CESifo WORKING PAPERS
8263
2020
April 2020

Contagion of Pro- and Anti-Social Behavior among Peers and the Role of Social Proximity

Eugen Dimant

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8263*

Contagion of Pro- and Anti-Social Behavior among Peers and the Role of Social Proximity

Abstract

This paper uses a novel experimental design to study the contagion of pro- and anti-social behavior and the role of social proximity among peers. Across systematic variations thereof, we find that anti-social behavior is generally more contagious than pro-social behavior. Surprisingly, we also find that social proximity amplifies the contagion of anti-social behavior more strongly than the contagion of pro-social behavior. Anti-social individuals are also most susceptible to the behavioral contagion of other anti-social peers. These findings paired with the methodological contribution inform the design of effective norm-based policy interventions directed at facilitating pro-social behavior and reducing anti-social behavior in social and economic environments.

JEL-Codes: C910, D640, D900.

Keywords: behavioral contagion, peer effects, anti-social & pro-social behavior.

Eugen Dimant
University of Pennsylvania / Philadelphia / USA
edimant@sas.upenn.edu

This version: April 15, 2020
This paper has been accepted for publication at the Journal of Economic Psychology.
I want to thank Lata Gangadharan and two anonymous referees, as well as Jim Andreoni, Dan Ariely, Max Bazerman, Robert Cialdini, Rachel Croson, René Fahr, Uri Gneezy, Rachel Kranton, John List, Rosemarie Nagel, David Rand, Imran Rasul, Arno Riedl, Tim Salmon, and Marie Claire Villeval for helpful discussions and suggestions. I am also greatly indebted to the input I received from Ed Glaeser, Dan Houser, Larry Katz, David Laibson, Ulrich Schmidt, and Wendelin Schnedler on the early version of the experimental design. I am particularly thankful for valuable input from Cristina Bicchieri, Gary Bolton, Judd Kessler, Muriel Niederle, and Al Roth during my (visiting) research positions at the Harvard University, Stanford University, the University of Pennsylvania, and the University of Texas at Dallas, respectively. This work has greatly benefited from input at numerous conferences and workshops, including the 2016 London Experimental Workshop, the Norms, Actions and Games Workshop at the Institute for Advanced Study in Toulouse, the 2016 CESifo Venice Summer Institute, the 2017 Norms and Behavioral Change Workshop at Nottingham University, the 2017 Behavioral Science Hub Seminar at the LSE, the 2017 Science for Philanthropy Initiative (SPI) conference at the University of Chicago, as well as invited presentations at UC Berkeley, UC San Diego, and Wharton Business School. As part of my dissertation, this paper was awarded the 2015/2016 prize for the best experimental contribution by the German Society for Experimental Economic Research (GfeW). Financial support by the German Research Foundation (DFG) through the SFB 901 is gratefully acknowledged.

1. Introduction

This paper is part of a growing body of literature on social norms, peer-effects, and conformity that has generated substantial scientific interest in the last two decades. Scholars in economics and psychology have advanced our understanding of peer effects and their underlying mechanisms by capitalizing on comprehensive laboratory and field studies, particularly with respect to pro-social behavior, cooperation, and reciprocity (e.g., Frey and Meier, 2004; Croson and Shang, 2008; Shang and Croson, 2008; Gächter et al., 2013; Thöni and Gächter, 2015; Dimant and Hyndman, 2019). Peer effects on anti-social behavior have been examined with respect to doping (Gould and Kaplan, 2011), dishonesty (Gino et al., 2009; Fortin et al., 2007; Lefebvre et al., 2015; Buckenmaier et al., 2018; Gross et al., 2018) and theft (Falk and Fischbacher, 2002). Peers have also been found to affect academic gains (Duflo et al., 2011), investment decisions (Bursztyn et al., 2014), littering behavior (Cialdini et al., 1990), productivity at work (Ichino and Maggi, 2000; Falk and Ichino, 2006; Mas and Moretti, 2009), juvenile behavior (Damm and Dustmann, 2014), and charitable giving (Meer, 2011; Smith et al., 2015). Recently, Andreoni et al. (2017) and Fatas et al. (2018) found compelling evidence for preference conformism among peers. All these insights have enriched our understanding of peer effects across social and economic domains.

We extend the existing literature in a number of ways and addresses three main questions by studying the drivers of behavioral contagion among peers, which we refer to as the change in behavior in response to social interaction (Wheeler, 1966). First, does the contagion of pro-social behavior (defined as contributing to a social good) differ from the contagion of anti-social behavior (defined as subtracting from a social good)? Second, how does social proximity to peers affect behavioral contagion? Third, are pro-social individuals affected differently by contagion than anti-social individuals? The third question is particularly relevant from a policy point of view, as it allows policymakers to micro-target the right individuals and improve the effectiveness of norm-based and peer-based interventions to reduce deviance (see, e.g., Miller and Prentice, 2016; Bolton et al., 2019).

A number of aspects set this paper apart from existing research investigating peer effects or conformity, which are often used interchangeably. For one, most of the related literature has studied this topic in interdependent and interactive environments (i.e. public goods games or trust games), in which strategic components and monetary concerns play important roles, and behavior is affected through those channels as well (de Oliveira et al., 2015; Charness et al., 2019). We purposely study behavior in a novel non-strategic setup to isolate the drivers of behavioral contagion. This methodological contribution is inspired by and avoids the well-documented challenges of studying peer effects (e.g. endogeneity and reflection problems, as well as selection effects), in which own and peer behaviors are circular and ultimately challenge causality (Manski, 1993, 2000; Angrist, 2014).[1] In addition, we extend our setup to the realm of anti-social behavior in order to study contagion in a richer setup that broadens the action space beyond (non-)giving in a pro-social setting (see, e.g., Croson and Shang, 2008; Shang and Croson, 2008), the importance of which has been convincingly argued by List (2007) and Bardsley (2008). In our setup, pro- and anti-social behavior can thrive simultaneously.

