significant at a significance level of 5%. Women in our sample are around 16.8% less specialized than men. However, one drawback of the variable Specialization is that grades are capped, which mainly becomes relevant for very good students. Since many people with very good grades have a very low standard deviation in grades, whereas few people are very bad

[4] Students with a grade average of 6 would also have a standard deviation of 0. However, none of our subjects has this grade average.

in everything (they probably would have had to repeat the year), we additionally control for the average grade of all school subjects and the squared average grade in Column 2. The coefficient of Female is still negative but increases slightly. Results stay almost unchanged when additionally controlling for age in Column 3. Our findings are therefore in line with the first part of Hypothesis 2.2.

Next, we look at the general relation between specialization and preferences over the interpersonal correlation between uncertain outcomes. Figure 2.2 shows boxplots of the variable Specialization, one plot for subjects who choose the negatively correlated risky payoffs and one plot for subjects who choose the positively correlated risky payoffs. The point estimates for the median of the distribution of specialization are slightly lower for subjects choosing the positive correlation relative to subjects choosing the negative correlation. This is in line with Hypothesis 2.2, but the effects are weak and insignificant. The same holds when looking at the means. A Mann-Whitney *U* Test is also unable to reject the null hypothesis of the distributions being identical in both subsamples (p-value = 0.16).




*Note:* Boxplots of variable Specialization for subjects who choose the other coin side and subjects who choose the same coin side. The lines inside the box represent the medians, the lower and upper bounds of the box the lower and upper quartiles, respectively. The lower and upper lines outside the box represent the minimum and maximum values inside 1.5 times the interquartile distances, respectively. Dots outside are outliers.

Figure 2.2: Boxplots of specialization

We now turn towards the question if the gender difference in specialization can partially be explained by the gender difference in the preferences over the correlation be-

tween uncertain outcomes. This question is analyzed in Table 2.3. Column 1 is the same as Column 1 of Table 2.2; specialization is regressed on the dummy variable Female. In Column 2, we additionally add the dummy variable for choice of the same side of the coin to the regression. The variable Same does not enter significantly and the coefficient of Female stays almost unchanged. In Column 3 and Column 4, we additionally control for the grade average, the squared grade average, and for age. The coefficient for Same remains insignificant. However, the point estimate becomes absolutely larger, and the point estimate for the coefficient of Female becomes correspondingly absolutely smaller at the same time. This is generally consistent with Hypothesis 2.2, but due to lack of statistical power the results at this point remain inconclusive.

We can therefore not confirm Hypothesis 2.2. Women in our sample choose the positively correlated risky payoffs more often than men, and they are weakly significantly less specialized. However, we do not find sufficient evidence for a relationship between specialization and preferences over the correlation between uncertain outcomes.

| | *Dep. variable:* Specialization | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female | -0.146** | -0.141** | -0.108** | -0.103* |
| | (0.0580) | (0.0585) | (0.0530) | (0.0537) |
| Same | | -0.0770 | -0.0894 | -0.0940 |
| | | (0.0616) | (0.0562) | (0.0574) |
| Grade Average | | | 0.512*** | 0.516*** |
| | | | (0.177) | (0.178) |
| Grade Average Squared | | | -0.0620* | -0.0636* |
| | | | (0.0341) | (0.0343) |
| Age | | | | 0.00198 |
| | | | | (0.00291) |
| Constant | 0.869**** | 0.922**** | 0.0454 | -0.00456 |
| | (0.0488) | (0.0623) | (0.232) | (0.237) |
| Observations | 195 | 195 | 195 | 195 |
| $R^2$ | 0.0350 | 0.0432 | 0.179 | 0.181 |

*Note:* OLS regression; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 2.3: Effect of gender, choosing the same coin side, grade average, squared grade average, and age on specialization

## 2.5 Conclusion

We run a simple laboratory experiment to elicit preferences over the interpersonal correlation between uncertain outcomes. In the experiment, subjects choose between a

lottery in which payoffs are positively correlated with the payoffs of another subject and a lottery in which payoffs are negatively correlated with the payoffs of another subject.

We find that around 69% of all subjects choose the positive correlation; only 31% choose the negative correlation. We furthermore find that significantly more women than men prefer positively correlated uncertain outcomes. Women in our sample are also less specialized than men. However, this gender difference in specialization cannot be explained by the gender difference in preferences over the correlation between uncertain outcomes.

We see our study as the first step in a larger research agenda investigating preferences over the interpersonal correlation between uncertain outcomes. One next and essential step would be to find more evidence that the decision subjects make in our experiment reflect a stable preference. In order to do so, we suggest an experiment in which people make similar decisions at different points in time and to test the correlation between those decisions. Furthermore, it would be important to find out which role inequity aversion plays for agents when deciding about correlated payoffs. In our experiment, choosing the positively correlation also lowers inequality in payoffs between the two agents. This could shed light on the degree to which we are dealing with a behavioral primitive or a consequence of underlying preferences that are already well-known. It would therefore be important to check if preferences over correlated risky outcomes and inequity aversion are two independent concepts. Another step would be measuring preferences over the interpersonal correlation between uncertain outcomes in more detail and eliciting people's willingness to pay for a positive correlation. This could be done by offering not only a binary choice between perfectly negative and perfectly positive correlations but different correlations in between. This would also allow for a more precise measurement of gender differences.

Our hypothesis on the relation between preferences on the interpersonal correlation between uncertain outcomes and specialization requires more investigation. First of all, our findings regarding the determinants of specialization could be improved immensely and be shown more robustly by using a more precise measure of specialization. One caveat of the measure we are employing is that it is not clear if a higher standard deviation in school grades actually translates into occupations that require more specialized human capital. When working with a student population, however, it is not clear which other information could be used. A more robust approach could be to run the experiment with a non-student population eliciting data on actual occupations and combine this with information on which occupations require higher levels of specialization and which ones lower levels. Alternatively, one could use data on real job specialization together with elicited school grades to find support that a higher standard deviation in

school grades translates into more specialized occupations.

Finally, not only preferences over the interpersonal correlation between uncertain outcomes should matter for choosing the amount of specialization, but also risk preferences themselves. In our experiment, risk is not an issue because personal risks are the same for both lotteries between which participants choose. It would still be interesting to see if there is some sort of correlation between preferences over the correlation between uncertain outcomes and risk preferences in themselves. We would suggest eliciting personal risk preferences to shed more light on this question.

## References

Anscombe, Frank J., and Robert J. Aumann. 1963. "A Definition of Subjective Probability." *The Annals of Mathematical Statistics* 34 (1): 199–205.

Arrow, Kenneth. 1974. *Essays in the Theory of Risk Bearing*. Amsterdam: North-Holland.

Belle, Deborah. 1991. "Gender Differences in the Social Moderators of Stress." In *Stress and coping: An anthology*, ed. by A. Monat and R. S. Lazarus, 258–274. New York: Columbia University Press.

Chambers, Matthew, Carlos Garriga, and Don E. Schlagenhauf. 2009. "Accounting for Changes in the Homeownership Rate." *International Economic Review* 50 (3): 677–726.

Charness, Gary, Uri Gneezy, and Alex Imas. 2013. "Experimental Methods: Eliciting Risk Preferences." *Journal of Economic Behavior & Organization* 87:43–51.

Cooper, William S., and Robert H. Kaplan. 1982. "Adaptive 'Coin-Flipping': a Decision-Theoretic Examination of Natural Selection for Random Individual Variation." *Journal of Theoretical Biology* 94 (1): 135–151.

Curry, Philip A. 2001. "Decision Making under Uncertainty and the Evolution of Interdependent Preferences." *Journal of Economic Theory* 98:357–369.

Devenow, Andrea, and Ivo Welch. 1996. "Rational Herding in Financial Economics." *European Economic Review* 40:603–615.

Fehr, Ernst, Michael Naef, and Klaus M. Schmidt. 2006. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review* 96 (5): 1912–1917.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

Flaherty, Joseph, and Judith Richman. 1989. "Gender Differences in the Perception and Utilization of Social Support: Theoretical Perspectives and an Empirical Test." *Social Science and Medicine* 28 (12): 1221–1228.

Galor, Oded, and Viacheslav Savitskiy. 2018. *Climatic Roots of Loss Aversion*. NBER Working Paper 25273. Cambridge: National Bureau of Economic Research.

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

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–292.

Kaplan, Berton H., John C. Cassel, and Susan Gore. 1977. "Social Support and Health." *Medical Care* 15 (5): 47–58.

Oswald, Andrew J. 1996. *A Conjecture on the Explanation for High Unemployment in the Industrialized Nations: Part I*. Warwick Economic Research Paper 475. Coventry: The University of Warwick.

Pratt, John W. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32 (1-2): 122–136.

Robson, Arthur J. 1996. "A Biological Basis for Expected and Non-Expected Utility." *Journal of Economic Theory* 68 (2): 397–424.

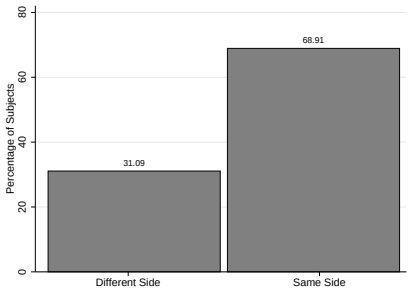
Savage, Leonard J. 1954. *The Foundation of Statistics*. New York: Dover.

Sias, Richard W. 2004. "Institutional Herding." *Review of Financial Studies* 17 (1): 165–206.

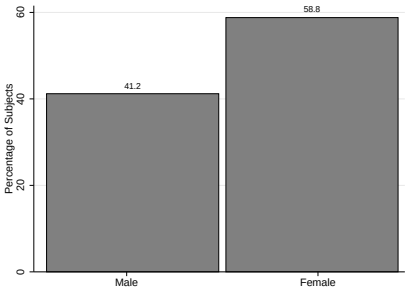
Tam, Tony. 1997. "Sex Segregation and Occupational Gender Inequality in the United States: Devaluation or Specialized Training?" *American Journal of Sociology* 102 (6): 1652–1692.

Von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

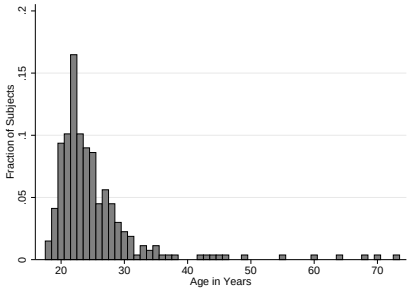
# 2.A Summary of Variables



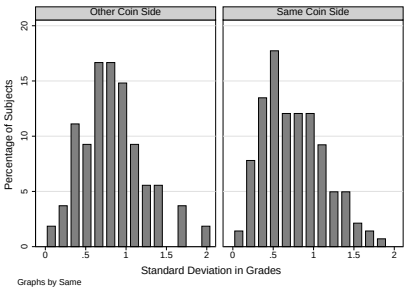

(a) Main Choice from Part 1



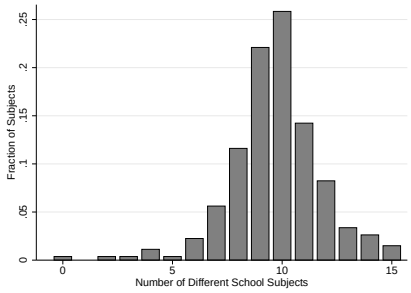

(b) Gender



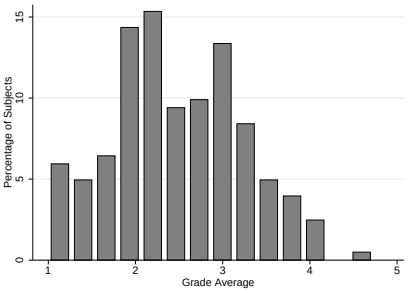

(c) Age




(d) Specialization




(e) Number of School Subjects




(f) Grade Average

Figure 2.3: Summary histograms

## 2.B Instructions

**Welcome!**

Please turn off your phone and do not communicate with other people.
If you have any questions, please raise your hand. An experimenter will then come to your place. Please stay seated quietly for the duration of the experiment in order to not disturb anyone.
Please enter your cubicle number into the following field:

**General Information**

This experiment will last up to 15 minutes and is made up of two parts.
In Part 1, you will make one decision.
In Part 2, you will answer a questionnaire.

**Your Compensation**

You will receive €5 for your participation. Additionally, you can earn a further payoff.
Two (out of a maximum of 32) participants of this experiment will be randomly selected by the computer at the end of the experiment.
The additional payoff of the two randomly selected participants will be determined by Part 1. All other participants do not receive any additional compensation.
For the sake of clarity, instructions for Part 1 are formulated as if the computer had selected you at the end of the experiment.

**Part 1**

Please read the following instructions carefully and answer the control questions.

**Your Teammate**
The second person that is randomly selected by the computer will be your teammate.

**Coin Toss**
In Part 1, your payoffs will be determined by a fair coin toss. The result of the coin toss is "Heads" with a probability of 50% and "Tails" with a probability of 50%.
The coin toss determines if you receive €10 as an additional payoff or if you do not receive any additional payoff. You have a winning chance of 50%.
Furthermore, the coin toss determines also if your teammate receives an additional payoff of €10 or if he or she does not receive any additional payoff. Your teammate also has a winning chance of 50%.

At the end of the experiment, after all other participants have left the laboratory, the coin toss will be performed in front of you and your teammate by an experimenter.
Which side of the coin ("Heads" or "Tails") will be your winning side and which side of the coin will be your teammate's winning side will be determined by coincidence, your roles, and your decisions during Part 1.

**Active and Passive Role**
There are two roles, the active role and the passive role. The computer will randomly determine who of you and your teammate has the active and who the passive role at the end of the experiment.
The person with the passive role will have the winning side ("Heads" or "Tails"), which is randomly drawn by the computer.
The person with the active role can decide her- or himself if she or he wants the same winning side as the passive person or if she or he wants the other side.

**Additional Payoff**
If the active person chooses the **same** winning side as the passive person, the participants either **both win €10** or **both win nothing**.
If the active person chooses the **other** side as winning side, then always **one of the participants wins €10** (either the active or the passive person), and **the other one wins nothing**.

**Example 1**
Assume the following: you drew the active role and your teammate randomly received the winning side "Heads". You chose the same winning side. If the coin toss now resulted in "Heads", you would receive €10 and your teammate would receive €10. If the coin toss now resulted in "Tails", then you and your teammate would both receive nothing.

**Example 2**
Assume the following: you drew the active role and your teammate randomly received the winning side "Heads". You chose the other winning side. If the coin toss now resulted in "Heads", you would not receive anything and your teammate would receive €10. If the coin toss now resulted in "Tails", you would receive €10 and your teammate would not receive anything.

**Control Question 1**
Assume the following: you drew the active role and your teammate randomly received the winning side "Heads". You chose the same winning side. The coin toss resulted in "Heads". Which answer is correct?

- You and your teammate do not receive an additional payoff.

- You and your teammate each receive additional €10.
- You receive additional €10 and your teammate does not receive anything.
- You do not receive anything and your teammate receives additional €10.

**Control Question 2**
Assume the following: you drew the active role and your teammate randomly received the winning side "Tails". You chose the other winning side. The coin toss resulted in "Heads". Which answer is correct?

- You and your teammate do not receive an additional payoff.
- You and your teammate each receive additional €10.
- You receive additional €10 and your teammate does not receive anything.
- You do not receive anything and your teammate receives additional €10.

**Your Decision**

You can now decide if you want the same winning side as your teammate or if you want the other winning side. Should you be randomly assigned to the active role, your decision will be implemented.
Should you have the active role and your teammate the winning side "Heads" and you choose the same winning side, then you and your teammate both receive €10 with "Heads" and both nothing with "Tails".
Should you have the active role and your teammate the winning side "Tails" and you choose the same winning side, then you and your teammate both receive €10 with "Tails" and both nothing with "Heads".
Should you have the active role and your teammate the winning side "Heads" and you choose the other winning side, then you do not receive anything and your teammate receives €10 with "Heads" and you receive €10 and your teammate nothing with "Tails".
Should you have the active role and your teammate the winning side "Tails" and you choose the other winning side, then you do not receive anything and your teammate receives €10 with "Tails" and you receive €10 and your teammate nothing with "Heads".

- Same Winning Side
- Other Winning Side

**Part 2**

We are now starting the questionnaire.
Age (in years):
Gender:
How much money do you have available every month?

Which subject do you study?
Please tell us for the following school subjects if you were enrolled in them during your last years of school.

- Math
- German
- English
- French
- Latin
- Greek
- Spanish
- Geography
- Art
- History
- Sports
- Computer Science
- Physics
- Biology
- Chemistry
- Religion
- Politics/Economics/Social Science

Please tell us for every school subject that you were enrolled in during your last years of school the last grade you received in this subject. Please tell us your grades on a scale from 1 (best grade) to 6 (worst grade). If you do not remember exactly, please give us your best approximation.

# Chapter 3

# On the Dynamics of Prosocial Behavior[1]

## 3.1 Introduction

People often behave in ways that are not in line with behavior of the selfish, output-maximizing *homo oeconomicus*: they volunteer, give money to charities, help others, and cooperate. This type of behavior has often been attributed to social preferences, which reflect the idea that a person's utility depends not only on their own payoff but also on other people's payoffs and behaviors.[2] In past research, it has been shown that social preferences positively affect economic success (Carpenter and Seki, 2011; Becker et al., 2012; Algan et al., 2014), and well-being (Dunn et al., 2008; Park et al., 2017) in several contexts. Furthermore—challenging the traditional understanding of preferences as being fixed—prosocial behavior has been found to be highly context dependent (Dana et al., 2007; Grossman, 2014) and even sustainably malleable (Kosse et al., 2019). The importance and malleability of prosociality naturally gives rise to questions on its development and on how this development can be fostered.

We experimentally test a theoretical approach conceptualizing the formation of prosocial attitudes that goes back to Aristotle's Nicomachean Ethics. According to Aristotle, virtues are formed by committing virtuous acts. Based on this idea, we argue that people's preferences become more prosocial when these people engage in prosocial behavior via the following mechanism: when people do something prosocial due to a stimulating environment, they adjust their self-perception. In order to avoid cognitive dis-

[1] This project is joint work with Frederik Schwerter.

[2] Common manifestations of social preferences are altruism (Becker, 1974, 1976), inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), or warm glow (Andreoni, 1990).

sonance, they thereafter maintain the more prosocial behavior even if there is no stimulating environment anymore. We call this effect *altruistic capital effect*, in which the term altruistic capital refers to the person's prosocial self-image. We test altruistic capital formation by investigating the effect of an environment that encourages prosociality on subsequent prosocial behavior. We hypothesize that people who act prosocially due to such an encouraging environment behave more prosocially in subsequent situations compared to people who were not in the same stimulating environment beforehand. We furthermore aim to find out more about the underlying change in self-image and test for changes in self-reported altruistic preferences.