The novel experimental design is a sequential two-stage variant of a give-or-take donation game (a variant of List, 2007; Bardsley, 2008) with a revision option. The experiment consists of four stages, details of which were only revealed once the participant reached the respective stage. At **Stage 0**, participants fill out a 25-item personality questionnaire taken from a major US dating website. At **Stage 1**, every participant is matched with a charity and is given the opportunity to give money to or take money away from the charity. At **Stage 2**, to account for the aforementioned challenges, participants are randomly divided into either *active* or *passive* roles. Passive participants do not make any further decisions and are also not able to change their initial behavior from stage 1. Conversely, participants in the active role observe the stage 1 behavior of exactly one peer in the passive role at random and – depending on the treatment – may in addition observe the social proximity score with that particular peer (based on the number of common answers at stage 0). Finally, at **Stage 3**, the participants in the active role may revise their initial decision and consequently replace their initial behavior from stage 1.

Across systematic variations of observable proximity to peers and opportunities to react to peer exposure, we find that behavioral contagion is asymmetric in that anti-social behavior is generally more contagious than pro-social behavior. Surprisingly, we also find that social proximity amplifies the contagion of anti-social behavior more strongly than the contagion of pro-social behavior, and that anti-social individuals are most susceptible to behavioral contagion, especially through socially close peers, both in frequency (preva-

[1] Following the literature, we refer to endogeneity and reflection problems as an issue that "arises when a researcher observing the distribution of behavior in a population tries to infer whether the average behavior in some group influences the behavior of the individuals that comprise the group." (Manski, 1993, p. 532). Selection effects indicate that individuals are rarely allocated randomly to peer groups in real life. Rather, individuals choose their peers, hence avoiding an exogenous variation of peer exposure.

lence of contagion) and magnitude (extent of contagion). Beyond adding to our general understanding of peer effects, these findings point to an important interaction between social proximity and behavioral contagion. From a policy perspective, this interaction is highly consequential for designing behavioral interventions, since it provides a lever to influence the behaviors of different types of individuals and has the potential to improve the effectiveness of norm-interventions (Hallsworth et al., 2017; Bicchieri and Dimant, 2019).

The rest of this paper is structured as follows: we discuss the experimental design and procedure and derive testable hypotheses in Section 2 and present the results in Section 3. We discuss the results and conclude in Section 4.

2. Design of the Experiment

2.1. Payoff Structure

In our experiment both participants and the charity with which they interact begin with a provisional endowment of €15 each. Participants are able to choose any integer value within the range [0, 15] and hence can make up to €30 if they decide to take all the money from the charity, or a minimum of €0 if they decide to give all of their money to the charity. In order to deal with the aforementioned literature that identifies reflection, selection, and strategic interaction as challenges for peer effect studies, the implemented payoff structure in the experiment contains a series of important precautionary features.

For one, the participant's decisions only affected one's own and the chosen charity's payments, but never those of other participants. The payoff procedure is clearly explained to every participant in the instructions: *At the beginning of the experiment* participants learn that only one participant would be randomly chosen *at the end of the experiment* and only the decisions of this randomly chosen participant would be relevant in determining his or her payoff and the payoff of one charity. The remaining n-1 participants receive a flat payment of €7.50 (corresponding to €10 per hour given the 45 minutes duration of the experiment).[2] With this, we employ a payoff function that was orthogonal to the

[2] It is worth clarifying why participants who were not chosen at random did not simply receive the initial endowment of €15 but €7.50 instead: to achieve proper calibration in the experiment, a pre-test was run with the same student population (but different subject pool), which indicated an extreme willingness to donate money to charities. This helped us to inform our decision and served as a means to get a symmetric distribution of taking and giving, which would be useful for the observation stage. As our results indicate, this adjustment achieved the desired effect. In principle, however, our design would have still been able to cope with an asymmetric distribution of behavior as long as at least a few participants would have engaged

treatment variations using the established ‘pay one’ random payment mechanism that is often preferred in setups like ours (Charness et al., 2016). This method retains incentive compatibility as theoretically argued by Azrieli et al. (2018) in which the authors show that in general environments (including strategic choices) paying for only one decision is the best mechanism. Ample experimental evidence exists finding practically no difference between ‘pay one‘ and ‘pay all’ methods (for a comprehensive discussion see Charness et al. 2016). In light of this compelling theoretical and experimental evidence, these design choices served the purpose of dealing with the challenges of studying peer effects while retaining incentive-compatibility and appropriate stake sizes, although the latter has not found to be crucial for behavior in dictator settings like ours (Carpenter et al., 2005).

For another, we attempt to minimize wiggle-room with respect to the charity choice: after the experiment ended, only the participant randomly chosen to have made payoff-relevant decisions towards the charity was able to suggest her charity of choice that would receive the respective payment. Following their last decision, all participants were shown a summary screen and a note stating whether they were the one randomly chosen to make the payoff-relevant decision towards the charity. Logistically, this participant was the last one to be paid by the experimenters and the actual donation to the chosen charity was done with this participant in attendance.[3]

Taken together, we purposely implemented a conservative incentive structure, which works against finding contagion effects in the first place. This lower-bound approach was deemed necessary to ensure that the previously mentioned challenges in peer effect research are accounted for. A detailed description of the experiment follows.

2.2. Experimental Procedure

We implement a two-step give-or-take donation game, in which participants are matched with a charity, including a possibility to revise their initial behavior. Across three treatments, we vary the social proximity signal that participants observe about their peers. This design allows us to study behavioral contagion as the result of active participants (with non-fixed initial behavior) observing passive participants (with fixed initial behavior) with varying social proximity. Figure 1 depicts the experimental design and is explained in

in extreme pro- and anti-social behavior.

[3]In one case, a participant who was chosen at random could not think of a particular charity. The experimenters then showed the participant a list of the most reputable charities according to CharityWatch.org and asked the participant to choose a charity from that list.

detail below. Participants are sequentially presented with four stages, which are the same for all treatments: **Stage 0** - social proximity questionnaire, **Stage 1** - action (initial give-or-take decision), **Stage 2** - observation (of exactly one passive peer's behavior with varying proximity), **Stage 3** - reaction (revision of initial decision from Stage 1).

Importantly, participants learned the details of each stages once they were reached. To minimize surprise effects, participants were told at the beginning of the experiment that the experiment consists of multiple stages and that specific information about each stage would only be given once they get there. To further make Stage 1 and Stage 3 decisions as comparable as possible, in Stage 1, participants were informed that their decisions will remain anonymous in the sense that neither the experimenter nor other participants will be able to link their give-or-take decision to their identity.

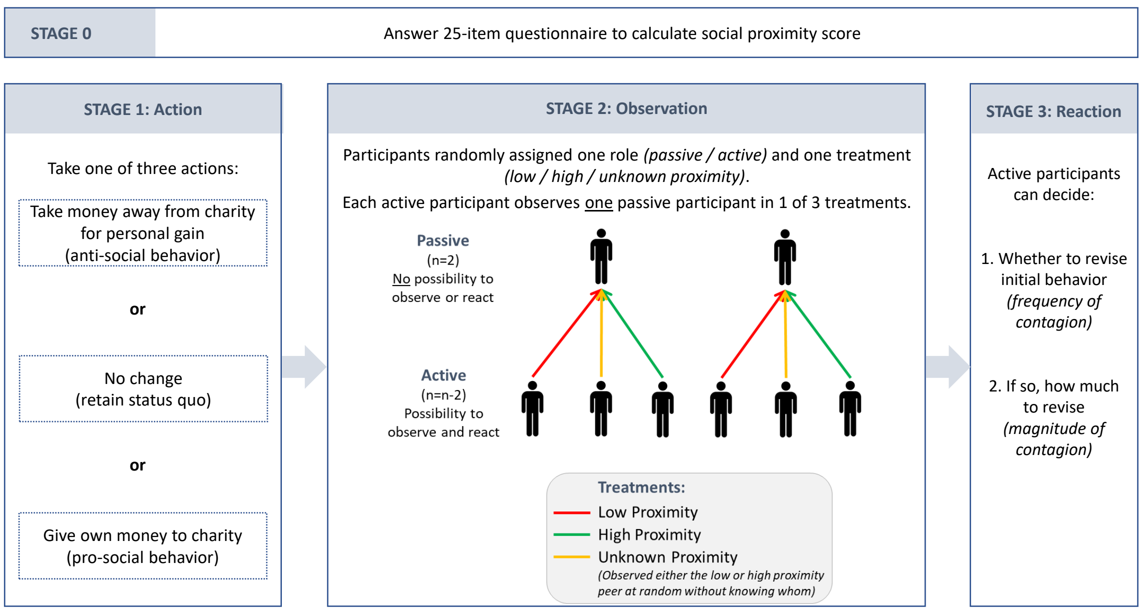


Figure 1: Experimental Design and Procedure. In all treatments, the observed information is the passive player's initial decision towards the charity. Note that in the baseline (unknown proximity condition), active players are unaware of the proximity score and only observe behavior.

In sum, our design allows us to study behavioral contagion in two ways: in *frequencies* and in *magnitudes*. The *frequency* of contagion is a binary measure and equals 1 if active participants revised their behavior following peer observation and 0 otherwise. In contrast, and in line with the previous motivation of the paper, the *magnitude* of contagion is a measure of the extent to which individuals changed their behavior following peer exposure.

A value greater (smaller) than zero indicates that the participant has become more pro-social (anti-social) after peer observation compared to that participant's initial behavior prior to peer observation. We will use this classification of behavior to analyze our results in several ways that are in accordance with the research questions motivated earlier: Does behavior spread? If so, to what extent? How does proximity affect this contagion, if at all?

Stage 0 - Prior to First Decision

We mimic social proximity through similar interests and preferences. We chose 25 real questions from a major American dating website (see the Appendix for details), and let the participants in our experiment give binary answers as to whether or not the content of the questions applies to them. The answers to these questions were then used to calculate a proximity score among participants in later stages.[4]

Stage 1 - Action

Each participant is paired individually with a charity[5] that is endowed with €15, and is then asked to take one of three mutually exclusive decisions towards that charity (amounts were on a continuous scale, allowing any integer value):

1. donate part or all of their own endowment to the charity's account
2. maintain the initial equal distribution
3. take part or all from (and hence harm) the charity and keep the amount for themselves

Note that in order to retain a clean observation of initial behavior, at this point the participants are *not* aware of any forthcoming specifics of the experiment that would otherwise affect their behavior (including exact number of actions throughout the experiment, the random allocation to passive and active participant and peer observation that has yet

[4]Common preferences or interests are often the first step in creating a common identity and successful matching (Hitsch et al., 2010). This aligns with our goal to create relative strengths of proximity beyond binary minimal-group-like measures. In economics, the role of social information in affecting individual behavior has previously been studied by Bohnet and Frey (1999); Charness et al. (2007); Chen and Li (2009); Benjamin et al. (2010); Eckel and Petrie (2011), among others.

[5]Following Eckel and Grossman (1996); we use a charity instead of a participant within each session to increase the salience of the pro- or anti-social decisions, since the behavior can either benefit or harm a credible institution delivering a public good. In what follows, we refer to the act of giving money to a charity (at personal cost) or taking money from a charity (for personal gain) as an act of pro- or anti-sociality, respectively. Participants were made aware that the endowment of the charity represents the experimenter's money that was allocated to a fund, which was prepared for donation to the respective charity right after the experiment ends (including a proof of payment uploaded to the website of the laboratory).

to occur in Stage 2, or the revision opportunity that has yet to occur in Stage 3.)

Stage 2 - Observation

We employ two randomizations as soon as participants reach this stage: (ii) role allocation (passive or active) and (ii) treatment allocation of the participants. Participants were not aware of either existence beforehand.

(i) *Role allocation*: we announce for the first time that there will be two randomly determined roles in the experiment, *passive* and *active*. The specifics of each role were explained to all participants on their computer screen before announcing which role one is assigned to them for the remainder of the experiment (see Appendix for screenshots). The assignment to two different roles is a key component of our design to deal with the aforementioned methodological challenges (Manski, 1993, 2000; Angrist, 2014).