This paper is inspired by recent theoretical literature on prosocial behavior and morality formation. Ashraf and Bandiera (2017) derive a model in which people form altruistic capital defined as an asset that enables individuals to internalize the effect of their actions on others and makes future altruistic behavior more likely. Importantly, altruistic capital is formed by engaging in altruistic behavior. Bénabou and Tirole (2011) offer an underlying mechanism that can explain why assuming altruistic capital formation is sensible. They develop a model in which agents gain utility from high self-esteem and make inferences about their true moral type by observing past moral and immoral actions. The model also comes to the conclusion that, under certain conditions, good actions build up *moral capital*[3] and lead to further good actions, whereas bad actions destroy moral capital and lead to further wrongdoing.

Especially in psychology, there is an abundant amount of empirical literature investigating the formation of prosocial or moral behavior in dynamic contexts. Evidence, however, is quite ambiguous. On the one hand, there is the *Foot-in-the-Door Effect* that refers to the phenomenon that the acceptance of a small initial request leads to a more probable acceptance of a greater subsequently made request. The effect has been empirically investigated and positive evidence has been found (Freedman and Fraser, 1966). Self-perception theory is most often used to explain it: the respondent feels helpful for accepting the small initial request and accepts again because he or she wants to maintain the new self-image (DeJong, 1979; Gorassini and Olson, 1995; Burger and Caldwell, 2003). Also in line with altruistic capital formation is the broader literature on habit formation in which people stick to well-pracitced behaviors (habits) because the processing that monitors their execution has become automatic (Ouellette and Wood, 1998). Past actions therefore have an effect on future actions, which has been shown for different domains, for example voting (Fujiwara et al., 2016). However, even though the observed effects are the same (performing an action in the past increases the likelihood of repeat-

[3]Moral capital is the term used by Bénabou and Tirole (2011) referring — in line with the term altruistic capital that we use — to the morality of the self-image.

ing the same action in the future), the underlying mechanism of habit formation is different to the mechanism underlying altruistic capital formation (automatisation vs. change of self-perception).

However, there is also evidence pointing in the opposite direction. People show *moral licensing* and *moral cleansing* behavior (see Merritt et al. (2010) for an overview). The term moral licensing refers to the effect that people who just did something they consider moral behave less morally afterwards because of the previous moral action—presumably because they feel "excused". Moral cleansing refers to the opposite effect. It refers to situations in which people who did something immoral show more moral behavior in subsequent situations because of the previous immoral action. These effects have been found for such different domains as discrimination and racism (Monin and Miller, 2001; Effron et al., 2009; Bradley-Geist et al., 2010; Merritt et al., 2012), environmental decisions (Mazar and Zhong, 2010), and consumption of luxury goods (Khan and Dhar, 2006).

Recently, also economists have taken interest in these phenomena and found that prosocial behavior decreases during repeated prosocial interactions in economic setups (Sass and Weimann, 2012; Brosig-Koch et al., 2017), but that this negative effect is weaker the longer the time period between these actions (Schmitz, 2019). However, it has been found empirically that the negative moral licensing effect shows when the first prosocial action was not costly, whereas the positive altruistic capital effect prevails given that the first prosocial action was costly and therefore actually carried a signal on the prosociality of the participant (Gneezy et al., 2012).

But how could these opposing effects coexist? We have the following mechanism in mind, which we explain in more detail in Section 3.3: individuals have a self-image, a certain belief about their level of prosociality, and act according to this self-image to avoid cognitive dissonance. This leads to a "normal" level of prosociality, which stays the same on average and comprises people's altruistic preferences. Moral licensing and moral cleansing effects are short term fluctuations around this natural level due to moral accounting, offsetting earlier outliers. However, altruistic capital formation refers to changing the self-image and therefore the whole level of "natural" prosociality. We imagine this effect to kick in after prosocial behavior has been abnormally high (or low) over a longer time span, or after actions carrying strong signals (e.g. if they were very costly or observed by others), which might be induced and enhanced by an environment fostering prosocial behavior.

In order to test our hypothesis, we conduct a laboratory experiment in which participants work on different tasks over several weeks. We vary between treatments if solving the tasks leads to prosocial charitable donations or payments for the subjects them-

selves. In a final session, we measure subjects' levels of prosociality via a real-effort task. We find a weakly significant negative treatment difference in this subsequent prosociality when controlling for self-reported altruism. Furthermore, we do not find a change in altruistic preferences for the treatment group. We discuss these findings—that contradict our hypotheses—in the last part of this chapter.

This chapter proceeds as follows: In Section 3.2, we describe the experimental design, Section 3.3 derives our hypotheses, Section 3.4 contains the analysis and our results and Section 3.5 concludes.

## 3.2 Experiment

We ran an experiment using the infrastructure of the BonnEconLab in May, June, and July 2016. 50 participants were invited using hroot (Bock et al., 2014); the surveys were run using Qualtrics. The duration of the experiment was six weeks and it consisted of ten online sessions (two per week), as well as of one final laboratory session. We restricted the experiment to subjects who had never participated in an experiment of the BonnEconLab before and were 18, 19, or 20 years old. Before starting the experiment, participants were randomized into one of two treatment groups: the Control Treatment (C), in which subjects completed tasks for receiving vouchers to shop for themselves at the online store Amazon, and the Altruistic Treatment (A), in which subjects completed tasks in order to generate donations for a good cause. All subjects were reimbursed with a payment of €100 for participating in all eleven sessions. Missing more than one session resulted in lower payoffs in order to avoid high attrition rates. The compensation was paid after the last laboratory session.

During the first five weeks of the experiment, subjects received an email containing a link to an online survey at 9 a.m. on every Monday and on every Thursday. They had time to finish the survey until 11:59 p.m. of the respective day. Afterwards, the links lost their validity. Subjects who did not finish the first or the second survey were excluded from the experiment. Subjects who did not finish one of the other surveys received a reminder email and all following survey links. The overall structure and timeline of both treatments is illustrated in Figure 3.1. Instructions translated from German can be found in Appendix 3.B.

In Session 1, subjects answered a short survey (S1) containing demographic questions. In Session 2, subjects answered a longer survey (S2), in which we elicited risk, time, and social preferences, as well as life satisfaction and psychological personality measures (Big Five (Gerlitz and Schupp, 2005)). Furthermore, subjects were asked to complete five sets of the real-effort task *Counting Zeros* by Abeler et al. (2011), which was also used to

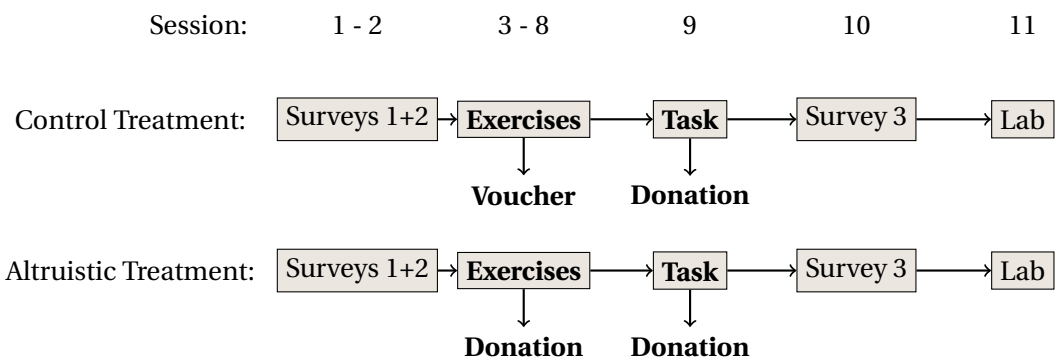



Figure 3.1: Timeline of experiment

generate the main donation later on. We describe the task in detail below. Session 1 and Session 2 were the same for both treatment groups.

In Session 3 to Session 8, subjects worked on simple exercises that took between five and fifteen minutes to complete per session. Upon solving a minimum of exercises correctly, subjects in the Control Treatment received a €5-Amazon voucher via email the following day. Subjects in the Altruistic Treatment did not receive anything themselves. Instead, upon solving the same minimum of exercises correctly, a donation was made to a project of the charity *SOS Kinderdörfer* the following day. After the donation had been made, the participants received a confirmation email. They were also informed about the possibility to review the donation receipts during the final laboratory session. The specific project the donation was made to changed with every session. All charitable projects were related to humanitarian causes. Upon starting each session, subjects in Treatment A received a short description of the project[4], which subjects of the Control Treatment did not read. Otherwise, instructions were identical. To stay in the experiment and to receive the final reimbursement of €100, a successful completion of the exercises was not required. Subjects in both treatments therefore had the choice to just click through the exercises and not work at all. Working for the donation or the voucher was therefore not mandatory, but depended on each subject's own decision.

One possible caveat of choosing to pay subjects in Treatment C more than subjects in Treatment A via Amazon vouchers is that the increased payoff could have led to the effect that subjects in Treatment C felt more obliged to work in return and therefore increased their later prosocial behavior. However, we still decided to do so for two reasons. First, donations and prosocial behavior in general are usually costly and therefore go along with a decreased endowment. Second, the gift exchange effect goes into the opposite

[4]Project descriptions were all based on descriptions of *SOS Kinderdörfer*'s official homepage (SOS Kinderdörfer, 2016). Information on poverty was based on information provided by Aktion Deutschland Hilft (2016).

direction of our hypothesized effect; therefore, our design is a conservative test in that regard.

Figure 3.2 shows an overview of which exercises were used in Session 3 to Session 8. In Session 3 and Session 6, the exercise was solving easy Raven matrices. In Session 4 and Session 7, the exercise consisted of translating a sequence of ciphers into a word using a translation key. In Session 5 and Session 8, the exercise was to partially solve a Sudoku riddle (see Appendix 3.C for screenshots of the exercise screens and Appendix 3.B for detailed instructions and exercise descriptions). Exercises were chosen to be easy, yet somewhat time consuming in order to make them solvable for everyone but requiring to work for the payoff or donation nonetheless.

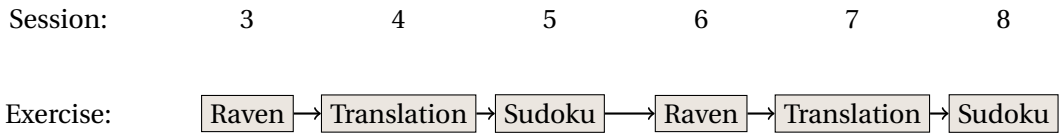



Figure 3.2: Exercises in Session 3 to Session 8

Session 9, Session 10, and Session 11 were again identical for both treatment groups. In Session 9, all subjects participated in the same real-effort task in order to generate a donation. The real-effort task was the task Counting Zeros. On each screen, subjects saw a 12×12-matrix with each field containing either a zero or a one. Subjects were asked to count the zeros. For every correctly solved task, a donation of €0.20 was made to a project supporting Ebola-orphans of the charity SOS Kinderdörfer. They also received a short description of the project, which was new to all of the participants. For every falsely counted matrix, €0.05 were subtracted from the donation to prevent people from guessing.[5] Subjects could stop working at any moment by entering the word "stop" into a field and were made aware that this would in no way effect their own final payment. Subjects could solve up to 40 tasks, which would have taken approximately 40 to 60 minutes. Therefore, the maximum donation was €8. Subjects could see their current donation at any time and received feedback on their answer after every table. The task was chosen such that it required a significant amount of concentration and effort to make it costly and cumbersome. As in any previous sessions, the donations were transferred the day after Session 9, and every participant received a confirmation email. Furthermore, a receipt of the overall donation was made available to all participants during Session 11.

In Session 10, subjects answered a survey (S3), which was the same as the one that subjects had answered in Session 2 (S2), to test for changes in preferences and personality measures between before and after the intervention.

[5] We allowed for a margin of error of +/−1. Also, the overall donation could not become negative.

Session 11, the last session, took place during week six at the BonnEconLab. In this session, subjects answered 20 multiple choice quiz questions concerning topics that were either general knowledge or had some connection to the charitable projects to which donations had been made. Subjects earned €0.20 per correctly answered question. They received direct feedback and the correct answer after answering each question and could see their current earnings at any given time. We included this session because we wanted to test for the possibility of people in the Altruistic Treatment acquiring information on the charitable projects and therefore increasing their willingness to donate. Since we do not find any difference in quiz results between treatments, we do not discuss the results further. Upon finishing the questionnaire, subjects received their final payment, which consisted of their reimbursement for the overall participation plus the amount earned during the quiz. Subjects could also review all the donation receipts. This concluded the experiment.

## 3.3 Conceptual Framework and Hypotheses

In deriving our hypotheses and designing our experiment, we had the following mechanism in mind. People hold a certain belief about who they are, or put differently, they have a self-image. Part of this self-image is an idea about how prosocial they are. In order to avoid cognitive dissonance (Akerlof and Dickens, 1982), people try to act in accordance with this self-image. If a person now performs a good deed, this satisfies the desire to re-affirm the self-image and leads to the effect that in the immediate time afterwards, prosocial behavior is not as necessary and can be reduced. This effect wears off after some time, after which a new good deed is necessary to keep up the good self-image. The other way around, in a short time frame, antisocial deeds will increase prosocial behavior directly afterwards. These short term fluctuations around the "inner" or "natural" level of prosociality can lead to the typically found moral licensing and moral cleansing effects that wear off after some time has passed.

However, we argue that this self-image determining the "normal" amount of prosocial behavior is malleable. More specifically, by placing a person in an environment that brings her to behave repeatedly prosocially over a longer time frame, she will update her self-perception after some time. This could happen due to a self-serving bias in attribution of positive events. This self-serving bias has been documented in psychological literature (see Malle (2006) for a summary) and refers to the finding that people tend to explain their successes (or positive behaviors) with personal characteristics, while they explain failures (or negative behaviors) with situational factors. Therefore, even if people are induced to behave prosocially mainly by the environment (by making prosocial

behavior cheap or occur as a side effect), people overestimate the role of their own personality and preferences and use the positive signals to update their self-image. This updated self-image will then lead to an increase in the overall level of prosociality, which leads to our first hypothesis about the effort provided in generating the donation.

**Hypothesis 3.1.** *Subjects who repeatedly commit prosocial acts due to a stimulating environment (Altruistic Treatment) behave more prosocially afterwards.*

Our first hypothesis concerns the effort that subjects exert in Session 9 when generating a donation for a good cause by working on the real-effort task Counting Zeros. The hypothesis says that subjects from the Altruistic Treatment, who donated up to six times in the previous sessions, will generate higher donations in this final task.

If the mechanism we describe above is what leads to an increase in prosocial behavior, we should also be able to see changes in survey responses measuring prosocial attitudes. This motivates our second hypothesis about the survey responses on altruistic preferences.

**Hypothesis 3.2.** *Subjects in the Altruistic Treatment report stronger altruistic preferences at the end of the experiment than at the beginning.*

The second hypothesis concerns the survey responses we elicit in Session 2 and in Session 10. In both sessions, we ask subjects how willing they would be to donate to a charity. The hypothesis says that we expect reports of altruistic attitudes to be higher at the end of the experiment in Treatment A, but not in Treatment C.

## 3.4 Results

Overall, 50 subjects signed up for the experiment of whom three subjects did not complete the first or second survey and are therefore excluded from the analysis. Of the remaining 47 participants, only five subjects missed one session and one subject missed two sessions. 23 subjects were randomized into Treatment A and 24 subjects into Treatment C. Table 3.1 shows summary statistics of key parameters for Treatment A and Table 3.2 for Treatment C.

Subjects are on average 19.7 years old, 64% of them are women. The variable Task Ability measures how many out of five tables of the real-effort task Counting Zeros subjects solve correctly when practicing the task in Session 2. On average, participants solve 3.2 tasks correctly. Life satisfaction and altruism are each measured on a scale from 0 to 10, values reported here refer to the responses given in Session 2 at the beginning of the experiment before the intervention. Subjects report on average a life satisfaction of 7.0 and altruism of 7.3. None of the differences in these variables between treatment groups

| Variable | Mean | Minimum | Maximum | Standard deviation | Observations |
|---|---|---|---|---|---|
| Female | 0.609 | 0 | 1 | 0.499 | 23 |
| Age | 19.784 | 18.529 | 20.712 | 0.647 | 23 |
| Task Ability | 3.217 | 1 | 5 | 1.204 | 23 |
| Happiness | 7.391 | 3 | 9 | 1.530 | 23 |
| Altruism Session 2 | 7.391 | 3 | 10 | 2.126 | 23 |

Table 3.1: Summary statistics Treatment A

| Variable | Mean | Minimum | Maximum | Standard deviation | Observations |
|---|---|---|---|---|---|
| Female | 0.667 | 0 | 1 | 0.482 | 24 |
| Age | 19.664 | 18.310 | 20.759 | 0.754 | 24 |
| Task Ability | 3.125 | 0 | 5 | 1.361 | 24 |
| Happiness | 6.708 | 2 | 10 | 2.095 | 24 |
| Altruism Session 2 | 7.250 | 3 | 10 | 1.939 | 24 |

Table 3.2: Summary statistics Treatment C

are significant to a significance level of 1% (Wilcoxon Mann-Whitney Test). We include histograms for all elicited variables in Appendix 3.A. All 47 subjects solve — given they participated in the respective session — the required number of exercises in Session 3 to Session 8 and therefore all either generate a donation or receive an Amazon voucher.

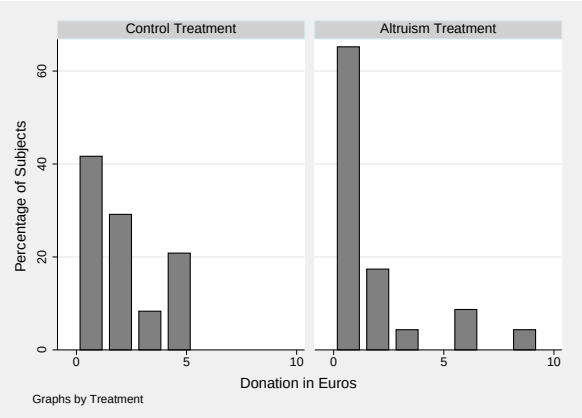



Figure 3.3: Donation Session 9

Figure 3.3 shows histograms of Session 9 donation, one for Treatment C and one for Treatment A. Overall, we do not find the predicted positive effect of Treatment A on Session 9 donation, but a negative moral licensing effect. On average, subjects in Treatment

A generate a donation of €1.81, subjects in Treatment C a donation of €2.15. We can see that there is a higher variation in donations in Treatment A, which we discuss in detail below. The distribution of donations is heavily skewed to the right. Almost all donations (85%) are between €0 and €4, which is in the lower half of possible donations; the maximum possible donation being €8. Therefore, we focus our analysis on logarithmic donations to decrease the effect of outliers and focus on differences in the lower ranges. We set donations of €0 (only three observations) to €0.05.

| | *Dep. variable:* Log Donation | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treatment A Dummy | -0.598 | -0.622* | -0.630* |
| | (0.366) | (0.355) | (0.355) |
| Altruism Session 2 | | 0.175** | 0.159** |
| | | (0.0817) | (0.0773) |
| Task Ability | | | 0.112 |
| | | | (0.150) |
| Constant | 0.415* | -0.852 | -1.087 |
| | (0.213) | (0.631) | (0.801) |
| Observations | 47 | 47 | 47 |
| $R^2$ | 0.0565 | 0.133 | 0.145 |

*Note:* OLS regression; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 3.3: Effect of treatment, initial altruism, and task ability on log donation

Table 3.3 shows the main regression result. In Column 1, log donation from Session 9 is regressed on a Treatment A dummy, which is 1 if participants were in the Altruistic Treatment and 0 if subjects were in the Control Treatment. The coefficient is negative and not significant. However, when controlling for self-reported altruism elicited in Session 2, i.e., before the treatment, in Column 2, the coefficient only changes slightly, but becomes significant at a significance level of 10%. Hardly anything changes when additionally controlling for task ability (measured during the practice phase in Session 2) in Column 3. This result is not in line with Hypothesis 3.1 and we discuss it below.