- *Passive Role*: in each experimental session, exactly two participants were chosen to be passive and remain so throughout the next parts of the experiment. The feature of this role is that no further actions towards the charity could be taken. Instead, this participants' Stage 1 behavior would be displayed to the remaining n-2 *active* participants to which those could choose to react knowing that the passive participant's behavior will be held fixed. To retain maximum control and increase statistical power, passive participants were chosen through a quasi-random procedure from the extreme tail of the distribution (either giving away their entire endowment or taking the charity's entire endowment). This means that in all treatments and sessions, one of the two passive observees was a fully pro-social participant, and the other passive observee was a fully anti-social participant.[6]

- *Active Role*: the remaining n-2 participants are randomly allocated into one of three treatments (unknown/low/high proximity). Active participants were able to revise

[6]Choosing behavior from the extreme tail of the distribution across all sessions and treatments allows us to make behavior comparable in its simplest form. In each session, we were able to choose among multiple participants who initially engaged in the maximum amount of pro- and anti-social behavior, which allowed us to randomize among them. This approach is equivalent to existing experimental approaches in which selective behavior was revealed to participants (as is done in, e.g., Bicchieri and Xiao, 2009; Bicchieri et al., 2018). Although it did not occur, had session not have enough extreme pro- or anti-social behavior, participants would have observed the most extreme peer behavior in that session. Because all treatments were conducted within in each session, this would have avoided any differential peer exposure and could have been controlled for in our regression analysis. Nonetheless, we did not run into this potential obstacle.

their initial behavior from Stage 1 following the observation of exactly *one* passive participant. No inference about the behavior of any other active or the second passive participant was possible by design.

(ii) *Treatment allocation*: Treatment randomization occurs at the level of social proximity observed towards *one* passive participant. In all treatments, the observation of a passive peer contains two pieces of information: the type of behavior (taking or giving) and the exact monetary amount of the behavior, which was always take-all or give-all. Treatment variations include the extent to which social proximity information is provided to active participants. They are either given: (a) no information on proximity (baseline condition) and could observe, without knowing which, the behavior of either a high or low proximity peer with equal probability.[7] Across all treatments, the unique advantage of this design is that the same passive participant can be either high or low proximity to active participants:

- **Unknown Proximity Treatment:** The active participant randomly observes the actual behavior of one of the two passive participants, but does not receive any information about the actual proximity scores between them and either of the observees.
- **High (Low) Proximity Treatment:** Active participants are told that the calculated proximity score between them (the observer) and the observed passive participant is *higher* (*lower*) compared to the second unobserved passive participant.

Stage 3 - Reaction

After observation, only players in the *active role* are given the option to revise their Stage 1 decision towards the charity and, if choose to revise, the revision decision replaces the initial decision.[8] The experiment ends with all participants observing a summary screen containing their own decisions as well as a note of whether they were randomly chosen to have made payoff-relevant decisions towards the charity (see screenshots in Appendix).

[7] At no point participants receive information about the exact proximity score with their peer. Instead, the information is limited to the relative proximity score, indicating that the proximity score with the observed passive player is higher or lower than with the second (but unobserved) passive player. The exact computation method of high and low proximity was disclosed to every participant and is based on the number of congruent answers in the initial questionnaire.

[8] Importantly: unbeknown to the participant at the moment of making a revision decision, the experiment ends after this stage and no additional decisions towards their charity will take place. This was necessary in order to keep Stage 1 and Stage 3 decisions comparable as participants were unaware of whether and how many additional (revision) decisions would take place throughout the experiment.

2.3. Theoretical Arguments and Hypotheses

In our experiment, observed behavior can reasonably be interpreted as a trade-off between one's own preferred behavior and the distance in behavior to the observed peer. Any utility function that captures this trade-off is a candidate model. For our purposes, we nest our testable hypotheses within a simple two-stage extension of the quadratic utility conformism model as introduced by Akerlof (1997), which is motivated by the classic work on the economics of conformism by Jones (1984).[9] We present an intuition of the model predictions below and discuss the model formally in the Appendix.

At the heart of our extension of Akerlof's model is the introduction of a social proximity parameter $\alpha_t^{ij} \in (0, 1)$, which depicts the social proximity of individual i with a peer j at time $t \in [1, 2]$, which represents the two points in time at which the active participant makes a decision. The proximity parameter moderates the trade-off between an individual's inherent preference for taking or giving behavior θ^i on the one hand and the actual behavior of the active peer x_t^i and the passive peer x_t^j (which will be fixed in our experiment to retain comparability across treatments) on the other hand.

Following our experimental design, peer behavior is not observed at t_1. Hence, in the absence of any other information except own preferences, an individual will mould his expectations of the preferences of others on his own, $x_t^i - x_t^j$, which is in line with literature on the value of information (such as Grout et al. 2015). From this follows directly that the behavior in stage 1 simply reflects one's inherent preference $x_t^i = \theta^i$.

After peer observation, individual i maximizes her utility in t_2 in the following way:

$$\max_{x_t^i} U_t^i = I - (1 - \alpha_t^{ij})(x_t^i - \theta^i)^2 - \alpha_t^{ij}(x_t^i - x_t^j)^2 \tag{1}$$

As detailed in the appendix more formally, we can now derive predictions for what we refer to as *Contagion Gap*. We define a *Contagion Gap* as the difference in behavior

[9] In particular, we focus on the quadratic utility version of the conformity model by Akerlof (1997) because it generates unique equilibria from which we can derive testable hypotheses for our experiment and it is more in line with the general story of our paper. While Akerlof's general model puts more emphasis on one's own decisions, the quadratic utility model has an a priori assumption that puts individual and peer behavior on an equal footing with respect to how behavior affects individual utility. Assuming that individuals are on a continuum between extremely selfish and extremely altruistic, the quadratic utility approach is conclusive. It should be noted that this assumption is not crucial to our model's predictions because we mainly focus on the relevance of two factors, i.e. social proximity and the pro-/anti-sociality of observed behavior, in affecting behavioral contagion. Thus, the equal-weight assumption is not decisive in predicting the direction in which individual behavior changes as a function of those two factors.

between the observer and observee after revision compared to the difference in behavior prior to peer observation, formally $|x_1^i - x_1^j| - |x_2^i - x_1^j|$.

As derived in the appendix, the intuition behind the role of social proximity in affecting the contagion gap is straightforward: the extent to which the behavior of active peer i after peer exposure (t_2) converges to the passive peer's j initial behavior (t_1) is largely driven by the strength of the social proximity α_t^{ij} between peers. The further away individual i's inherent preference for pro-/anti-social behavior is from individual j's initial behavior, the stronger α_t^{ij} has to be in order to close the contagion gap. From this, we can generate our testable hypotheses:

***H_1*:** *Exposure to peer behavior will generate substantial behavioral contagion.*

***H_2*:** *Social proximity amplifies both pro- and anti-social behavioral contagion.*

We can also turn to existing experimental research to substantiate our claims and find support for our hypotheses independent of our formalization. Following the existing literature on social proximity, it is reasonable to assume that observing the behavior of people who are socially closer matters more in terms of what is socially accepted or an existing norm (e.g., Charness et al. 2007). Existing research indicates that social proximity predicts behavior in different contexts related to charitable giving, fairness, punishment, reciprocity, and trust (Akerlof, 1997; Charness et al., 2007; Chen and Li, 2009; Hugh-Jones et al., 2017; Bicchieri et al., 2019a), as well as neighborhood effects (Damm and Dustmann, 2014). Norms or behavioral prescriptions have also been shown to be associated with one's identity, thus rendering behavioral contagion more likely in situations in which one's identity is salient to peers (Akerlof and Kranton, 2000; Benjamin et al., 2010). Results on advice seeking and advice giving substantiate the claim that norm signaling of what is socially acceptable is more relevant coming from people who are socially closer (Gino and Moore, 2007). From an evolutionary point of view, a number of ultimate-proximate reasons (i.e., kin selection, inclusive fitness, and evolution of fairness) stress the role of proximity and observability of behavior in affecting the likelihood of altruism (Henrich et al., 2010).

3. Results

We conducted the experiment at the BaER-Lab at the University of Paderborn, Germany. Participants were recruited using ORSEE (Greiner, 2015). The experiment was run on zTree (Fischbacher, 2007). We obtained behavioral data for 185 participants through 9 sessions, in which participants were randomly assigned to one of the three treatments (unknown proximity, high proximity, and low proximity). Each session lasted about 45 minutes and the hourly average earnings were roughly $13 (see Table 1 for details).

Table 1: Descriptive Statistics

	Treatments		
	Unknown Proximity	*High Proximity*	*Low Proximity*
Total Participants (active and passive)	76	46	80
Active Participants	71	40	74
Female	60%	54%	59%
Age (average)	24.0 (4.86)	23.9 (3.76)	23.2 (2.80)

Notes: Standard deviations in parentheses. Across all treatments, there are no significant differences among all characteristics that can be compared prior to the treatment manipulation.

The histogram in Figure 2 indicates that prior to observing their peer's behavior, active participants' decisions mainly clustered around 0%, which represents the decision to not change the initial equal distribution between oneself and the charity. After the observation, however, we notice a perceptible reaction towards anti-social behavior, particularly in the high proximity condition, yielding significant differences at the 5% and 1% level, respectively. This finding provides a first indication that anti-social behavior is more contagious than pro-social behavior and will be examined in more detail shortly.

In what follows, we test whether and to what extent contagion is driven by both peer behavior and proximity, both in isolation and combined. The dependent variables of interest are the decision to revise (*frequency of contagion*) and the exact amount of the revision in response to observing a peer (*magnitude of contagion*).

3.1. Behavioral Contagion: Role of Peer Behavior

Our first approach to answering our hypotheses is to look at the frequencies (upper part of Figure 3) and magnitudes (lower part of Figure 3) of behavioral contagion conditional

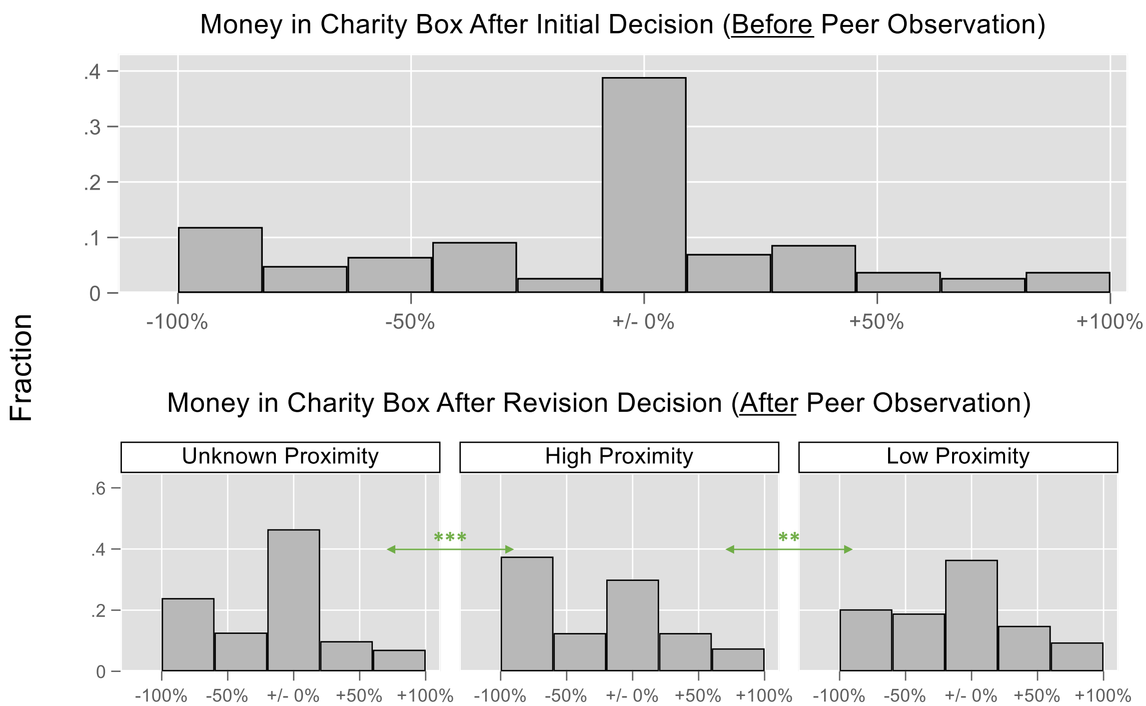


Figure 2: Distribution of behavior (money left in the charity box) before and after peer observation, with 0% representing the status-quo allocation. Green arrows and stars represent significant differences of distributions using $\tilde{\chi}^2$-tests. **p<0.05, ***p<0.01.

on the type of observed *peer behavior*.[10]

We employ a standard population proportion test to examine significant differences in frequencies of behavioral contagion. We find that observing anti-social behavior triggers behavioral contagion significantly more often than observing pro-social behavior (33.7% vs. 19.4%, p=0.027). For the magnitudes of behavioral contagion, we find that after observing anti-social behavior, participants become 13.6% more anti-social than they were before observing peer behavior. After observing pro-social behavior, participants become more pro-social, but to a lesser extent, by 1.9%. Using Mann-Whitney U (MWU) statistics, we find this difference to be significant at the 1% level (-13.6% vs. 1.9%, p<0.001).[11]

[10]For robustness purposes, we perform such an analysis for a number of cases and present them in Section B of the Appendix. Exemplary, one alternative way of looking at the data is taking into account the relative difference of own and observed behavior by analyzing participants who observed better behavior (i.e. more pro-social or less anti-social) or worse behavior (i.e. less pro-social or more anti-social) separately.