We now turn to Hypothesis 3.2, which claims that reports on altruistic attitudes should increase more for subjects in Treatment A. To measure participants' altruistic preferences, we ask them to state their willingness to donate to a good cause without expecting something in return on a scale from 0 to 10 in Session 2 and in Session 10. Table 3.4 repeats the regressions shown in Table 3.3 with variable Altruism Session 10 as the dependent variable. As can be seen, when controlling for Altruism measured in Session 2, the effect of the treatment on self-reported altruistic preferences is negative, but not significant. Again, the result is not in line with Hypothesis 3.2.

| | *Dep. variable:* Altruism Session 10 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treatment A Dummy | -0.0417 | -0.129 | -0.112 |
| | (0.606) | (0.481) | (0.470) |
| Altruism Session 2 | | 0.617**** | 0.650**** |
| | | (0.116) | (0.111) |
| Task Ability | | | -0.231 |
| | | | (0.229) |
| Constant | 7.042**** | 2.568*** | 3.053** |
| | (0.392) | (0.916) | (1.174) |
| Observations | 47 | 47 | 47 |
| $R^2$ | 0.000106 | 0.367 | 0.387 |

*Note:* OLS regression; * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 3.4: Effect of treatment, Session 2 altruism, and task ability on Session 10 altruism

To evaluate the change in altruistic preferences further, we subtract Session 2 answers from Session 10 answers. Figure 3.4 shows two histograms showing the distribution of the difference in altruistic preferences, one for Treatment C and for Treatment A.

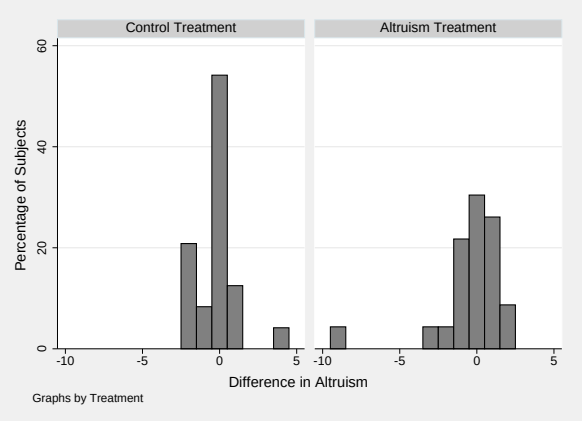



Figure 3.4: Difference in altruism measures between Session 10 and Session 2

On average, the difference is -.39 for subjects in Treatment A and -.21 in Treatment C (both measures of altruism are not significantly different between treatments), which does not confirm Hypothesis 3.2. Measures of altruistic preferences actually decrease in both treatments (even though measures are highly correlated between Session 2 and Session 10), and slightly stronger in Treatment A. However, we find an unexpected lower Session 9 donation in Treatment A. As we explain later on, we interpret this effect as moral licensing behavior. In this case, we would not expect a significant change in altru-

istic preferences. We also do not find any interesting treatment differences in any other variables.

Given the low sample size and the unexpectedness of results, it is difficult to draw conclusions from our results. However, let us point to a finding for the purpose of careful speculation. The standard deviations of the variables Session 9 donation and reported changes in altruism measures are significantly higher for Treatment A. For Session 9 donation, the standard deviation is 2.19 for Treatment A and 1.55 for Treatment C (p-value of 0.05 for a one-sided variance ratio test). For the difference in altruistic preferences (c.f. Figure 3.4), the standard deviation is 2.23 for Treatment A and 1.32 for Treatment C (p-value of 0.008 for a one-sided variance ratio test). The significantly higher variance for both variables represents that in Treatment A, there are subjects with very low donations as well as with very high donations. Similarly in Treatment A, there are subjects who dramatically decrease their reported altruism, but also those who increase it. In Treatment C, donations and altruism differences are less extreme.

This could possibly be due to two reasons. As we hypothesized before, we think that in order for our intervention to work and actually increase prosocial behavior over time, a long enough time frame for the self-image to adjust is essential and that in a too short time frame moral licensing effects could prevail. If it were the case that our treatment was just long enough for some people to show the expected effect, whereas for others it would have taken more time, the observed pattern could be explained. Another option would be that the treatment works for some type of people, whereas for others it does not work. For example, some people could have a more malleable self-image, or could be very open to prosocial behavior. However, as mentioned above, all of this is speculative, but might constitute interesting research questions for the future.

## 3.5 Discussion

We conduct a long-run laboratory experiment to shed light on the development of prosocial behavior and on the spillover effects of environments that stimulate prosociality. We find a weakly significant negative effect of such a stimulating environment on a subsequent donation decision. Moreover, we find that subjects in the treatment group do not increase their self-reported altruistic attitudes relative to subjects in the control group.

In deriving our hypotheses, we emphasize that a long enough time fame is crucial for altruistic capital formation to work. It could certainly be the case that in order to find a positive treatment effect, a period of three weeks of active intervention is not long enough for subjects to adapt their self-image. As mentioned in Section 3.4, this goes in line with our finding of higher standard deviations in donations and changes in altru-

istic attitudes in the treatment group: some people might have already adjusted while others might have needed more time. Our evidence at this point is merely suggestive and requires further exploring to find out which types of environments can sustainably increase prosocial behavior. It is also debatable if self-image and preferences are still malleable enough during adulthood. We therefore suggest a longer study with children in order to find out if stimulating environments can, over time, lead to more altruistic preferences.

Another possible issue with our experiment is the different compensation between treatments. Subjects in Treatment C receive €30 more in Amazon vouchers than subjects in Treatment A. With this design choice, our experiment more closely mimics real world environments, in which altruistic behavior is typically costly. However, this might lead to a gift exchange effect making people in Treatment C feeling obliged to donate more in order to compensate their high payoff. Indeed, when we ask subjects in a final comment section about their choice of work effort for generating the donation in Session 9, some of the participants in Treatment A mention that they tried to pay back the €30 they received in Session 3 to Session 8. This effect could mask the altruistic capital effect, which we are interested in. We would therefore suggest an experiment in which endowment is kept equal between treatment groups or varied in a 2×2 design.

Furthermore, it would be interesting to find out if there are other factors in an environment that can influence if altruistic capital is built up or not. One option is to vary if initial prosocial behavior is observed. We imagine that being observed during good behavior can accelerate the change in self-image by making the behavior more salient. It could also be the case that not only self-image but also social image plays a role when determining prosocial behavior.

Overall, we believe that more research will be needed in order to understand how people develop prosocial preferences and which environments are suited to foster them.

# References


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review* 101 (2): 470–492.

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *American Economic Review* 72 (3): 307–319.

Aktion Deutschland Hilft. 2016. *https://www.aktion-deutschland-hilft.de/de/fachthemen/armut/ (accessed 1 April 2016)*.

Algan, Yann, Elizabeth Beasley, Frank Vitaro, and Richard E. Tremblay. 2014. *The Impact of Non-Cognitive Skills Training on Academic and Non-Academic Trajectories: From Childhood to Early Adulthood*. Mimeo.

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

Ashraf, Nava, and Oriana Bandiera. 2017. "Altruistic Capital." *American Economic Review: Papers & Proceedings* 107 (5): 70–75.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse. 2012. "The Relationship Between Economic Preferences and Psychological Personality Measures." *Annual Review of Economics* 4 (1): 453–478.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82 (6): 1063–1093.

— . 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14 (3): 817–826.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126:805–855.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review* 71:117–120.

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

Bradley-Geist, Jill C., Eden B. King, Jeanine Skorinko, Michelle R. Hebl, and Casey McKenna. 2010. "Moral Credentialing by Association: The Importance of Choice and Relationship Closeness." *Personality and Social Psychology Bulletin* 36 (11): 1564–1575.

Brosig-Koch, Jeannette, Thomas Riechmann, and Joachim Weimann. 2017. "The Dynamics of Behavior in Modified Dictator Games." *PLoS ONE* 12 (4): 1–18.

Burger, Jerry, and David F. Caldwell. 2003. "The Effects of Monetary Incentives and Labeling on the Foot-in-the-Door Effect: Evidence for a Self-Perception Process." *Basic and Applied Social Psychology* 25 (3): 235–241.

Carpenter, Jeffrey, and Erika Seki. 2011. "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay." *Economic Inquiry* 49 (2): 612–630.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory* 33 (1): 67–80.

DeJong, William. 1979. "An Examination of Self-Perception Mediation of the Foot-in-the-Door Effect." *Journal of Personality and Social Psychology* 37 (12): 2221–2239.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2): 268–298.

Dunn, Elizabeth W., Lara B. Aknin, and Michael I. Norton. 2008. "Spending Money on Others Promotes Happiness." *Science* 319 (5870): 1678–1688.

Effron, Daniel A., Jessica S. Cameron, and Benoît Monin. 2009. "Endorsing Obama Licenses Favoring Whites." *Journal of Experimental Social Psychology* 45 (3): 590–593.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2): 293–315.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

Freedman, Jonathan L., and Scott C. Fraser. 1966. "Compliance without Pressure: The Foot-in-the-Door Technique." *Journal of Personality and Social Psychology* 4 (2): 195–202.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2016. "Habit Formation in Voting: Evidence from Rainy Elections." *American Economic Journal: Applied Economics* 8 (4): 160–188.

Gerlitz, Jean-Yves, and Jürgen Schupp. 2005. "Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP." *Research Notes 4*: 1–44.

Gneezy, Ayelet, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton. 2012. "Paying to Be Nice: Consistency and Costly Prosocial Behavior." *Management Science* 58 (1): 179–187.

Gorassini, Donald R., and James M. Olson. 1995. "Does Self-Perception Change Explain the Foot-in-the-Door Effect?" *Journal of Personality and Social Psychology* 69 (1): 91–105.

Grossman, Zachary. 2014. "Strategic Ignorance and the Robustness of Social Preferences." *Management Science* 60 (11): 2659–2665.

Khan, Uzma, and Ravi Dhar. 2006. "Licensing Effect in Consumer Choice." *Journal of Marketing Research* 43 (2): 259–266.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2019. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy (forthcoming)*.

Malle, Bertram F. 2006. "The Actor-Observer Asymmetry in Attribution: A (Surprising) Meta-Analysis." *Psychological Bulletin* 132 (6): 895–919.

Mazar, Nina, and Chen-Bo Zhong. 2010. "Do Green Products Make Us Better People?" *Psychological Science* 21 (4): 494–498.

Merritt, Anna C., Daniel A. Effron, Steven Fein, Kenneth K. Savitsky, Daniel M. Tuller, and Benoît Monin. 2012. "The Strategic Pursuit of Moral Credentials." *Journal of Experimental Social Psychology* 48 (3): 774–777.

Merritt, Anna C., Daniel A. Effron, and Benoît Monin. 2010. "Moral Self-Licensing: When Being Good Frees Us to Be Bad." *Social and Personality Psychology Compass* 4 (5): 344–357.

Monin, Benoît, and Dale T. Miller. 2001. "Moral Credentials and the Expression of Prejudice." *Journal of Personality and Social Psychology* 81 (1): 33–43.

Ouellette, Judith A., and Wendy Wood. 1998. "Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior." *Psychological Bulletin* 124 (1): 54–74.

Park, Soyoung Q., Thorsten Kahnt, Azade Dogan, Sabrina Strang, Ernst Fehr, and Philippe N. Tobler. 2017. "A Neural Link Between Generosity and Happiness." *Nature Communications* 8:15964.

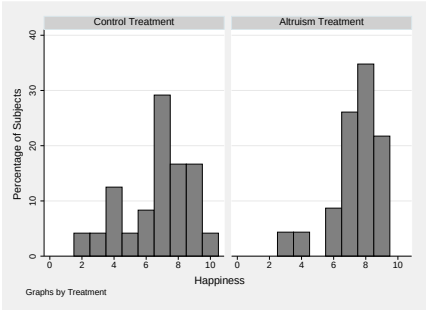
Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–1302.

Sass, Markus, and Joachim Weimann. 2012. *The Dynamics of Individual Preferences in Repeated Public Good Experiments*. Working Paper 2/2012. Magdeburg: Universität Magedeburg.

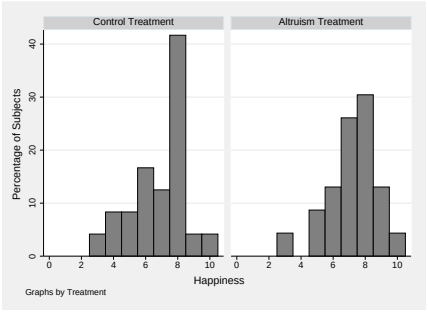
Schmitz, Jan. 2019. "Temporal Dynamics of Pro-Social Behavior: An Experimental Analysis." *Experimental Economics* 22 (1): 1–23.

SOS Kinderdörfer. 2016. *https://www.sos-kinderdoerfer.de/unsere-arbeit/hilfsprojekte (accessed 1 April 2016)*.