[11]Because we run pairwise comparisons instead of analyzing the full factorial design in Sections 3.2 and

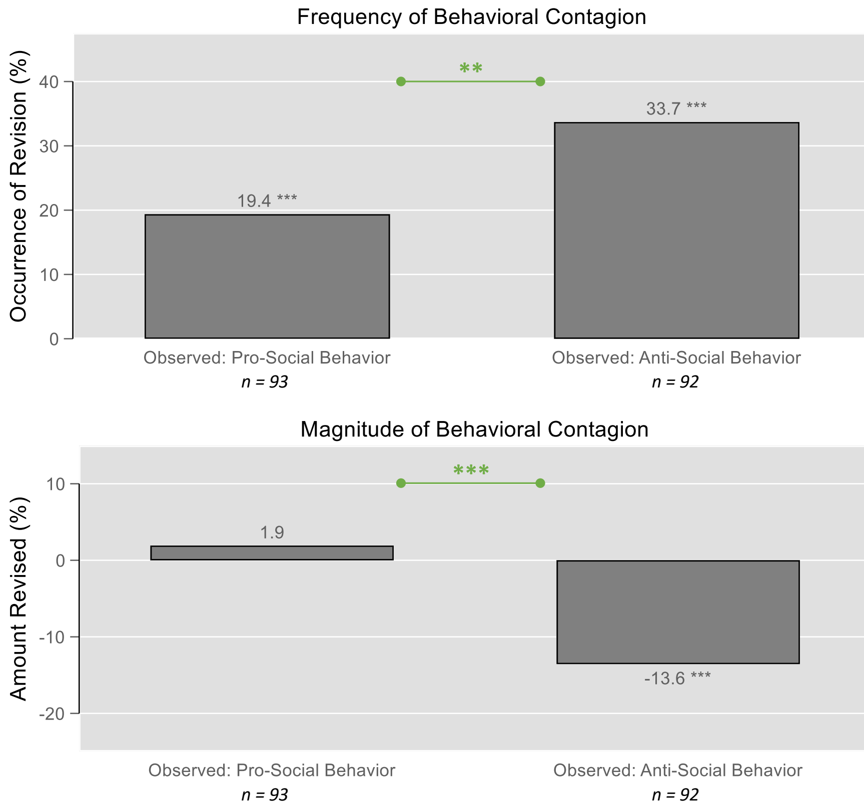


Figure 3: Frequency (upper section) and magnitude (lower section) of behavioral contagion. Frequency: prevalence of behavioral contagion (revision of initial behavior into the direction of observed behavior). Magnitude: extent of the revised behavior relative to one's initial behavior. Observed pro-social/anti-social: behavior of the observed passive peer. Green arrows and stars represent significant differences of compared means, while stars next to the mean number of each bar depict significant difference from zero. **p<0.05, ***p<0.01

Overall, these results suggest that while exposure to peer behavior indeed leads to substantial contagion, and is thus in line with hypothesis H_1, the extent to which contagion occurs in both behavioral domains is asymmetric.[12] In fact, with a ratio of about 7:1 (= 13.6/1.9) the downward adjustment of behavior is much more pronounced than the

3.3, we take into account that such multiple comparisons cause an inflation of the type-I-error. To counteract this inflation, we capitalize on the false discovery rate correction as introduced by Benjamini and Hochberg (1995), which has shown to be superior to Bonferroni with respect to statistical power. Our results of interest are not dependent on the type of the error correction due to the highly significant differences.

[12]The extent to which individuals react to peer information is in line with other research (e.g., Thöni and Gächter, 2015; Bicchieri et al., 2019b, see also Herbst and Mas, 2015)

upward adjustment, indicating that anti-social behavior is more contagious than pro-social behavior. This is true for both the frequency and, particularly, the magnitude of contagion.

3.2. Behavioral Contagion: Role of Social Proximity

Next, we examine behavioral contagion conditional on the type of observed *social proximity* of the peer. Our results, as shown in Figure 4, robustly indicate that higher social proximity indeed triggers stronger behavioral contagion, particularly contagion of anti-social behavior: with increasing social proximity, the frequency and magnitude of revised behavior increases as well. The equality of proportion statistics indicate that the differences are all highly significant. In particular, the occurrence of behavioral contagion in the high proximity condition is more than twice as likely compared to the low proximity condition (57.5% vs. 25.7%, $p<0.001$), more than five times as likely in the high proximity condition compared to the unknown proximity condition (57.5% vs. 9.9%, $p<0.001$), and more than twice as likely in the low proximity condition compared to the unknown proximity condition (25.7% vs. 9.9%, MWU, $p=0.013$). We attribute the latter finding to a saliency effect of observation (Bolton et al., 2019; Bradley et al., 2018) in combination with the fact that participants were matched on similarities rather than dissimilarities, and hence low proximity seems to be perceived as socially closer than unknown proximity.[13] This is also in line with the arguments of our theoretical model and Dimant and Hyndman (2019).

Furthermore, we see a similar picture with regards to the magnitude of behavioral contagion. With negative values across all treatment specifications, the results indicate that participants over-proportionally adjust their behavior downwards (becoming more anti-social) than upwards. Again, behavior in the high proximity condition is significantly different from the unknown proximity condition (-18.8% vs. -0.4%, MWU, $p=0.012$) and the low proximity condition (-18.8% vs. -4.0%, MWU, $p=0.025$). We do not observe magnitudes in the low proximity condition to be significantly different from the unknown proximity condition. From these results, we conclude that high social proximity triggers significantly higher behavioral contagion in both frequency and magnitude, and is particularly true for the contagion of anti-social behavior. This supports hypothesis H_2.

[13] Due to the nature of our proximity measure, these results suggest that providing clues about similarities renders peer behavior more relevant and, thus, leads to more contagion than situations where only the peer's behavior is observed. This is supported by answers to the post-experimental questionnaire. We can only speculate as to whether this effect would be reversed when participants were forced to focus on dissimilarities with their peer, which was not the topic of our project.

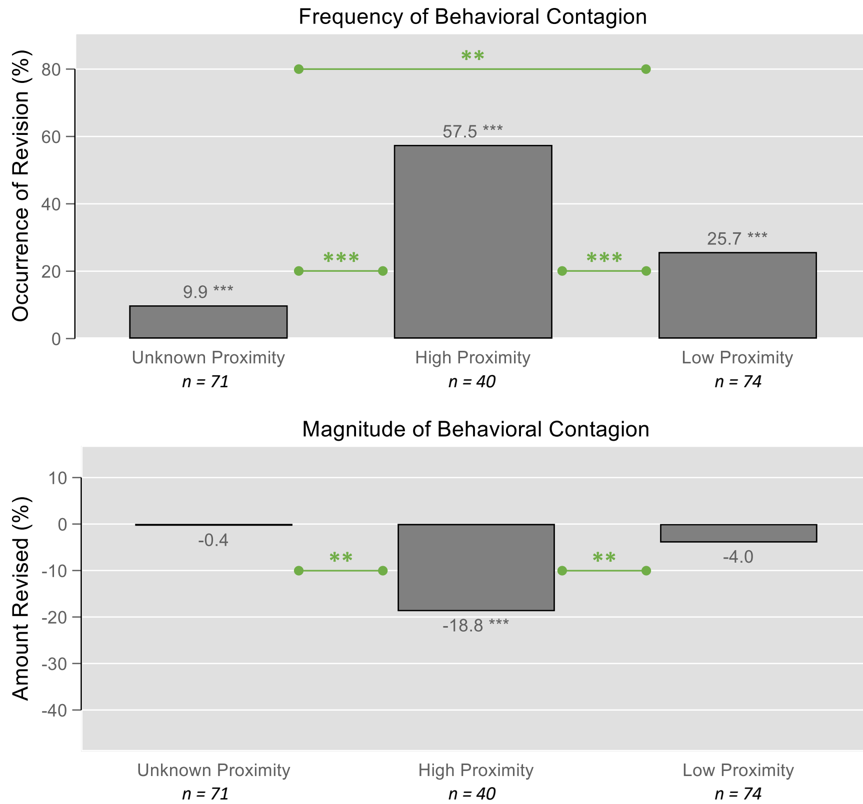


Figure 4: Frequency (upper section) and magnitude (lower section) of behavioral contagion conditional on social proximity to peer. Definitions of frequency and magnitude of behavioral contagion are identical to those used in Figure 2. Unknown / high / low proximity refers to the respective proximity treatment conditions. Green arrows and stars represent significant differences of compared means, while stars next to the mean number of each bar depict significant difference from zero. **p<0.05, ***p<0.01

3.3. Behavioral Contagion: Composite Effect of Peer Behavior and Social Proximity

Next, we combine the results from the previous two subsections and examine the composite effect of observed *behavior and social proximity* combined (see Figure 5).

For the frequency of contagion, we observe that behavioral contagion is asymmetric only where social proximity is high, with the difference being highly significant according to the equality of proportion statistic (88.2% vs. 34.8%, p<0.001). Contagion in the other proximity conditions is less frequent and not significantly different between the types of observed behavior. The results for the magnitude of behavioral contagion indicate that for both high and low proximity specifications the magnitude of anti-social contagion is larger than that of pro-social contagion, with the test statistics being significantly different at the 1% level (-51.3% vs. 5.3%, MWU, $p<0.001$) and 5% level (-10.0% vs. 1.5%, MWU, $p=0.013$), respectively. Once again, while our results indicate that behavioral contagion exists, we find that the effect is significantly larger and more frequent in the anti-social domain compared to the pro-social domain. In support of hypothesis H_2, we find that social proximity indeed amplifies contagion, particularly for anti-social behavior.

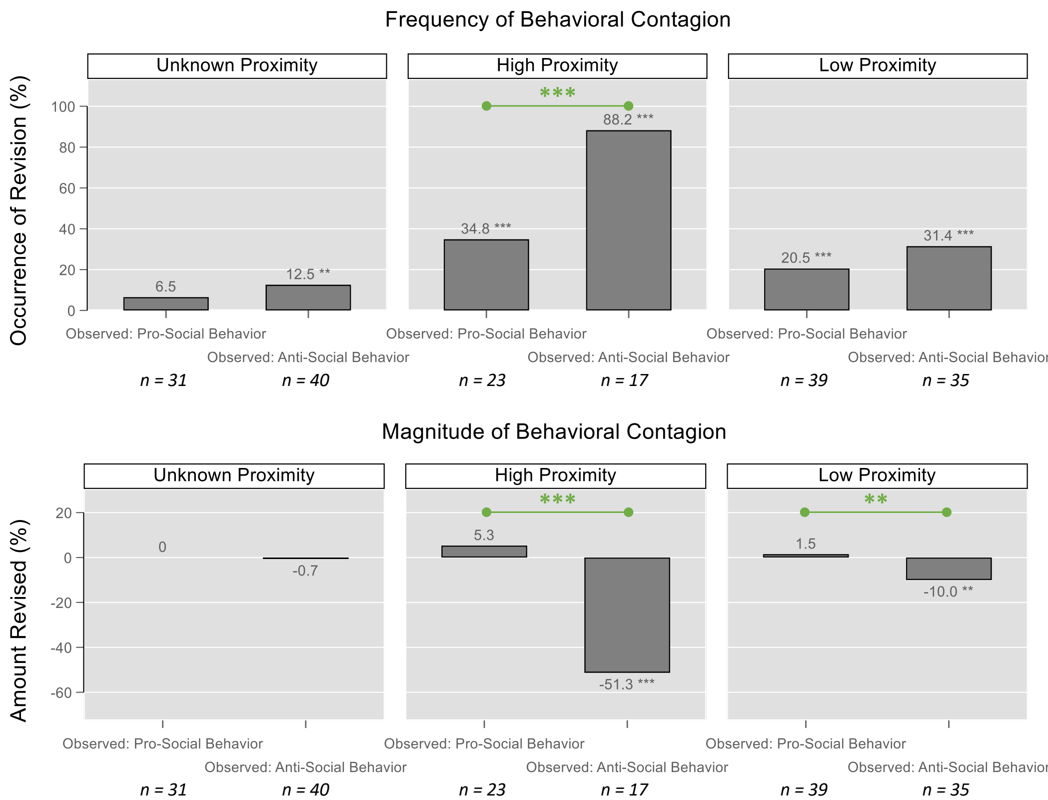


Figure 5: Frequency (upper section) and magnitude (lower section) of behavioral contagion conditional on social proximity to peer. Definitions of proximity as well as frequency and magnitude of behavioral contagion are identical to those used in previous figures. Green arrows and stars represent significant differences of compared means, while stars next to the mean number of each bar depict significant difference from zero. **$p<0.05$, ***$p<0.01$