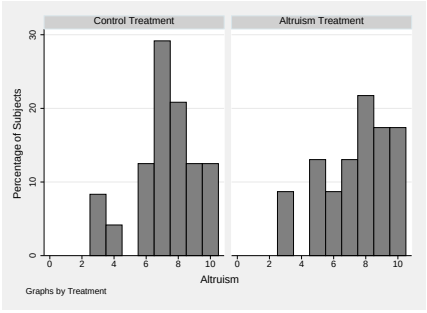
# 3.A Summary of Variables



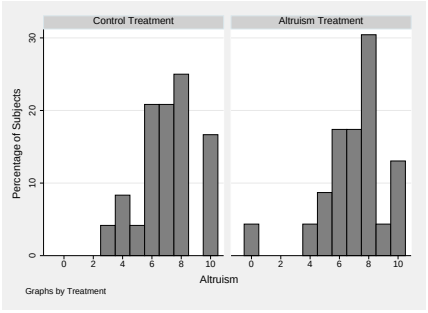

(a) Happiness Session 2



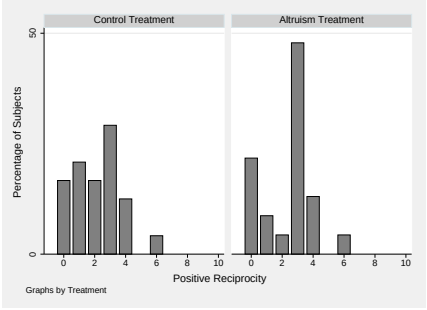

(b) Happiness Session 10



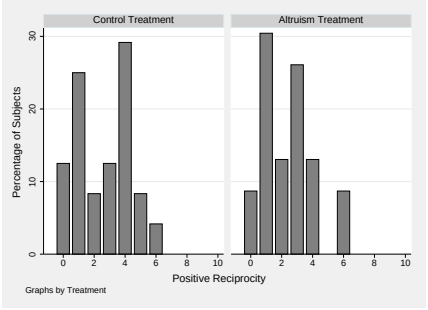

(c) Altruism Session 2



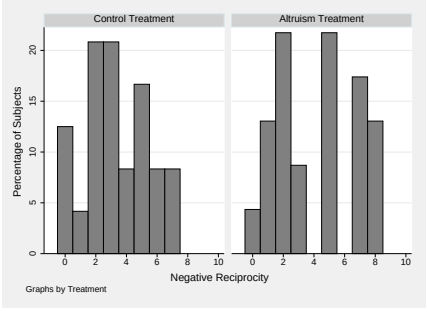

(d) Altruism Session 10



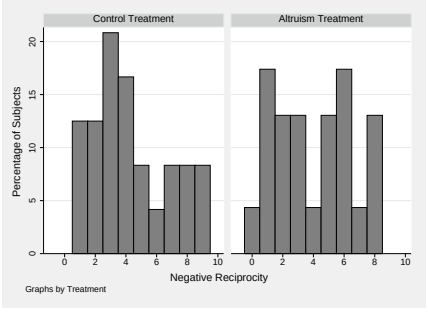

(e) Positive Reciprocity Session 2



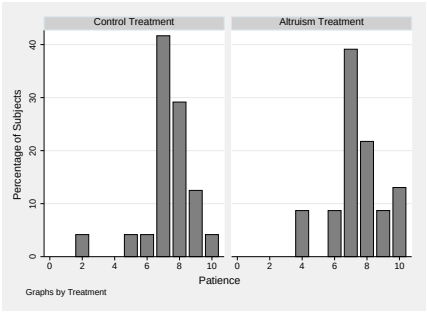

(f) Positive Reciprocity Session 10



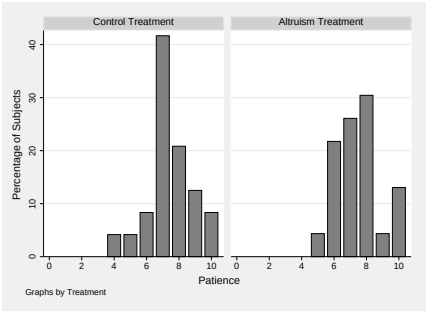

(i) Negative Reciprocity Session 2



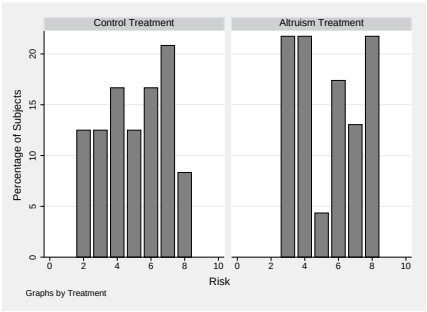

(j) Negative Reciprocity Session 10



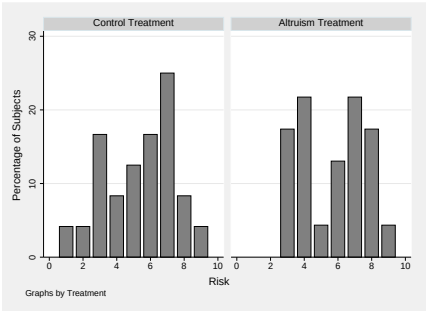

(a) Patience Session 2



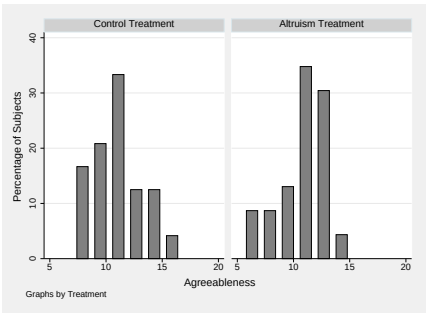

(b) Patience Session 10



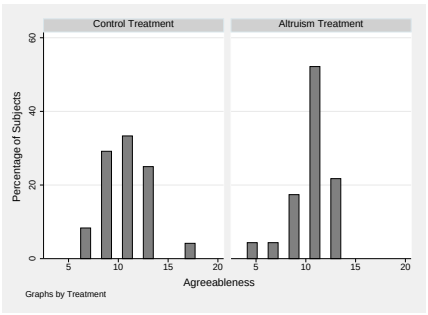

(c) Risk Session 2



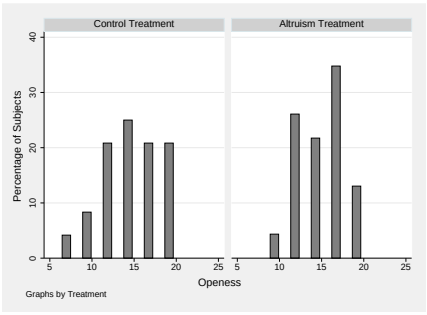

(d) Risk Session 10



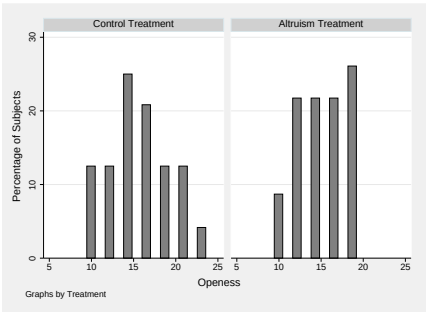

(e) Agreeableness (Big Five) Session 2



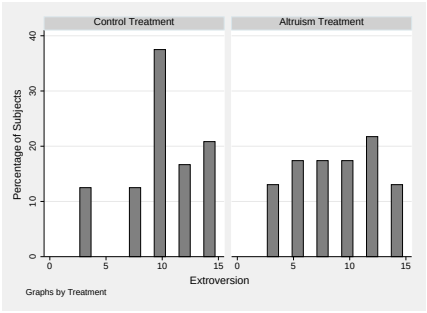

(f) Agreeableness (Big Five) Session 10



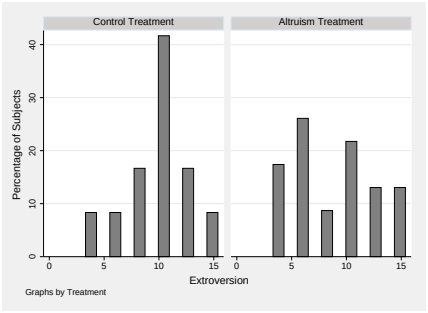

(q) Openess (Big Five) Session 2



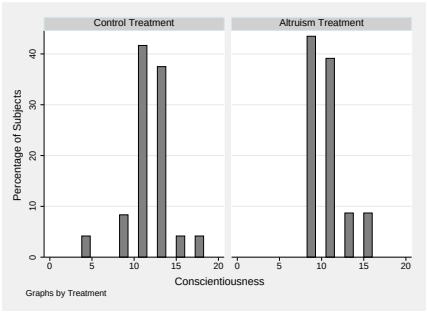

(r) Openess (Big Five) Session 10



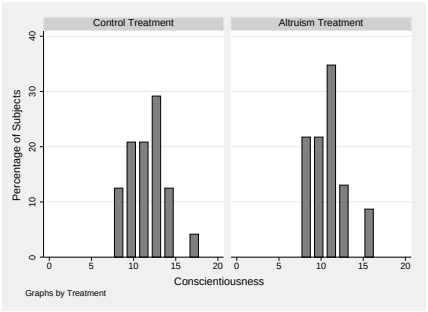

(a) Extroversion (Big Five) Session 2



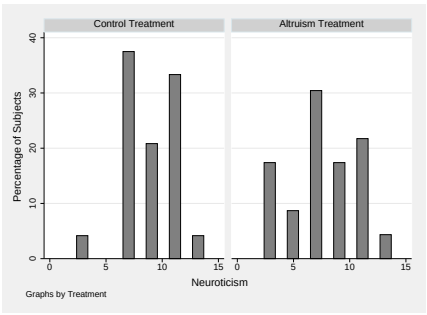

(b) Extroversion (Big Five) Session 10



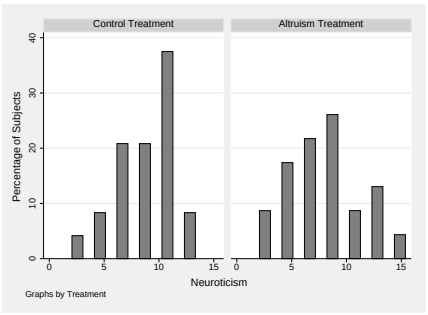

(c) Conscientiousness (Big Five) S. 2



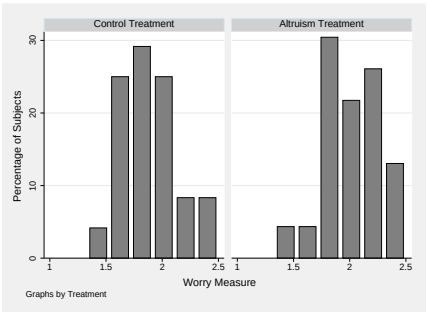

(d) Conscientiousness (Big Five) S. 10



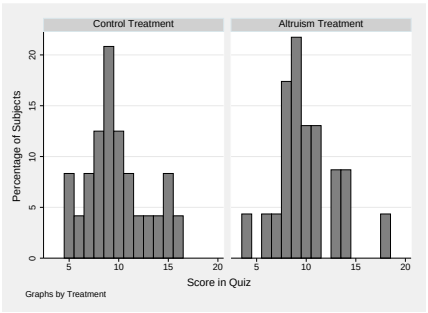

(e) Neuroticism (Big Five) Session 2




(f) Neuroticism (Big Five) Session 10




(y) Worry




(z) Correctly Answered Questions

Figure 3.7: Histograms

## 3.B Instructions

Instructions have been translated from German. Project descriptions were all based on descriptions of SOS Kinderdörfer's (2016) official homepage. Information on poverty was based on information provided by Aktion Deutschland Hilft (2016).

**Session 1 (31 May 2016)**

Thank you for participating in our study and welcome to the first part of the experiment!

Before we start with today's experiment, we have some information for you. As already mentioned in the invitation to this study, we are going to send you a link via email every Monday and Thursday during the following weeks. The addresses from which these emails will be sent are noreply@qemailserver.com and bonneconlab-online@uni-bonn.de.

You can then participate in the experiments following the corresponding links. You have time until 23:59 of the respective day to participate in each part of the experiment. Afterwards, the link sent by us loses validity. You should therefore be careful to take part in the experiments until 23:59 of the respective day. **For participating in all parts of the experiment, you will receive €100**.

The last part of the experiment, as mentioned in the invitation email, will take part during the first week of July. In a few weeks, you will be able to choose the date. This last part of the experiment will take 30 minutes at most.

Due to the data protection regulations of the BonnEconLab, we have to ask for your email address again. We will use this email address for the remainder of the study. **Please be sure to enter the same email address that you used for registering with the BonnEconLab. Participation is only possible with this specific email address.**

Please enter your email address in the following field:

Email address:

Please repeat your email address:

Email address:

Coming to the end of today's part, we are asking you for some information on your person:

Date of birth (DD/MM/YYYY):

Field of Study:

Number of Semesters:

Gender: male/female

Please click “Continue” to finish this part.

**Session 2 (2 June 2016)**

Welcome to the second part of the experiment!

In this part of the experiment it will be your task to count the number of zeros in a table und answer a questionnaire at the end.

You will now see a table on each following screen that contains zeros and ones. You have to count how many zeros are in the table. You can enter your answer directly on the same screen. As soon as you confirm your answer, we will tell you if it was correct or not, and a new table will be generated.

Please try to answer the tasks as quickly as possible.

Please click “Continue” to begin with the tasks.

Count the zeros in the following table and state the exact number in the field below:

How many zeros are in the table?

*[See Screenshot 3.11 in Appendix 3.C]*

Your answer was not correct.

Please click “Continue” to continue with the tasks.

*[or:]*

Your answer was correct.

Please click “Continue” to continue with the tasks.

In the following, we are asking you to answer some questions.

Here are different characteristics a person can have. Probably, some of the characteristics will apply to you personally whereas others do not. For some, you may be undecided.

Please use the following scale to answer: value 1 means “does not apply at all” and value

7 means “applies very much”. With the values between 1 and 7, you can grade your evaluation.

I see myself as someone, who...

- is a reliable worker ..... 1 2 3 4 5 6 7
- is talkative ..... 1 2 3 4 5 6 7
- is sometimes rude to others ..... 1 2 3 4 5 6 7
- is original, comes up with new ideas ..... 1 2 3 4 5 6 7
- worries a lot ..... 1 2 3 4 5 6 7
- has a forgiving nature ..... 1 2 3 4 5 6 7
- tends to be lazy ..... 1 2 3 4 5 6 7
- is outgoing, sociable ..... 1 2 3 4 5 6 7
- values artistic, aesthetic experiences ..... 1 2 3 4 5 6 7
- gets nervous easily ..... 1 2 3 4 5 6 7
- does things efficiently ..... 1 2 3 4 5 6 7
- is reserved ..... 1 2 3 4 5 6 7
- is considerate and kind to almost everyone ..... 1 2 3 4 5 6 7
- has an active imagination ..... 1 2 3 4 5 6 7
- is relaxed, handles stress well ..... 1 2 3 4 5 6 7

In the following, we are interested in how much you are willing to take on risks. Please state your evaluation on a scale from 0 to 10. 0 means “not at all willing to take on risks” and 10 means “very willing to take on risks”. You can grade your evaluation with the values in between.

0 1 2 3 4 5 6 7 8 9 10

We are now asking you for your willingness to behave a certain way. Please state your evaluation on a scale from 0 to 10. 0 means “not at all willing to do this” and 10 means “very willing to do this”. You can grade your evaluation with the values in between.

How much are you willing to forego something that carries utility for you in order to benefit from it in the future?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to punish someone who has treated you unfairly even though this has negative consequences for you?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to punish someone who has treated somebody else unfairly even though this has negative consequences for you?

0 1 2 3 4 5 6 7 8 9 10

To what extent would you be willing to donate to a good cause without expecting something in return?

0 1 2 3 4 5 6 7 8 9 10

Please think about how you would act in the following situation. You are in an unknown area und notice that you got lost. You are asking a stranger for the way. The stranger offers to accompany you to your destination.

Helping you costs the stranger approximately €20. However, the stranger says that he does not want money from you. You have six gifts with you. The cheapest gift costs €5, the most expensive gift costs €30. Would you offer the stranger one of the gifts as a thank you?

Yes/No

Which gift would you give to the stranger?

The gift worth €5
The gift worth €10
The gift worth €15
The gift worth €20
The gift worth €25
The gift worth €30
I don't know

Imagine the following situation: today, you received an unexpected €1.000.

How much of the money would you donate to a good cause? Donation:

At the end, we want to ask you how satisfied you are with your life.

Please state your evaluation on a scale from 0 to 10. 0 means "completely satisfied" and 10 means "completely unsatisfied".

How satisfied are you currently with your life?

0 1 2 3 4 5 6 7 8 9 10

**Session 3 (6 June 2016)**

**Control Treatment**

Welcome to the third part of the experiment

In today's part of the experiment, it is your task to solve several riddles.

If you solve a certain amount of riddles correctly, we will send you an Amazon voucher with a value of €5,00 via email tomorrow.

Please click "Continue" to start working on the riddles.

**Altruistic Treatment**

Welcome to the third part of the experiment.

A repeating topic of the experiment will be help and support for suffering children around the world. Children around the world are suffering from consequences of natural catastrophes, poverty, and wars. According to Aktion Deutschland Hilft, in 2015, 570 million of children lived in extreme poverty, around 150 million are orphans. Despite considerable efforts, the situation of many children in the world is still catastrophic in many ways.

In today's part of the experiment, it is your task to solve several riddles.

If you solve a certain amount of riddles correctly, the experimenters will donate an amount of €5,00 to a project of SOS Kinderdörfer tomorrow. As soon as the donation has been made, we will send you a confirmation email and you can check the respective bank statement after the laboratory session in July. Should you not solve enough riddles correctly the donation will not be made.

Purpose of the donation: today's donation goes to the project **Fight against Hunger** of SOS Kinderdörfer. Wordwide, almost one billion of people do not have enough food to eat. Especially children and families suffer from this; they are therefore supported by SOS Kinderdörfer. They give children in need a perspective and enable parents step by step to self-help: with food, seeds, and education. Via SOS-Familienhilfe, they fight against worldwide poverty and hunger, e.g. in Bangladesh, Niger, or Nicaragua.

Please click "Continue" to start working on the riddles.

**Both Treatments**

Your score:

Which of Figures 1 to 8 fits in the framed field?

*[See Figure 3.8 in Appendix 3.C]*

Your answer was correct.

Please click “Continue” to continue with the tasks.

*[or:]*

Your answer was not correct.

Please click “Continue” to continue with the tasks.

**Control Treatment**

Congratulations, you solved two riddles correctly.

You will receive your Amazon voucher with a value of €5,00 via email tomorrow.

Please click “Continue”, so that the experimenters can purchase the voucher.

*[or:]*

You answered all available riddles, but unfortunately you solved less than two of them correctly.

Please click “Continue” to end the survey.

**Altruistic Treatment**

Congratulations, you solved two riddles correctly.

A donation of €5,00 will be made to the project **Fight against Hunger** of SOS Kinderdörfer.

Please click “Continue”, so that the experimenters can make the donation generated by you.

*[or:]*

You answered all available riddles, but unfortunately you solved less than two of them correctly.

Please click “Continue” to end the survey.

**Session 4 (9 June 2016)**

**Control Treatment**

Welcome to the fourth part of the experiment.

*[See Session 3]*

**Altruistic Treatment**

Welcome to the fourth part of the experiment.

*[See Session 3]*

Purpose of the donation: today's donation will go to the project **Support of AIDS-Orpans** of SOS Kinderdörfer. After having to witness their parents dying, for many children the battle to survive begins. Many Aids-orphans end up on the streets. Adolescents, many of them still half children, are trying to take care of younger siblings. Others grow up with strained grandparents. They do not have enough to eat, not to mention money for school. Especially in Africa, HIV/Aids destroys the future of millions of children. SOS Kinderdörfer supports affected families via education, medication, money for school, or new homes in one of their villages.

Please click "Continue" to start working on the riddles.

**Both Treatments:**

Today's task will be to decode words that have been coded into ciphers back into letters.

For this, you receive the following key of decoding (you will see the key on every screen):

a = 17 n = 19 b = 8 o = 3 c = 26 p = 24 d = 1 q = 21 e = 11 r = 6 f = 4 s = 7 g = 18 t = 22 h = 2 u = 5 i = 10 v = 20 j = 13 w = 25 k = 14 x = 9 l = 16 y = 15 m = 23 z = 12

You will be asked to decode a sequence of ciphers, which each stands for a word, back into letters.

For example, you see the following sequence of ciphers:

17 24 4 11 16

Given the above key of decoding, the solution is „apfel", which you have to enter in a field. Please notice that neither mutated vowels, space characters, nor punctuation marks will appear in the words. Upper and lower cases are not relevant. Please click "Continue" to start with the tasks.

Your score:

Key of decoding:

Please enter the translation of the following sequence of ciphers in the field below:

*[See Figure 3.9 in Appendix 3.C]*

Your answer was correct.

Please click “Continue” to continue with the tasks.

*[or:]*

Your answer was not correct.

Please click “Continue” to continue with the tasks.

**Control Treatment**

Congratulations, you solved three riddles correctly.

You will receive your Amazon voucher with a value of €5,00 via email tomorrow.

Please click “Continue”, so that the experimenters can purchase the voucher.

*[or:]*

You answered all available riddles, but unfortunately you solved less than three of them correctly.

Please click “Continue” to end the survey.

**Altruistic Treatment**

Congratulations, you solved three riddles correctly.

A donation of €5,00 will be made to the project **Support of AIDS-Orphans** of SOS Kinderdörfer.

Please click “Continue”, so that the experimenters can make the donation generated by you.

*[or:]*

You answered all available riddles, but unfortunately you solved less than three of them correctly.

Please click “Continue” to end the survey.

**Session 5 (13 June 2016)**

**Control Treatment**

Welcome to the fifth part of the experiment.

*[See Session 3]*

**Altruistic Treatment**

Welcome to the fifth part of the experiment.

*[See Session 3]*

Purpose of the donation: today's donation goes to the project **Support of Girls in India** of SOS Kinderdörfer. They are abandoned or killed as babies, others are starved as small children – because their families do not want daughters. In India, girls are reckoned to be a burden and a curse, because daughters have to be endowed with a dowry, which many cannot afford. SOS Kinderdörfer protects girls in danger and advocates their rights. In India's 41 villages, girls that were abandoned, cast out, or left alone, find a new home, grow up with equal rights, and receive an education.

Please click "Continue" to start working on the riddles.

**Both Treatments**

It will be your task to solve so-called Sudoku riddles. For this, you will see a 9×9-matrix as the following on each screen:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 8 | 6 | | | 4 | | | | 9 |
| | | | | 5 | 8 | 6 | | |
| | 5 | | 2 | | 6 | 1 | | |
| | | 1 | | | | 3 | | 5 |
| | | 4 | 3 | | 2 | 9 | | |
| 3 | | 9 | | | | 4 | | |
| | | 6 | 4 | | 5 | | 7 | |
| | | 5 | 6 | 2 | | | | |
| 2 | | | | 3 | | | 6 | 1 |

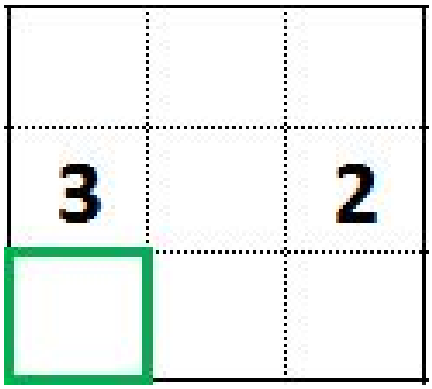
Your task is to find out which cipher belongs in the green field. To solve a Sudoku, the following rule applies:

Each row, each column, and each 3×3-box must contain the ciphers 1 to 9 exactly once.

The cipher we are looking for is definitely defined by the already existing black ciphers and can be deduced from them directly.

In our example, the cipher 5 is missing. This can be deduced as follows from the black ciphers.

Let's look at the 3×3-box in the middle of the riddle, which contains the green box:

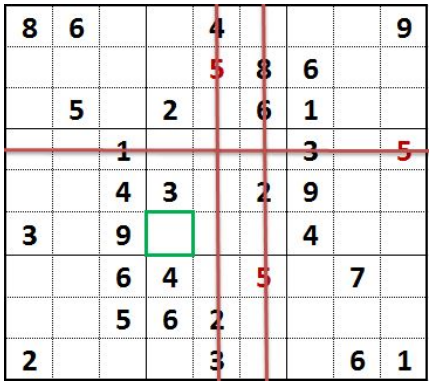



Here, the ciphers 3 and 2 have already been entered. From these, it can be deduced that the cipher 5 must be written into one of the remaining seven boxes, since every 3×3-box has to contain the ciphers 1 to 9 exactly once. If we now look at the whole riddle again, we can see that cipher 5 cannot stand in the upper row of the 3×3-box since this row already contains a cipher 5 and one cipher must not appear twice in one row.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 8 | 6 | | | 4 | | | | 9 |
| | | | | 5 | 8 | 6 | | |
| | 5 | | 2 | | 6 | 1 | | |
| | | 1 | | | | 3 | | 5 |
| | | 4 | 3 | | 2 | 9 | | |
| 3 | | 9 | | | | 4 | | |
| | | 6 | 4 | | 5 | | 7 | |
| | | 5 | 6 | 2 | | | | |
| 2 | | | | 3 | | | 6 | 1 |

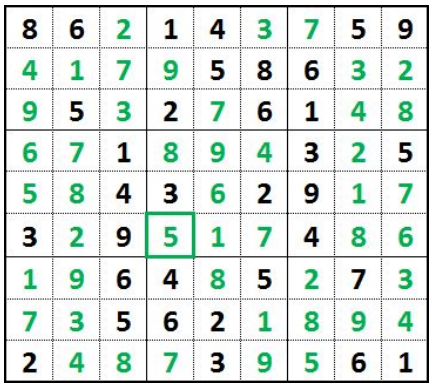
Cipher 5 can neither be written in the middle column and the right column of the central 3×3-box, since 5 already appears once in each of the two columns:

Now, the only spot left available for cipher 5 in the 3×3-box is the green box, therefore 5




is the solution.

The rest of the Sudoku can be solved using similar logic. The completely solved riddle (just as an example, you will not be asked to solve a whole riddle yourself) looks like the following:




Please click “Continue” to start with the tasks.

*[See Figure 3.10 in Appendix 3.C]*

Your answer was not correct.

Please click “Continue” to continue with the tasks.

*[or:]*

Your answer was correct.

Please click “Continue” to continue with the tasks.

**Control Treatment**

Congratulations, you solved two riddles correctly.

You will receive your Amazon voucher with a value of €5,00 via email tomorrow.

Please click "Continue", so that the experimenters can purchase the voucher.

*[or:]*

You answered all available riddles, but unfortunately you solved less than two of them correctly.

Please click "Continue" to end the survey.

**Altruistic Treatment**

Congratulations, you solved two riddles correctly.

A donation of €5,00 will be made to the project **Support of Girls in India** of SOS Kinderdörfer.

Please click "Continue", so that the experimenters can make the donation generated by you.

*[or:]*

You answered all available riddles, but unfortunately you solved less than two of them correctly.

Please click "Continue" to end the survey.

**Session 6 (16 June 2016)**

*[See Session 3, except the project description:]*

Purpose of the donation: today's donation goes to the project **Building a Children's Village in South Sudan** of SOS Kinderdörfer. The terrible civil war in South Sudan has almost destroyed everything that made up the lives of the habitants of the children's village in Malakal. At the moment, they are living in a temporary village in Juba made of mud huts. SOS Kinderdörfer wants to give them a perspective for the future. They want to build a permanent and save home for the 89 children and 23 adolescents and their SOS-mothers in the next two to three years.

**Session 7 (20 June 2016)**

*[See Session 4, except the project description:]*

Purpose of the donation: today's donation goes to the project **Help after Earthquake in Nepal** of SOS Kinderdörfer. The devastating earthquake of 25 April 2015 has brought suffering and destruction to the people of the Himalaya region. SOS Kinderdörfer supports children and families in Nepal: via first lifesaving emergency aid right after the catastrophe, now via helping and supporting families and students by building new schools and family homes.

**Session 8 (23 June 2016)**

*[See Session 5, except the project description:]*

Purpose of the donation: today's donation goes to a **Clinic of SOS Kinderdörfer in Mogadishu, Somalia**. Hunger, a civil war and missing infrastructure are threatening the lives of children in Somalia. The mortal rate of boys and girls under age five is high. Every tenth child in Somalia dies before its first birthday. In the country, which basically does not have a public health care system, the SOS-mother-child-clinic in the capital is an important place to go. In the clinic, amongst others, malnutrition and Malaria are treated, vaccinations are done, and support for births is offered.

**Session 9 (27 June 2016)**

Welcome to the ninth part of this experiment!

The purpose of this part will be help and support for suffering children around the world. Children around the world are suffering from consequences of natural catastrophes, poverty, and wars. According to Aktion Deutschland Hilft, in 2015, 570 million of children lived in extreme poverty; around 150 million are orphans. Despite considerable efforts, the situation of many children in the world is still catastrophic in many ways.

In this part of the experiment, your task will be to count zeros in tables.

You will now see a table on each following screen that contains zeros and ones. You have to count how many zeros are in the table. You can enter your answer directly on the same screen. As soon as you confirm your answer, we will tell you if it was correct or not, and a new table will be generated.

For every task you solve correctly, a donation of €0.20 will be made to a project of SOS Kinderdörfer by the experimenters tomorrow. This means that the more tasks you solve, the higher the total donation.

Notice: for every task for which you give a wrong answer, €0.05 will be subtracted from the donation. You should therefore not guess, but try to solve all tasks carefully.

As soon as the donation has been made, we will send you a confirmation email and you can check the respective bank statement after the laboratory session in July.

Purpose of the donation: today's donation goes to the project **Support of Ebola-Orphans** of SOS Kinderdörfer. Thousands of people have died from the Ebola virus. Staying behind are traumatized children who have lost their parents and are treated as outcasts. The affected children are in need of protection and help. Ebola-orphans are stigmatized. Out of fear of contagion, many of them are not taken in by their relatives and end up on the streets. Especially the weakest do not have anyone standing by their side. SOS Kinderdörfer takes care of them now and in the coming years.

You yourself decide how long you want to work on the tasks. You can end the task whenever you want. While you are working on the task, you will always see an empty field. If you want to end the task, please enter the word "stop" into the field and the donation that you have generated up to that point will be made by the experimenters.

Please click "Continue" to start with the task.

Your donation: €

If you want to end working on the tasks, please enter the word "stop" into the following field:

Count the zeros in the following table and enter the exact number into the field below:

*[See Figure 3.11 in Appendix 3.C]*

Your answer was correct.

Please click "Continue" to continue with the task.

*[or:]*

Your answer was not correct.

Please click "Continue" to continue with the task.

*[or:]*

You have stopped working on the task.

Your donation is €.

Thank you for pariticpating!

Please click “Continue”, so that the experimenters can donate the amount generated by you.

*[or:]*

You solved all tasks.

Your donation is €.

Thank you for participating!

Please click “Continue”, so that the experimenters can donate the amount generated by you.

**Session 10 (30 June 2016)**

Welcome to the tenth part of this experiment!

In this part of the experiment, we are asking you to fill out a questionnaire.

Please click “Continue” to start answering the questionnaire.

*[See questionnaire Session 2]*

**Session 11 (4, 5, or 7 July 2016.)**

Welcome to the eleventh and final part of this experiment!

Before we start the study, we want to know how satisfied you are with your life.

Please state your evaluation on a scale from 0 to 10. 0 means “completely satisfied” and 10 means “completely unsatisfied”.

How satisfied are you currently with your life?

0 1 2 3 4 5 6 7 8 9 10

Please click “Continue” to start the study.

In today’s part of the experiment it will be your task to take part in a short quiz and answer questions.

For every correctly answered question you receive an additional payoff of €0.20.

Please click “Continue” to start answering the questions.

*[See Figure 3.12 in Appendix 3.C]*

Thank you, you answered all questions.

You are receiving an additional payment of €.

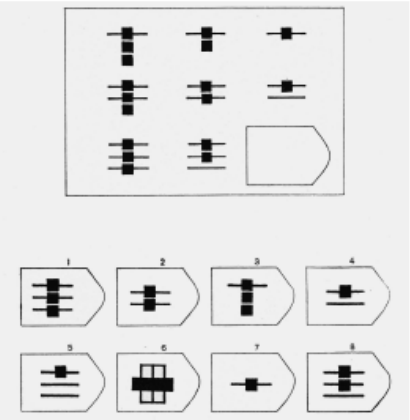
Please click “Continue” to end the experiment and go to the adjoining room to receive your payment.

## 3.C Screenshots

Figure 3.8: Screenshot of real-effort task Raven Matrices of Sessions 3 and 6

Ihr bisheriger Score: 0

Welche der Abbildungen 1 bis 8 passt in das umrandete Feld?



1 2 3 4 5 6 7 8

weiter

Figure 3.9: Screenshot of real-effort task Translation of Sessions 4 and 7

Ihr bisheriger Score: 0

Kodierungsschlüssel:

| | |
|---|---|
| a = 17 | n = 19 |
| b = 8 | o = 3 |
| c = 26 | p = 24 |
| d = 1 | q = 21 |
| e = 11 | r = 6 |
| f = 4 | s = 7 |
| g = 18 | t = 22 |
| h = 2 | u = 5 |
| i = 10 | v = 20 |
| j = 13 | w = 25 |
| k = 14 | x = 9 |
| l = 16 | y = 15 |
| m = 23 | z = 12 |

Geben Sie bitte die Übersetzung der folgenden Zahlenkombination in das darunterliegende Feld ein:

**17 19 1 11 19 14 11 19**

weiter

Figure 3.10: Screenshot of real-effort task Sudoku for Sessions 5 and 8

Figure 3.11: Screenshot of real-effort task Counting Zeros for Session 9

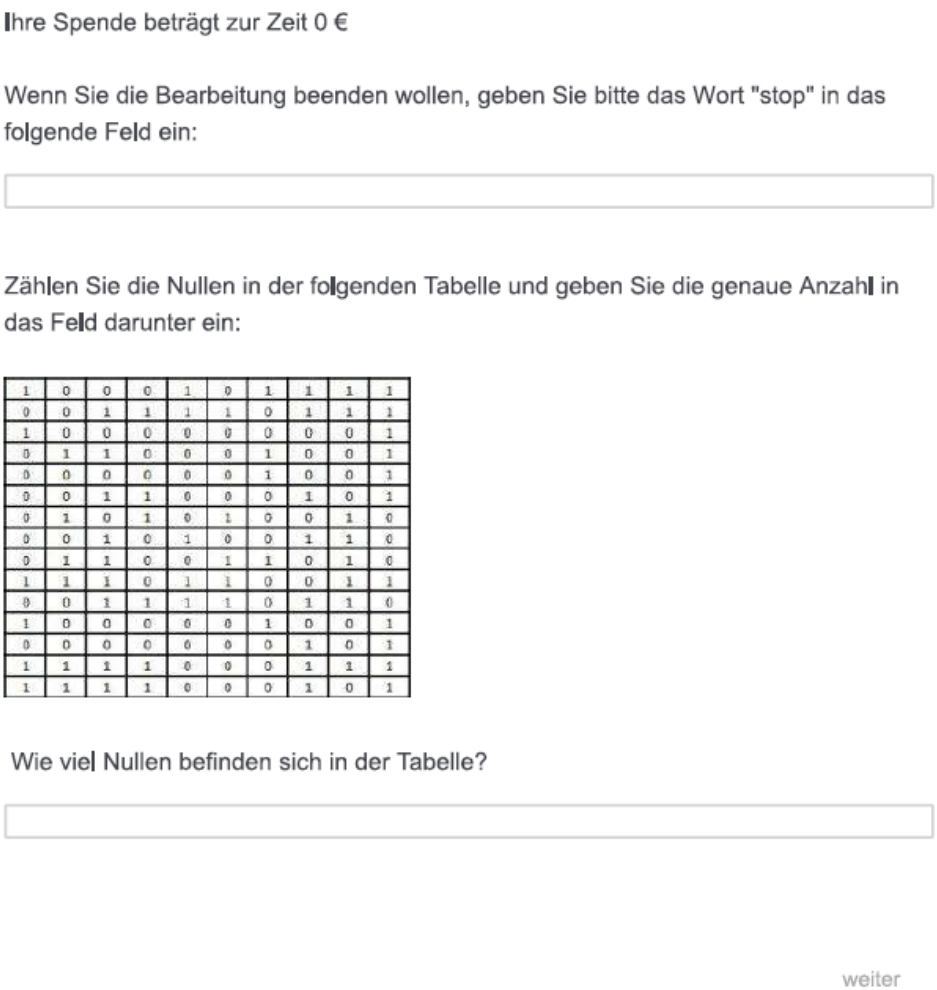
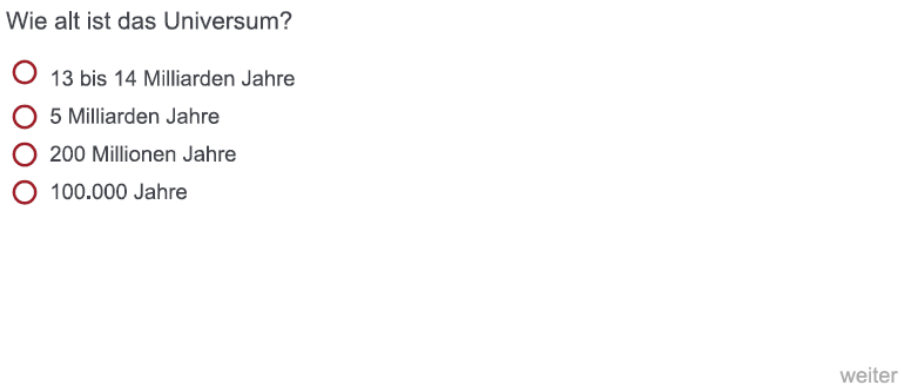


Figure 3.12: Screenshot of quiz question of Session 11



# Chapter 4

# Image Concerns and the Dynamics of Prosocial Behavior[1]

## 4.1 Introduction

Prosocial behavior is a pervasive facet of human interactions. Humans volunteer, give money to charities, donate blood, and help friends as well as strangers. All of these activities evoke personal costs but people are nonetheless willing to make sacrifices to increase social welfare (Charness and Rabin, 2002). Such behavior is often understood to reflect social preferences.[2] Ample evidence suggests that social preferences positively affect economic success (Carpenter and Seki, 2011; Becker et al., 2012; Algan et al., 2014) and well-being (Dunn et al., 2008; Park et al., 2017) in several contexts.[3]

Policy makers and corporations may hence wish to foster the prevalence of social preferences to obtain its benefits. However, the current state of knowledge on the malleability and the development of social preferences allow little guidance, as our understanding of the matter is still quite limited.

We experimentally investigate how prosocial behavior, one expression of social preferences, can be fostered over time. One particular variable that can affect prosocial behavior is observability. It has repeatedly been shown that people behave differently when others witness their actions (Zajonc, 1965; Guerin, 1983). In particular, being observed usually increases prosocial behavior because people want to be liked and respected by others (Ariely et al., 2009) or want to avoid resentment (DellaVigna et al., 2012). These

[1] This chapter is joint work with Louis Strang.

[2] Important manifestations of social preferences are, for instance, altruism (Becker, 1974, 1976), inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), and warm glow (Andreoni, 1989).

[3] Note that the correlation between social preferences and economic success could also be explained by a respective correlation of both variables with IQ (Burks et al., 2009).

studies report, however, only the change of behavior during the observation itself. Beyond that, little is known about the sustainability of these positive observability effects and it is unclear how being observed affects the dynamics of prosocial behavior. We contribute to the existing research by investigating spillover effects of being observed during the decision over a prosocial act on subsequent prosocial behavior. We hypothesize that observability not only increases immediate prosocial behavior but has positive spillover effects on later behavior as well.

This hypothesis is motivated by a new—and at the same time ancient—approach to conceptualize the formation of altruistic attitudes. According to Aristotle's Nicomachean Ethics, virtues are formed through the practice of virtuous actions. In modern terminology, engaging in prosocial behavior becomes a habit and eventually changes the person's self-image, meaning the way they think about themselves. They then keep up the prosocial behavior in order to avoid cognitive dissonance (Akerlof and Dickens, 1982). This idea is captured by the concept of *altruistic capital* that states that past altruistic behavior accumulates altruistic capital that enables individuals to internalize how actions affect others and finally increases future altruistic behavior (Ashraf and Bandiera, 2017). Being observed while doing something good should therefore increase later prosocial behavior for two reasons. First, being observed should, due to image concerns, increase immediate prosocial behavior compared to a situation in which one is not observed. This builds up altruistic capital, and has therefore positive spillover effects on subsequent behavior. Second, we argue, that performing good deeds in front of others makes a given action more salient, intensifies the experience and has therefore stronger effects on a person's self-image adjustment. Furthermore, it should also change people's beliefs about their social image; the way they think that others think about them. These image changes lead to a stronger increase of altruistic capital. We capture these mechanisms in a theoretical framework and derive our hypotheses formally.

We conduct two variants of a laboratory experiment to test if observability of earlier prosocial actions influences later levels of prosocial behavior. The experiments differ in the currency of giving in the later period (either money or effort) and in the mode of observability (either one single observer or a multi-people audience). In Experiment A, we find that prosocial behavior, as expected and in accordance with prior research, increases when subjects are observed. We do not find such a difference in Experiment B. Moreover, we find only small and insignificant positive effects of early observability on subsequent prosocial behavior in both experiments.

We proceed as follows: Section 4.2 reviews the relevant literature, Section 4.3 describes the two experimental designs, Section 4.4 presents a theoretical model and derives predictions, Section 4.5 presents the results, and Section 4.6 discusses the results

and concludes.

## 4.2 Literature

In economics, social preferences are traditionally understood to be persistent traits of individuals—complementing other dimensions of their enduring personality (Becker et al., 2012). They also have been found to be partially transmitted from generation to generation (Dohmen et al., 2012; Nunn and Wantchekon, 2011). However, there likewise exists evidence that social preferences can be altered, for instance when interacting and receiving attention from a socially-minded mentor during childhood (Kosse et al., 2019). Moreover, altruistic behavior is highly context-dependent (Dana et al., 2007; Grossman, 2014): certain features may trigger people to behave less prosocially—for instance, when contexts provide individuals with cues that potentially serve as excuses for not behaving prosocially or when the responsibility for certain outcomes is diffused. At the same time, other contexts promote prosocial behavior.

People have been shown to have image concerns, meaning they behave differently when others are present and can observe their actions. This can be due to an opportunity to display a convenient and normatively desired behavior, which is or is not in line with own preferences. Regarding prosocial behavior, this implies that individuals tend to behave more prosocially when they are observed, allowing them to obtain social recognition for their actions (Alpizar et al., 2008; Ariely et al., 2009; Powell et al., 2012; Bašić et al., 2018).

We seek to contribute to these findings by testing whether positive context effects of image concerns on prosocial behavior spill over to subsequent behavior, that is, spur circles of prosociality. In a broader context, we want to find out how prosocial behavior can be increased sustainably by gradually changing social preferences.

Our project builds on theoretical and empirical literature on dynamics of prosocial and moral behavior. When deriving our theoretical model of altruistic capital, we follow Ashraf and Bandiera (2017) who argue that past altruistic behavior accumulates altruistic capital which increases future altruistic behavior. Bénabou and Tirole (2011) offer an underlying mechanism that could explain such an accumulation process. In their model, agents gain utility from high self-esteem and make inferences about their true unknown moral type by observing their own past moral or immoral actions. Moral behavior is interpreted as an investment in one's self-image. The model yields the conclusion that, under certain conditions, good actions can build up *moral capital* and lead to further good actions, whereas bad actions destroy moral capital and lead to further bad actions.

Empirical evidence on the development of altruistic behavior stems from psycholog-

ical and recent economic research. There is evidence on people compensate early moral or immoral behavior; it is observed that early prosocial actions lead to decreased prosociality later on, whereas early selfish actions lead to an increase in prosocial behavior (see Merritt et al., 2010, for a summary). Schmitz (2019) reports results from an experiment on repeated social behavior in which subjects play a donation dictator game at two points in time. The second donation is smaller and this decrease is even stronger if both decisions happen within a day instead of having an extended period of one week between the two decisions.

However, there also exists evidence on the *Foot-in-the-Door Effect* that refers to the phenomenon that the acceptance of a small initial request leads to a more probable acceptance of a larger request that is made afterwards (Freedman and Fraser, 1966; DeJong, 1979; Beaman et al., 1983). It is argued that this effect shows due to a change in self-perception of individuals who accept the first small request, which is therefore in line with our argument.

Gneezy et al. (2012) experimentally investigate another dimension that is important for subsequent altruistic behavior. They claim that the development of a prosocial self-perception is only possible if prosocial acts involve personal costs. They find that people increase prosocial behavior only when the initial prosocial behavior was costly. Costless actions, in contrast, have no effect on subsequent prosocial decisions or can even decrease them. Our design incorporates this finding since subjects always have to invest time and effort or money in order to behave altruistically.

Building on these previous works on moral dynamics, social recognition, and the malleability of social preferences, we test not only the immediate effects of observability on prosocial behavior but in particular how later prosocial behavior is affected. We conjecture that social attention directed at one's good deeds leads to an adjustment of social-image and stronger adjustments of self-image. We therefore expect subjects to increase their later prosocial behavior if they have been observed beforehand.

## 4.3 Experimental Design

We investigate the causal effect of observability on present and future prosocial behavior by conducting two variants of a laboratory experiment. In both experiments, subjects face two sequential prosocial decisions within one session (see Figure 4.1 for an overview). We vary the observability of the subjects' first decision between treatments: in the Public-Private treatment, the first prosocial decision is observed by one observer or a group of observers, while the second prosocial decision is always made in private. In contrast, both decisions are made anonymously in the Private-Private treatment. We are

primarily interested in second-stage prosocial behavior to evaluate the spillover effects of being observed on subsequent non-observed prosocial behavior. We run two variants of the experiment, which differ in the way donations are made and how observability is implemented.

Both experiments were conducted at the BonnEconLab using oTree (Chen et al., 2016) and hroot (Bock et al., 2014). Experiment A was conducted in August and September 2017 and a total of 242 subjects participated (including 37 subjects who served as observers). Experiment B was conducted in December 2017 and 77 subjects participated. Section 4.A includes verbal and written instructions.

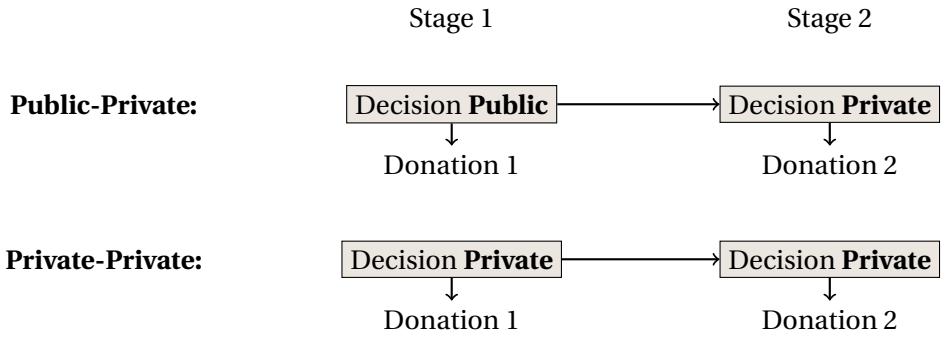



Figure 4.1: Treatment overview

## Experiment A

In Experiment A, subjects participate in one of two roles. A minority of the subjects functions as observers, that is they do not make any decisions themselves but solely monitor the behavior of other subjects. The remaining subjects, irrespective of the treatment, take the same two consecutive donation decisions.

Donation decisions are made as follows: in Stage 1, subjects can work on a real-effort task called *Counting Zeros* (first implemented by Abeler et al., 2011) to generate a donation to a project of the charity *SOS-Kinderdörfer*. In this task, subjects face 15×10 - tables, with all 150 cells each containing either the digits *0* or *1*. On each screen, containing exactly one table (see Figure 4.6 for a screenshot), subjects have to state the total number of zeros that the table contains. Per correctly counted table, the generated donation increases by a specific piece rate, which decreases in the number of completed tasks (see Table 4.1). To prevent subjects from simply guessing the correct number, we subtract €0.05 from the total donation for answers deviating by more than one from the correct

number.[4] Subjects can freely choose to stop working at any time and can leave earlier when doing so. This allows for higher opportunity costs of exercising and hence more costly prosocial acts.[5] There is a maximum time of 20 minutes and a maximum number of 25 tables, resulting in a maximum donation of €2.90.

Stage 2 consists of a double-blind dictator game. In this stage, subjects open an envelope, which they already receive at the beginning of the experiment. This envelope contains the subjects' compensation of €5 for participating in the experiment.[6] The envelope also contains written instructions and a smaller envelope. The instructions state that participants may leave any amount of the €5 in the small envelope to donate to a different project of the same charity as in Stage 1.[7]

We use a between-session treatment variation to prevent subjects from the Private-Private treatment being aware of any social component of the experiment. Sessions are conducted in turns; each one lasting at most 30 minutes. We now describe the exact procedure of each treatment.

**Private-Private** For each Private-Private session, we invite three participants to the BonnEconLab.[8] At the beginning, they receive the aforementioned envelope, verbal instructions to open it at the end of the experiment, and information about the size of their compensation. Afterwards, they are sent to three separate rooms with one working space and one computer each. They are told to choose their respective rooms themselves to ensure a double-blind procedure and complete anonymity. Instructions for Stage 1 are already displayed on the computer screens when subjects enter the room and they immediately start with the experiment. In Stage 1, subjects work on the Counting Zeros task described above to generate a donation between €0 and €2.90. After subjects decide to

[4] We allow for a margin of error of +/- 1. Also, the total amount cannot become negative.

[5] As mentioned before, Gneezy et al. (2012) emphasize the importance of positive costs.

[6] Observers in the Public-Private condition receive a flat payment of €5 as well.

[7] The €5 are provided in coins, such that all donations between €0 and €5 in steps of 10 Cents are possible.

[8] In case that less than three subjects show up for a Private-Private session, the session is run with fewer subjects.

Table 4.1: Experiment A: piece rates for correctly solving a table of Counting Zeros

| Tables solved | Piece rate |
|---|---|
| 1 – 5 | 30 Cents |
| 6 – 10 | 20 Cents |
| 11 – 15 | 5 Cents |
| 16 – 20 | 2 Cents |
| 21 – 25 | 1 Cent |

stop working, they have solved the maximum number of tables, or time is up, they are informed about their generated donation and open the envelope that leads to Stage 2, which was not announced beforehand. After deciding how much money to donate in the dictator game, subjects leave without talking to or seeing the experimenter or any of the other subjects again.

**Public-Private** For each Public-Private session, we invite one additional subject, resulting in a total of four subjects per session.[9] At the beginning of each session, all four subjects are seated at the same table and are asked to introduce themselves to each other by stating their first name and field of study.[10] Subsequently, one subject is randomly determined to act as an observer whose only role it is to monitor the performances of the remaining three subjects during Stage 1. After the observer is determined, he or she is separated from the other subjects and seated at a computer. On this computer, the other subjects' screens are displayed such that the observer can monitor their performances. Meanwhile, the other three subjects receive the same envelopes and the same information as subjects in the Private-Private treatment. Additionally, they are told that the observer will monitor their behavior and that each subject will have to report his or her outcomes to the observer in person. The observer is not aware of the envelopes to ensure the other subjects not feeling observed in Stage 2. From here on, the procedure of Stage 1 is identical to Public-Private. Only at the end of this stage, before moving on to Stage 2, when subjects are told about their donation, they are asked to go to the observer and report their generated donation.

Upon returning from the observer, they open the envelope that leads to Stage 2, which was not announced beforehand. The second stage proceeds in exactly the same way as in the Private-Private treatment, including complete anonymity. After deciding how much money to donate in the dictator game, subjects leave without talking to or seeing the experimenter, the observer or any of the other subjects again.

## Experiment B

In Experiment B, all subjects, irrespective of the treatment, take the same two consecutive donation decisions. For a tighter control of the dynamics of prosocial behavior, we change, compared to Experiment A, the nature of the donation decisions. Instead of using different donation decisions in Stage 1 and Stage 2, we now use the same real-

[9] In case that less than four but more than one subjects show up for a Public-Private session, we run the session with three or two subjects. If only one subject shows up, he or she is paid a show-up fee and is sent home.

[10] These personal interactions are used to create familiarity between subjects and have been used before. See, for instance, Gächter and Fehr (1999) and Ewers and Zimmermann (2015).

Table 4.2: Experiment B: piece rates for a correctly pressed key combination of Click for Charity

| Correct combinations | Piece rate |
|---|---|
| 1 – 100 | 1.00 Cent |
| 101 – 200 | 0.50 Cent |
| 201 – 350 | 0.25 Cent |
| 351 – 500 | 0.10 Cent |
| 501 – 700 | 0.05 Cent |
| > 700 | 0.01 Cent |

effort task in both stages. This allows detecting differences in prosocial behavior not only across treatments but also within-subjects between Stage 1 and Stage 2. We now can make a statement on the prevalence of the altruistic capital effect in observed and non-observed situations.[11] Moreover, we change the observational mechanism. Subjects have to report their donation in front of all other subjects of the same session rather than reporting to a single observer to further increase the salience of observability.[12]

We first describe the task used for making donation decisions. We closely follow the design of Ariely et al. (2009) using their real-effort task *Click for Charity* in both stages. The task consists of alternately pressing the keys "X" and "Y" on the computer keyboard[13] for five minutes. For each correct combination, a piece rate is donated to a project of the charity SOS-Kinderdörfer. Once again, the piece rate is concave and declines in the number of correct combinations (see Table 4.2). Figure 4.7 shows a screenshot of the task screen. Again, the projects differ between the two stages.

The experiment is conducted as follows: subjects arrive at the laboratory and are randomly assigned to one of the two treatments. When receiving the instructions, subjects in the Public-Private treatment additionally learn that they will have to announce their first name and their generated donation from Stage 1 at the end of the experiment in front of all other participants of the session. Subjects in the Private-Private treatment do not receive this information and are not aware of the other treatment condition.[14] After practicing the above described task Click for Charity, they can work on it for five minutes to generate their Stage 1 donation. Note that none of the subjects is aware of Stage 2 dur-

[11] Note that one could argue that repeating the task might lead to fatigue and hence to a decline in performance in the second stage, or, contrary, that learning might enhance performance. However, any time trend effects are orthogonal to our main treatment comparisons and therefore cannot explain any differences between observed and unobserved subjects.

[12] On average, 22 subjects participate in one session.

[13] Computer keyboards all have a German layout.

[14] Subjects in both treatments only learn about the other treatment at the end of the experiment after both stages are completed and subjects of the Public-Private treatment have to announce their donations.

ing this phase. Only after finishing Stage 1, subjects receive written instructions for Stage 2, which follows the same procedure as Stage 1. However, now all subjects are specifically informed that this stage's donation is completely anonymous.

Furthermore, we ask subjects for their level of happiness at the beginning and at the end (before the public announcement of donations) of the experiment.

Finally, participants receive a flat compensation of €6. Each session lasts at most 40 minutes.

## 4.4 Theory and Hypotheses

In this section, we derive a simple theoretical model and present the hypotheses that follow from it. According to Aristotle, people become virtuous by committing virtuous acts and thereby accustoming to it. We model this habitual formation with the assumption that people accumulate altruistic capital whenever doing something altruistic. When deriving our model of altruistic capital formation, we follow the approach of Ashraf and Bandiera (2017).

In period $t = 1, 2$, Agent $i$ chooses an altruistic action $a_{i,t} \in [0, \bar{a}]$. The altruistic action generates social welfare $W(a_{i,t})$, but creates a cost $c(a_{i,t}, A_{i,t})$ at the same time, where $W(a_{i,t})$ is increasing and concave in $a_{i,t}$ and $c(a_{i,t}, A_{i,t})$ increases linearly in $a_{i,t}$. The altruistic action $a_{i,t}$ does not only generate social welfare and create costs but also accumulates altruistic capital in the next period, denoted by $A_{i,t+1}$. Share $u$ of the altruistic action increases social welfare in the same period, whereas share $1-u$ increases altruistic capital of the following period (this borrows from Lucas (1988)). Apart from this, altruistic capital builds up faster, the higher the parameter $\kappa_t$, which reflects a particular form of self-awareness. It reflects our understanding that higher image concerns make altruistic acts more salient and therefore enhance the internal habit formation process. Image effects are common to all agents but are situation-specific, as they depend for instance on the presence of an audience. In our experiment, we vary the effect of image in the first period between treatments, assuming that $\kappa_t$ is increasing in public observability, that is $\kappa_t^{\text{Public}} > \kappa_t^{\text{Private}}$. In particular, altruistic capital in period $t$ is $A_{i,t} = (1-u)\kappa_{t-1} a_{i,t-1} + (1-\delta) A_{i,t-1}$, where $\delta$ captures the depreciation rate of altruistic capital.

We argue that greater altruistic capital reduces the cost of acting altruistically as one accommodates to altruistic behavior. Having a prosocial identity (due to self- and/or social image adjustments) makes behaving prosocially less costly since it reduces cognitive dissonance and because the decision process becomes less difficult. We therefore assume that altruistic capital decreases the cost of acting prosocially, that is, $\partial c / \partial a_{i,t} \partial A_{i,t} <$

0. [15]

Finally, agent $i$'s utility in period $t$ is equal to $(\sigma_i + \theta_t)W\big(a_{i,t}\big) - c(a_{i,t}, A_{i,t})$. The utility increases proportionally in $W$ for two reasons: first, the agent attaches a positive weight $\sigma_i$ on $W$ that represents her individual social preferences, such as pure altruism or warm glow. The second component, $\theta_t$, expresses a second kind of image effects, where an agent simply wants to make a better impression while being observed (social image). We exogenously vary the parameter in our experiment, and we expect that $\theta_t^{\text{Public}} > \theta_t^{\text{Private}}$. This image effect can be interpreted as the agent deriving utility from others thinking well of her (social image). The agent seeks to maximize her utility by choosing $a_{i,t}$.

**Stage 1** As subjects are randomly assigned to treatments, we assume that previously accumulated altruistic capital and altruistic preferences, $A_{i,1}$ and $\sigma_i$, are equally distributed for both treatment groups. Hence, the only difference between treatments consists of the social observability. In the Public-Private treatment, we increase the social image parameter $\theta_1$ and therefore the benefit of the generated social welfare.[16] Consequently, the agent has a higher return of her altruistic act and chooses a larger action $a_{i,1}$.

**Hypothesis 4.1.** *Subjects generate a greater donation in Stage 1 in the Public-Private treatment than in the Private-Private treatment.*

**Stage 2** In the Public-Private treatment, observability occurs only in Stage 1 while subjects make their first decision. The subsequent donation decision in Stage 2 is completely private for all subjects and $\kappa_2$ and $\theta_2$ should therefore be similar for both treatment groups. Altruistic capital $A_{i,2}$, however, is no longer equal as participants in the Public-Private treatment choose a larger action $a_{i,1}$ due to $\theta_1$ and experience an additional increase due to a higher $\kappa_1$ (heightened self-awareness). This increases their altruistic capital stock with a higher rate, which in turn decreases the cost $c(a_{i,2}, A_{i,2})$ in period $t = 2$. A reduced cost makes it comparatively more attractive to engage in prosocial activities, which leads to our second hypothesis.

**Hypothesis 4.2.** *Subjects generate a greater donation in Stage 2 in the Public-Private treatment than in the Private-Private treatment.*

[15] Ashraf and Bandiera (2017) assume that altruistic capital increases the marginal product of the altruistic action. Both assumptions are equivalent. We use cost reduction for the intuitive reason that habits reduce the cost of the decision process as well as of the action itself.

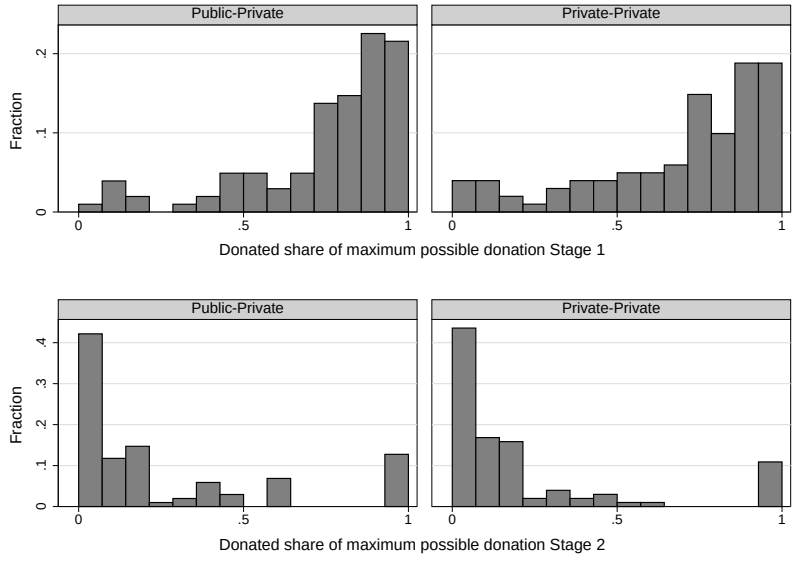
[16] As the existence of Stage 2 is unknown when making the decision for $a_{i,1}$, there is no incentive for any of the two treatment groups to choose a suboptimal $a_{i,1}$ in anticipation of an increased $A_{i,2}$.

## 4.5 Results

### Experiment A

In Experiment A, a total of 203 subjects participate as decision makers, 102 subjects in the Public-Private and 101 subjects in the Private-Private treatment. In Stage 1, in which subjects can generate a donation by correctly counting zeros in tables, about 75% of all subjects solve at least five tables correctly and subjects quit, on average, after 15.9 trials. This results in an average donation of €2.18 in the Public-Private treatment and €2.00 in the Private-Private treatment (out of a maximum of €2.90 if all 25 tables are solved correctly within 20 minutes). Subjects in the Public-Private treatment spend significantly more time working on the task (on average around 14 minutes in the Private-Private treatment and around 17 minutes in the Public-Private treatment; the difference is significant to a significance level of 1%). However, almost all subjects sacrifice some time for the charitable act. In Stage 2, where subjects are no longer asked to spend time and effort but money, only 62% of subjects donate a strictly positive amount at all, albeit 12% give their complete show-up fee of €5. Nonetheless, the average donation is €1.30 in the Public-Private and €1.03 in the Private-Private treatment. Figure 4.2 displays donated shares of the maximum possible amount separately for the two treatment groups and for Stage 1 and Stage 2.

Figure 4.2: Donations in Experiment A: donated share of maximum possible donation



| *Dep. variable:* | Donation 1 | Donation2 | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Public | 0.179* | 0.257 | 0.269 | 0.804 |
| | (0.105) | (0.226) | (0.230) | (0.803) |
| Donation 1 | | | -0.0662 | 0.0425 |
| | | | (0.169) | (0.218) |
| Public x Donation 1 | | | | -0.254 |
| | | | | (0.342) |
| Constant | 2.003**** | 1.039**** | 1.171*** | 0.954* |
| | (0.0796) | (0.154) | (0.389) | (0.489) |
| Observations | 203 | 203 | 203 | 203 |
| $R^2$ | 0.0143 | 0.00638 | 0.00731 | 0.0107 |

*Note:* Coefficients in all columns are OLS estimates. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 4.3: Regression results: Experiment A

Table 4.3 reports OLS estimates for Experiment A. In Column 1, Stage 1 donation is regressed on a treatment dummy, which is 1 if subjects are in the Public-Private treatment and 0 if they are in the Private-Private treatment. The coefficient is positive (subjects donate on average €0.18 more in the Public-Private treatment) and significant to a significance level of 10%. This is in line with Hypothesis 1. In Column 2, Stage 2 donation is regressed on the same treatment dummy. As stated in Hypothesis 2, the coefficient is positive (subjects donate on average €0.26 more in the Public-Private treatment), but not significant. In Column 3, Stage 2 donation is regressed on the treatment dummy, now additionally controlling for Stage 1 donation. The coefficient of the dummy variable stays almost the same compared to Column 2. The coefficient of Stage 1 donation is close to zero, which suggests that a higher giving of Stage 1 does not per se induce higher giving in Stage 2 but that observability itself induces higher giving. However, neither of the coefficients is significant. In Column 4, Stage 2 donation is regressed on the treatment dummy, Stage 1 donation, and the product of Stage 1 donation and the treatment dummy. The interaction term is negative, which could be a hint that for subjects in the Public-Private treatment, Stage 1 donation has a negative effect on Stage 2 donation, speaking against a general altruistic capital effect. However, again, none of the coefficients is significant.

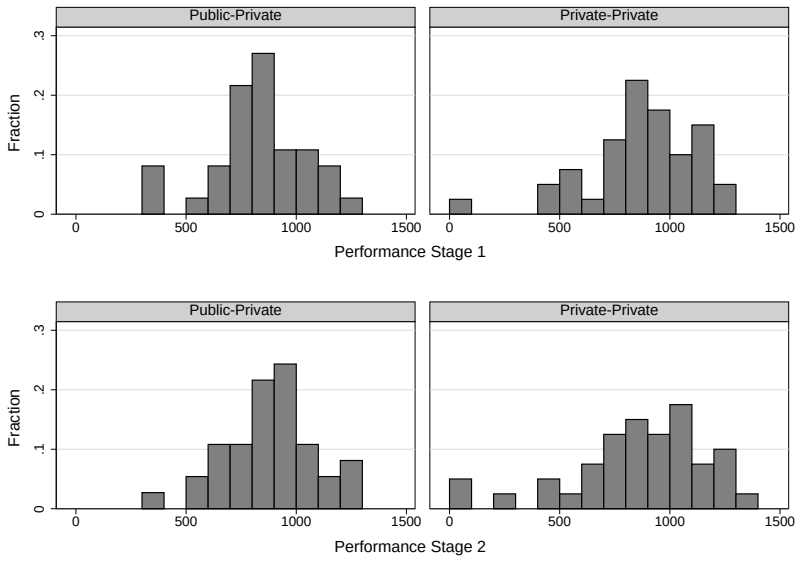
## Experiment B

In Experiment B, a total of 77 subjects participate. 37 subjects are in the Public-Private and 40 subjects in the Private-Private treatment. In this experiment, subjects generate

two donations by working on the real-effort task Click for Charity twice. Figure 4.3 shows the distributions of number of clicks (performance) per subject separately for each treatment group and each stage. We show graphics for performance levels instead of donations, since the concave piece rate leads to a low variation in actual donations. Therefore, performance levels give a more accurate picture of differences in behavior. We show below that qualitative results stay fairly unchanged when running the analysis for donations instead.

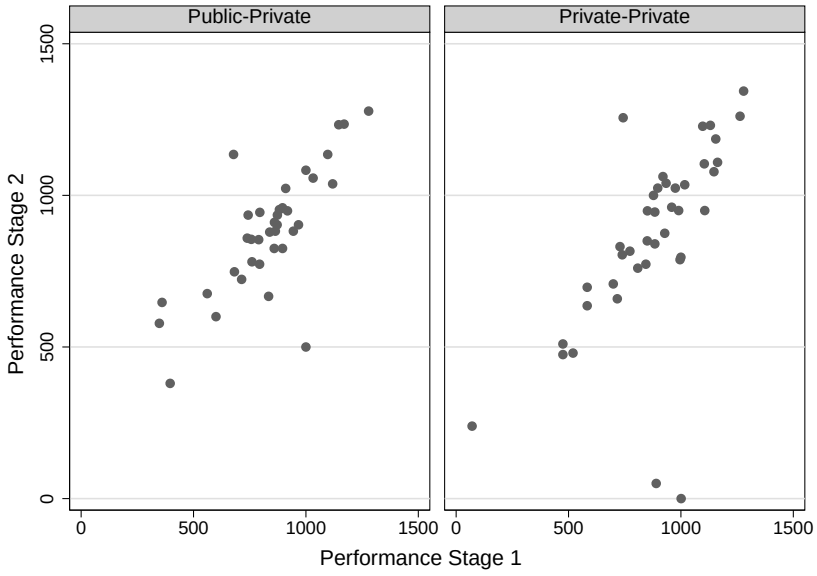
As in the previous experiment, almost all subjects engage in the task and generate a donation larger than zero. The average donation (pressed pairs) in Stage 1 is €2.12 (837.14) in the Public-Private and €2.1 (876.45) in the Private-Private treatment. We do not observe any decline in Stage 2 where the average donation (pressed pairs) is €2.13 (879.54) in the Public-Private and €2.03 (858.1) in the Private-Private treatment. Note that in Stage 1, average donations are higher in the Public-Private treatment, whereas average key combinations are lower. This is possible due to the concave piece rate, which increases donations strongly in the beginning and only weakly in the end. In the Public-Private treatment, subjects press a lower total number of key combinations than in the Private-Private treatment, but the minimum number of pressed pairs is higher. This results in higher average donations.

Figure 4.3: Performance levels (number of correctly clicked combinations) in Experiment B




As Figure 4.4 illustrates, we find a strong positive correlation ($\rho = 0.667$) of performance between stages. Also, the difference of performance between Stage 1 and Stage

Figure 4.4: Relation between performance levels for Stage 1 and Stage 2 in Experiment B

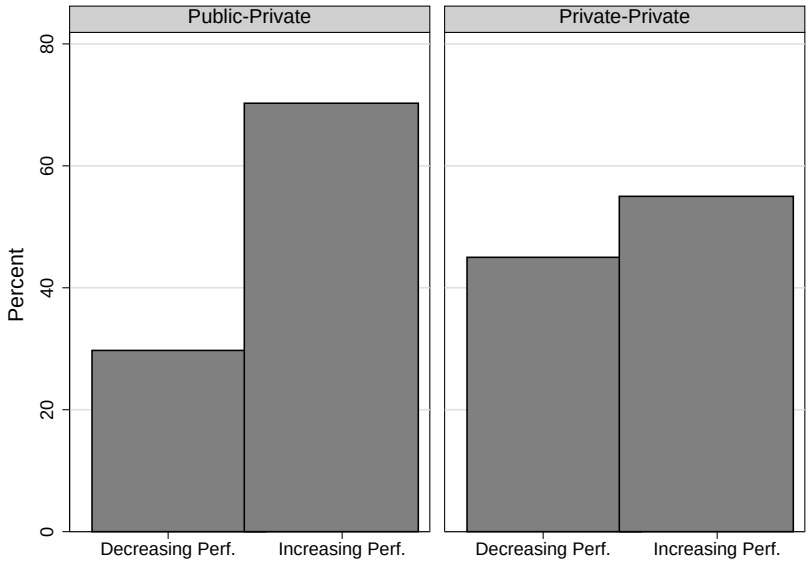



2 is not significantly different from zero (using a $t$-test, $p = 0.637$), which shows that subjects do not decrease their prosocial behavior over time. We also observe that correlations between Stage 1 and Stage 2 performance in the Public-Private treatment ($\rho = 0.76$) and in the Private-Private treatment ($\rho = 0.64$) are not significantly different from each other.

Analyzing individual changes in performances between Stage 1 and Stage 2, we find that in the Public-Private treatment around 70.3% of subjects increase their performance between Stage 1 and Stage 2, whereas only 55% of subjects do so in the Private-Private treatment. This finding is visualized in Figure 4.5. The difference of 15 percentage points between treatments goes in the expected direction but is not significant (Wilcoxon Rank sum Test, p-value of 0.17).

Table 4.4 replicates Table 4.3 for Experiment B and reports OLS estimates. In Column 1, Stage 1 donation is regressed on a treatment dummy, which is 1 if subjects are in the Public-Private treatment and 0 if they are in the Private-Private treatment. In Experiment B, the coefficient is also positive, but not significant. In Column 2, Stage 2 donation is regressed on the same treatment dummy. As stated in Hypothesis 2, the coefficient is positive (subjects donate on average €0.1 more in the Public-Private treatment), but not significant. In Column 3, Stage 2 donation is regressed on the treatment dummy, additionally controlling for Stage 1 donation. The coefficient of the dummy variable stays almost the same compared to Column 2. However, the coefficient is still not significant. The coefficient of Stage 1 donation is positive and highly significant which illustrates

Figure 4.5: Percentage of subjects who increase or decrease their performance between Stage 1 and Stage 2, separately for Public-Private and Private-Private




again that Stage 1 and Stage 2 donations are strongly correlated. In Column 4, Stage 2 donation is regressed on the treatment dummy, Stage 1 donation, and the product of Stage 1 donation and the treatment dummy. Again, the coefficients of the treatment dummy and the interaction term are not significant.

| *Dep. variable:* | Donation 1 | Donation2 | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Public | 0.0199 | 0.102 | 0.0954 | -0.237 |
| | (0.0382) | (0.0682) | (0.0688) | (0.408) |
| Donation Stage 1 | | | 0.317**** | 0.304**** |
| | | | (0.0645) | (0.0690) |
| Public x Donation Stage 1 | | | | 0.157 |
| | | | | (0.199) |
| Constant | 2.100**** | 2.030**** | 1.364**** | 1.392**** |
| | (0.0363) | (0.0677) | (0.0739) | (0.0792) |
| Observations | 77 | 77 | 77 | 77 |
| $R^2$ | 0.00339 | 0.0267 | 0.0571 | 0.0577 |

*Note:* Coefficients in all columns are OLS estimates. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 4.4: Regression results: Experiment B, donations

Since the correlation of performance and generated donations is only piecewise, Table 4.5 reports regressions in the domain of performance. Again, there are no significant treatment effects and none of the hypotheses can be supported.

Even in Stage 1, being observed does not have a significant effect on donation behavior, which we find surprising since we closely follow the design of Ariely et al. (2009) and thus cannot replicate their findings. In contrast to their study, subjects in Stage 1 of our Private-Private treatment actually achieve a higher performance. Both treatment groups accomplish numbers that are similar to those in the public condition of Ariely et al. (2009). Furthermore, we have enough statistical power; a treatment difference in performance similar to the one of Ariely et al. (2009) (on average 822 clicks in the public condition and 548 clicks in the private condition) would be significant to a significance level of 1% with our sample size. We use the same mechanism to implement social observability, as well as the same piece rates, even though the cutoffs are different as we decrease the piece rate in steps of 100 instead of 200. The increased concavity could potentially decrease the treatment difference in donations and therefore explain why we do not find the same results. However, we observe subjects to continue the task even if one click is worth only 0.01 Cent.[17] To summarize, despite closely following the design of Ariely et al. (2009), we are not able to detect any direct effect of public observability on prosocial activities.

| *Dep. variable:* | Performance 1 | Performance 2 | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Public | -39.31 | 21.44 | 52.06 | 124.5 |
| | (51.45) | (58.85) | (42.87) | (109.8) |
| Performance Stage 1 | | | 0.779**** | 0.814**** |
| | | | (0.0651) | (0.0831) |
| Public x Performance Stage 1 | | | | -0.0849 |
| | | | | (0.133) |
| Constant | 876.5**** | 858.1**** | 175.5*** | 144.9** |
| | (38.15) | (48.65) | (56.23) | (64.98) |
| Observations | 77 | 77 | 77 | 77 |
| $R^2$ | 0.00764 | 0.00171 | 0.455 | 0.456 |

*Note:* Coefficients in all columns are OLS estimates. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$; **** $p < 0.001$; robust standard errors in parentheses.

Table 4.5: Regression results: Experiment B, performances

[17] We do not believe that subjects are not capable of pressing more pairs in five minutes, as Ariely et al. (2009) themselves have a control condition in which subjects work for high monetary incentives and press, on average, 1290 combinations, which is also approximately the maximum level we observe.

## 4.6 Conclusion

The aim of this paper is to investigate spillover effects of observability on later unobserved prosocial behavior, thereby studying the concept and prevalence of altruistic capital formation. We hypothesize that being observed during a good deed has a positive effect on subsequent prosocial behavior because people build up altruistic capital. People feel obliged to maintain their positive social and self-image, even in situations in which their actions are not observed by others, and keep on behaving prosocially. We do not find such behavior, independent of the concrete nature of the prosocial act, either requiring a donation of money or investing time.

This lack of supporting evidence is mainly due to the reason that social image as a trigger of stronger prosocial behavior cannot be established in our experiments. We do find only a weakly significant positive effect in Experiment A and an insignificant effect in Experiment B. This result is not driven by a lack of prosocial behavior of subjects in the Public-Private treatment but if anything by a substantial prosocial attitude of the control group that does not face any social exposure in the first place. People are willing to (repeatedly) sacrifice own resources for social welfare, regardless of observability. This suggests other potential drivers of repeated prosocial activity: it is possible that people already have a high altruistic capital stock and a prosocial self-perception and therefore do not react to further motivation. It could therefore be interesting to run a similar study with a subject pool that has not developed high prosocial preferences (e.g. children or people who show very low altruistic preferences).

To summarize, people show substantial willingness to donate for a charitable reason. Surprisingly, image effects do not seem to make a large difference in our chosen setups, neither in the observed stages nor in the following non-observed stages. It might therefore be necessary to adjust the design. It seems that people are willing to spend time and effort to generate donations and keep on working even if they are not observed and even if the piece rate is extremely low. One solution would be to make donating more costly in all stages, for example by not using real-effort tasks but dictator games that actually lower participants' income instead or by making the real-effort task more difficult and thus more costly.

Given that the effect that we hypothesized can be found in a new study, it would be interesting to disentangle different channels. We elaborate in this chapter that we imagine observability to have an effect on subsequent prosocial choices via two channels. First, subsequent prosociality should increase because observability increases initial prosociality and this builds up altruistic capital. And second, because being observed fosters the altruistic capital formation directly. Disentangling these effects would be difficult but could be done, for example, by running an experiment in which, in one treat-

ment, initial prosociality is increased via observability and, in another treatment, via a different incentive.

A consequence of no observation effects is that making conjectures on basic altruistic capital formation is more difficult. In order to test how early prosocial behavior influences subsequent prosocial behavior, one perhaps requires another exogenous manipulation of prosocial behavior, for instance a direct manipulation (such as a manipulation of the prosociality of a task's consequences) or using social reference points. The influence of the behavior of a reference person has been shown in various studies (e.g. Falk and Ichino, 2006; Mas and Moretti, 2009; Schwerter, 2016). In the area of altruistic and reciprocal behavior, information about others' donations increases own contributions to a public good (Shang and Croson, 2009), and subjects match effort provision to associates' levels (Gächter et al., 2012, 2013; Thöni and Gächter, 2015).

In our design, subjects make decisions only within one session, which we see as a conservative test of altruistic capital formation. Nonetheless, it might be the case that the accumulation of altruistic capital requires a longer time horizon with repeatedly triggered prosocial acts to form a habit of acting prosocially or to change people's self-image. Hence, another potential extension would be to conduct a new experiment in which multiple stages take place in different weeks over a longer time horizon.

# References


Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference Points and Effort Provision." *American Economic Review* 101 (2): 470–492.

Akerlof, George A., and William T. Dickens. 1982. "The Economic Consequences of Cognitive Dissonance." *American Economic Review* 72 (3): 307–319.

Algan, Yann, Elizabeth Beasley, Frank Vitaro, and Richard E. Tremblay. 2014. *The Impact of Non-Cognitive Skills Training on Academic and Non-Academic Trajectories: From Childhood to Early Adulthood*. Mimeo.

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

Andreoni, James. 1989. "Giving With Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97 (6): 1447–1458.

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

Ashraf, Nava, and Oriana Bandiera. 2017. "Altruistic Capital." *American Economic Review: Papers & Proceedings* 107 (5): 70–75.

Bašić, Zvonimir, Armin Falk, and Simone Quercia. 2018. *Self-Image, Social Image and Prosocial Behavior*. Mimeo.

Beaman, Arthur L., C. Maureen Cole, Marilyn Preston, Bonnel Klentz, and Nancy Mehrkens Steblay. 1983. "Fifteen Years of Foot-in-the Door Research: A Meta-Analysis." *Personality and Social Psychology Bulletin* 9 (2): 181–196.

Becker, Anke, Thomas Deckers, Thomas Dohmen, Armin Falk, and Fabian Kosse. 2012. "The Relationship Between Economic Preferences and Psychological Personality Measures." *Annual Review of Economics* 4 (1): 453–478.

Becker, Gary S. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82 (6): 1063–1093.

— . 1976. "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14 (3): 817–826.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *Quarterly Journal of Economics* 126:805–855.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." *European Economic Review* 71:117–120.

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

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini. 2009. "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences of the United States of America* 106 (19): 7745–7750.

Carpenter, Jeffrey, and Erika Seki. 2011. "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay." *Economic Inquiry* 49 (2): 612–630.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117 (3): 817–869.

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9:88–97.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang. 2007. "Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness." *Economic Theory* 33 (1): 67–80.

DeJong, William. 1979. "An Examination of Self-Perception Mediation of the Foot-in-the-Door Effect." *Journal of Personality and Social Psychology* 37 (12): 2221–2239.

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

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde. 2012. "The Intergenerational Transmission of Risk and Trust Attitudes." *Review of Economic Studies* 79 (2): 645–677.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47 (2): 268–298.

Dunn, Elizabeth W., Lara B. Aknin, and Michael I. Norton. 2008. "Spending Money on Others Promotes Happiness." *Science* 319 (5870): 1678–1688.

Ewers, Mara, and Florian Zimmermann. 2015. "Image and Misreporting." *Journal of the European Economic Association* 13 (2): 363–380.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity." *Games and Economic Behavior* 54 (2): 293–315.

Falk, Armin, and Andrea Ichino. 2006. "Clean Evidence on Peer Effects." *Journal of Labor Economics* 24 (1): 39–57.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–868.

Freedman, Jonathan L., and Scott C. Fraser. 1966. "Compliance without Pressure: The Foot-in-the-Door Technique." *Journal of Personality and Social Psychology* 4 (2): 195–202.

Gächter, Simon, and Ernst Fehr. 1999. "Collective Action as a Social Exchange." *Journal of Economic Behavior & Organization* 39 (4): 341–369.

Gächter, Simon, Daniele Nosenzo, and Martin Sefton. 2013. "Peer Effects in Pro-Social Behavior: Social Norms or Social Preferences?" *Journal of the European Economic Association* 11 (3): 548–573.

— . 2012. "The Impact of Social Comparisons on Reciprocity." *Scandinavian Journal of Economics* 114 (4): 1346–1367.

Gneezy, Ayelet, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton. 2012. "Paying to Be Nice: Consistency and Costly Prosocial Behavior." *Management Science* 58 (1): 179–187.

Grossman, Zachary. 2014. "Strategic Ignorance and the Robustness of Social Preferences." *Management Science* 60 (11): 2659–2665.

Guerin, Bernard. 1983. "Social Facilitation and Social Monitoring: A Test of Three Models." *British Journal of Social Psychology* 22 (3): 203–214.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2019. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy (forthcoming)*.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.

Mas, Alexandre, and Enrico Moretti. 2009. "Peers at Work." *American Economic Review* 99 (1): 112–145.

Merritt, Anna C., Daniel A. Effron, and Benoît Monin. 2010. "Moral Self-Licensing: When Being Good Frees Us to Be Bad." *Social and Personality Psychology Compass* 4 (5): 344–357.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review* 101 (7): 3221–3252.

Park, Soyoung Q., Thorsten Kahnt, Azade Dogan, Sabrina Strang, Ernst Fehr, and Philippe N. Tobler. 2017. "A Neural Link Between Generosity and Happiness." *Nature Communications* 8:15964.

Powell, Kate L., Gilbert Roberts, and Daniel Nettle. 2012. "Eye Images Increase Charitable Donations: Evidence From an Opportunistic Field Experiment in a Supermarket." *Ethology* 118 (11): 1096–1101.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83 (5): 1281–1302.

Schmitz, Jan. 2019. "Temporal Dynamics of Pro-Social Behavior: An Experimental Analysis." *Experimental Economics* 22 (1): 1–23.

Schwerter, Frederik. 2016. *Social Reference Points and Risk Taking*. Mimeo.

Shang, Jen, and Rachel Croson. 2009. "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods." *Economic Journal* 119 (540): 1422–1439.

Thöni, Christian, and Simon Gächter. 2015. "Peer Effects and Social Preferences in Voluntary Cooperation: A Theoretical and Experimental Analysis." *Journal of Economic Psychology* 48:72–88.

Zajonc, Robert B. 1965. “Social Facilitation.” *Science* 149 (3681): 269–274.

# 4.A Instructions

These are the original instructions translated into English for both experiments. Instructions in italic were only for subjects in the Public-Private treatment.

## Experiment A

### Verbal Instructions to all participants (only Public-Private)

Welcome.
Before we start, I would like to point out that your decisions in this study might not be completely anonymous. If you do not agree to these terms, you now have the possibility to end your participation in this study.
I am now kindly asking you to introduce yourselves with your first name and subject of study or occupation.
In this study, one of you has the role of an observer. All others are decision makers. All decision makers will make a decision at the computer today. The observer is also seated at a computer. This computer shows the computer screens of the decision makers. The only task of the observer is to observe the decisions made by the decision makers. At the end of the experiment, all decision makers furthermore tell the observer personally about the decisions they made.
Everyone now please draw a card. The one of you with the red card takes the role of the observer. This person please stays seated for now.
All others, with the black cards, please follow me to the adjoining room.

### Verbal Instructions (all treatments)

You will receive €5 for your participation. The money is inside this envelope. Please do not open it yet, but only after the screen reads that the experiment is finished. Please sign here that you received the money. You can now go upstairs. You may choose an office room with the numbers 1, 2 or 3. Please close the office door behind you and take a seat at your working space. You can start immediately.

### Instructions Observer (only to observers in Public-Private)

Welcome to this experiment!

You will be taking the role of the observer. Your task will be to observe all participants of the study and their choices.
For your participation in this study you will receive €5 in cash at the end.
On your screen, you can see the screens of three other participants in real time. This means that you will be able to see the participants' choices throughout the course of the study. The other participants have the possibility to generate a donation for SOS Kinderdörfer by solving tasks. At any time the participants may choose to stop solving tasks. After they have stopped working on the tasks, the participants will come to you and inform you about their room number as well as the amount of their generated donation. For this, please leave your cubicle so that the participant cannot see your screen. Please act neutral and do not speak. After all three participants have come to you, your part in the experiment is finished and you will receive your payment. If you have any questions, please only address them to us: raise your hand and we will come to you. The violation of this rule will lead to you being excluded from this study and all its payments.

### Screen 1 – Welcome and General Information

**Welcome!**
Today you have the possibility to generate a donation for the project **Fight against Hunger** of **SOS Kinderdörfer**.
Worldwide, nearly one billion people do not have enough food. Especially children and families are suffering from this and are therefore specifically being supported by SOS Kinderdörfer. SOS Kinderdörfer provide a perspective to children in need and enable parents step by step to self-help: with food, seeds and through education. Through their SOS Family Help they are fighting against poverty and hunger around the globe, e.g. in Bangladesh, Niger and Nicaragua.
You can generate the donation by solving as many tasks as you like out of a maximum of 25 tasks. These tasks will be explained in detail on the following page. *Your observer can see your screen. Right now, your observer can see these instructions and will also see how many tasks you solve and therefore how much you donate.* As soon as you decide not to solve any more tasks, this experiment is over. *Before you may leave, you have to let the observer know if and how much you are donating.*
You will find more details concerning the amount of the donation and the explanation of the task on the following page. After the study is over, we will send you a copy of the confirmation of all donations.

### Screen 2 – Explanation Task

**Explanation Task**
Your task is to count all zeros in a table consisting of zeros and ones.
For every correctly entered total number of zeros in a table, we donate a certain amount to the aforementioned project of SOS Kinderdörfer. This amount varies with the number of the tables as follows:

| | |
|---|---|
| Tables 1–5: | 30 Cents |
| Tables 6–10: | 20 Cents |
| Tables 11–15: | 5 Cents |
| Tables 16–20: | 2 Cents |
| Tables 21–25: | 1 Cent |

For every wrong input, 5 Cents will be subtracted from your donation, whereas the total amount cannot become negative.
You can stop the task **at any time** by clicking the button "Stop task" in the bottom right corner of the screen.
In total, you can solve up to 25 tasks. For this, you have a maximum of 20 minutes.
*Your decision concerning your amount of work and the generated donation will be controlled by your observer. Additionally, after finishing the tasks, you will inform your observer in person about the total amount of your donation.*
Click "Continue" to start with the tasks.

### Screen 3 – Task

So far, you have generated a total donation of X. For correctly solving this task your donation increases by **Y Cents**.
For a wrong answer, 5 Cents will be subtracted.
Count all **zeros** in the following table.

### Screen 4 – Results

You solved X tasks correctly and thus generated a donation of €Y. We will transact this amount for you to the project **Fight against Hunger** of SOS Kinderdörfer.
*Please go now downstairs into the laboratory to your observer and inform him or her about your room number and your generated donation.*
Click "Continue" *as soon as you are back at your desk.*

### Screen 5 – End

The experiment is now over.

### Written Instructions in Envelope

For your participation in today's experiment, you receive €5. These €5 are also inside this DIN A5 envelope. Before the experiment is finished, you have another possibility to donate.
You can donate any part of your €5 to the **Project Bolivia: Children in Poverty Districts** of **SOS Kinderdörfer**. We will transact the donation for you.
In the slums of La Paz, children and their families are living in poverty. Crime, alcoholism and hopelessness are omnipresent. People do not have another possibility than to take on irregular, precarious and merely profitable jobs. One of the main reasons for the low income per household is the low level of education of many parents: only about 15 percent hold any degree. Most mothers and fathers therefore do not have any perspective – and the same fate threatens their children. SOS Kinderdörfer help by supporting the parents with 30 SOS day nurseries, educational projects, microcredits and psychological help.
You can donate any amount between €0 and €5 in steps of 10 Cents.
**Your decision is completely anonymous.**
Please take the money and the small white envelope out of the DIN A5 envelope and place the amount you want to donate inside the white envelope. Following this, close the small envelope and leave it on your desk.
Afterwards, the experiment is over and you can leave the laboratory.

## Experiment B

### Screen 1 – Welcome

**Welcome!**
Please wait until all participants are seated.

### Screen 2 – Mood

**First of all, we would like to know about your current mood.**

For this purpose, please indicate your answer on a scale of 0 to 10.
0 indicates that your current mood is really bad. 10 indicates that your current mood is really good.
You can choose any integer in between 0 to 10 to express your mood.
How is your current mood?

### Screen 3 – Explanation Task

Next, in the main part of the study, you will work on a task. In this task, you will have to **press the X-button and Y-button alternately.**
You now have the possibility to get to know the task by testing it for 60 seconds. By doing so, please try to press as many correct combinations as possible.
Please pay attention to pressing the buttons alternately. Otherwise it might happen that the combinations are not being counted.
Do you have any questions? If yes, please raise your hand. If this is not the case, please click "Continue" in order to test the task.

### Screen 4 – Practice Task

Please press the X-button and the Y-button alternately.
So far, you have pressed X correct combinations.

### Screen 5 – Result

The test phase is now over.
Within 60 seconds, you entered X correct key combinations. If you have any questions concerning the task, please raise your hand. If you do not have any questions, please click "Continue".

### Screen 6 – Information Donation

For your participation in this study, you receive €6. This amount will be paid to you in cash at the end of the study.
In addition, you have the possibility to generate a donation for the project "Fight against Hunger" by SOS Kinderdörfer.

Worldwide, nearly one billion people do not have enough food. Especially children and families are suffering from this and are therefore specifically being supported by SOS Kinderdörfer. SOS Kinderdörfer provide a perspective to children in need and enable parents step by step to self-help: with food, seeds and through education. Through their SOS Family Help they are fighting against poverty and hunger around the globe, e.g. in Bangladesh, Niger and Nicaragua.
You can generate the donation by working on the task you already got to know.
**Important: your donation is completely anonymous.**
***Important: at the end of the experiment, we will ask you to step out of your cubicle and inform all other participants about your name and your total generated donation.***
*While doing so, your generated donation will be indicated on your screen, so that all other participants can read it as well.*
You can find further details concerning the amount of donation on the next page.
After the end of the study, we will make a confirmation of the donation accessible to you.

### Screen 7 – Information Piece Rate

For every correctly entered combination, we will donate a certain amount to the project described on the previous page. This amount per combination varies with the number of already entered combinations.
You can find the exact values in the following table:

| Number of combinations | Donation per combination |
|---|---|
| 1–100 | 1 Cent |
| 101–200 | 0.5 Cent |
| 201–350 | 0.25 Cent |
| 351–500 | 0.1 Cent |
| 501–700 | 0.05 Cent |
| ab 701 | 0.01 Cent |

This means that we donate 1 Cent per combination to the project for the first 100 combinations. For the next 100 combinations, the donation amounts to 0.5 Cents per combination etc.
If you pressed, for example, 160 correct combinations we would donate 100 × 1 Cents +60 × 0.5 Cents = 130 Cents to the project.
To start working on the task, please click “Continue”.

### Screen 8 – Task

Please press the X-button and the Y-button alternately. So far you have pressed X correct combinations and therefore generated a donation of €Y.

### Screen 9 – Result

You entered X correct key combinations and thus generated a total donation of €Y.
We will transact this amount for you to the project "Fight against Hunger" of SOS Kinderdörfer.
Please wait until all participants are finished with the task.

### Screen 10 – Confirmation of Donation

As announced, you have the possibility to check the confirmation of donation for the total amount of donations generated over the course of this study. For this, we will upload a confirmation on the following website within the next couple of days, after transacting the donations:
LINK
You can now copy the website's address or take a picture of it.

### Screen 11 – Information Donation

You now, once again, have the possibility to generate a donation. This time, we will transact the money for you to the project "Bolivia: Children in Poverty Districts" by SOS Kinderdörfer.
In the slums of La Paz, children and their families are living in poverty. Crime, alcoholism and hopelessness are omnipresent. People do not have any other possibility than to take on irregular, precarious and merely profitable jobs. One of the main reasons for the low income per household is the low level of education of many parents: only about 15 percent hold any degree. Most mothers and fathers therefore do not have any perspective – and the same fate threatens their children. SOS Kinderdörfer help by supporting the parents with 30 SOS day nurseries, educational projects, microcredits and psychological help.
**Important: your donation is completely anonymous.** ***Important: in contrast to the first donation you will not have to inform anyone about your donation. Your second***

***donation is completely anonymous. However, you will have to announce your first donation to all other participants of the experiment hereafter!***
You can generate the donation by working on the same task as before. Also, the amount of donation per combination does not change and will be shown to you once again on the next page.
After the end of the study, we will make a confirmation of the donation accessible to you. This confirmation will also be provided on the website to which you have already received the link.

### Screen 12 – Reminder Piece Rate

Here, you can, once again, see the amount of donation for every entered combination:

| Number of combinations | Donation per combination |
|---|---|
| 1–100 | 1 Cent |
| 101–200 | 0.5 Cent |
| 201–350 | 0.25 Cent |
| 351–500 | 0.1 Cent |
| 501–700 | 0.05 Cent |
| ab 701 | 0.01 Cent |

To start working on the task, please click “Continue”.

### Screen 13 – Task

Please press the X-button and Y-button alternately. So far you have pressed X correct combinations and therefore generated a donation of €Y.

### Screen 14 – Result

You entered X correct key combinations and thus generated a donation of €Y.
We will transact this amount for you to the project “Bolivia: Children in Poverty Districts” of SOS Kinderdörfer.
Please wait until all participants are finished with the task.

### Screen 15 – Mood

**We now once again would like to know about your current mood.**

For this purpose, please indicate your answer on a scale of 0 to 10.
0 indicates that your current mood is really bad. 10 indicates that your current mood is really good.
You can choose any integer in between 0 to 10 to express your mood.
How is your current mood?

### Screen 16 – Announcement (Private-Private)

In the following, some of the other participants (you are not part of them) will stand up and announce their names as well as their generated donation of part 1. These participants already knew about this while making their decision.
We ask you to stay seated inside your cubicle during this time, and to open your curtain when we ask you to do so.
After this part is finished, we will start with the payment. Please stay seated inside your cubicle until we call your cubicle number.

### Screen 16 – Announcement (Public-Private)

As announced, before this study is over, you will now inform the other participants on your generated donation from Part 1. For this, we will ask you to stand up and stand in front of your cubicle, whereas the curtain is open.
Some of the other participants will also step out of their cubicle at the same time. You will then sequentially state your first names and your generated donations from part 1.
We ask you to be silent until your turn and to say the following sentence when requested to do so:
**My name is ___ and I generated a donation of ___ Euros in part 1.**
As a reminder, you will see the amount of your donation from Part 1 on the next page.
Please click “Continue”, check the amount of your donation, and step out of the cubicle when we ask you to do so.

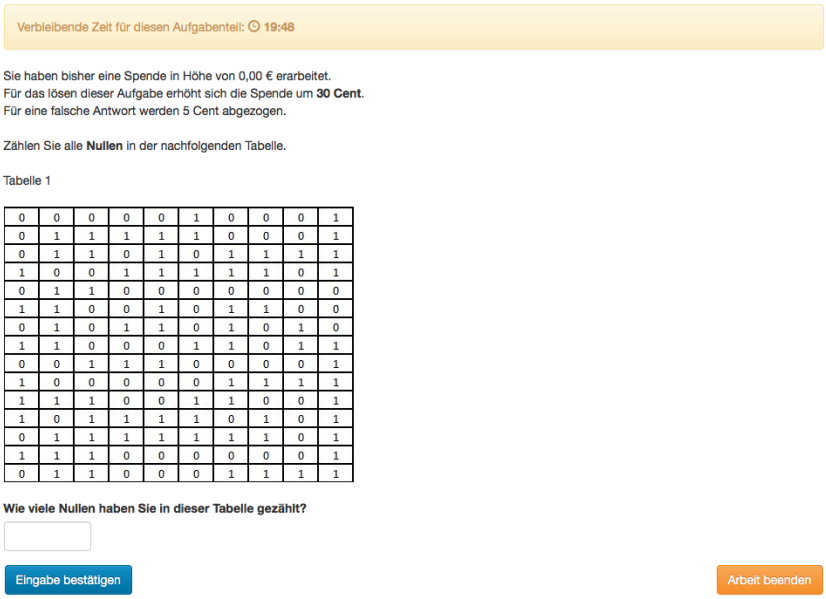
### Screen 17 – Announcement (Public-Private)

My name is ___ and I generated a donation of ___ Euros in part 1.

# 4.B Screenshots

## Experiment A

Figure 4.6: Screenshot of real-effort task Counting Zeros for Experiment A




## Experiment B

Figure 4.7: Screenshot of real-effort task Click for Charity for Experiment B