3.4. Whom Does Behavioral Contagion Affect and How?

Lastly, we are interested in understanding the heterogenous effects of behavioral contagion conditional on the individual's initial behavior, i.e. whether the participants initially engaged in anti-social, pro-social, or equal split behavior. Based on one's initial behavior, our results suggest that behavioral contagion is mostly driven by those who engaged in anti-social behavior, both in frequency and in magnitude.

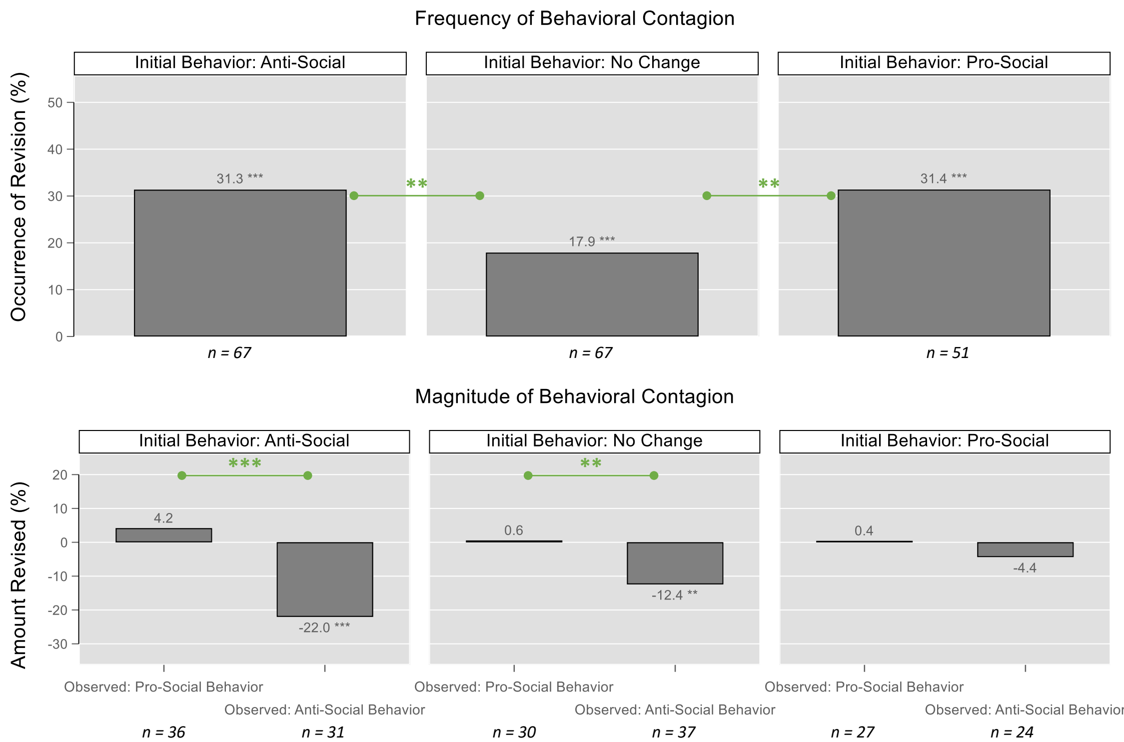


Figure 6: Frequency (upper section) and magnitude (lower section) of behavioral contagion conditional on initial anti- and pro-social behavior. Definitions of proximity as well as frequency and magnitude of behavioral contagion are identical to those used in previous figures. Green arrows and stars represent significant differences of compared means, while stars next to the mean number of each bar depict significant difference from zero. **p<0.05, ***p<0.01

As depicted in Figure 6, in terms of frequency, participants whose initial behavior was to maintain the equal split between themselves and the charity, express significantly less behavioral contagion than those who initially engaged in anti-social (17.9% vs. 31.3%, p=0.018) or pro-social behavior (17.9% vs. 31.4%, p=0.018). The frequency of contagion in the latter two cases is statistically indistinguishable.

The pattern is even more pronounced when looking at the magnitude of behavioral contagion. When participants initially engage in anti-social behavior, we observe the largest and most significant asymmetric spread of contagion, indicating that observing anti-social behavior is over five times (= |-22.0/4.2|) more contagious than observing pro-social behavior, with the difference being highly statistically significant (-22.0% vs. 4.2%, MWU, $p<0.001$). We observe a similar and highly significant pattern for the initial 'no change' participants (-12.4% vs. 0.6%, MWU, $p=0.035$), but not for participants who initially engaged in pro-social behavior (-4.4% vs. 0.4%, MWU, $p=0.429$).

3.5. Regression Analysis

Our previous findings support the notion that anti-social behavior is more contagious than pro-social behavior, which is amplified with proximity. We further investigate the robustness of our results using various regression specifications (Table 2) and find that our findings remain robust. We analyze the likelihood (Logit) and magnitude (OLS) of contagion with a core set of independent variables that are in accordance with existing theoretical and empirical research (social proximity and the observed anti- or pro-social behavior), an active participant's initial behavioral choice, the initial behavioral gap (amount taken from or given to charity) between oneself and the observed passive participant, and gender. We ensure robustness of our results by systematically adding a set of controls, including measures for greed, risk, gender, dummies for the questions used in the personality questionnaire, and relevant interactions that are coherent with our hypotheses.

In line with our previous results, we observe that behavioral contagion is both significantly more likely (model 1, 4.960, $p<0.01$) and also significantly stronger (model 4, -18.477, $p<0.001$) when observing anti-social behavior as compared to pro-social behavior. For the magnitude of contagion, the estimations also indicate that higher peer proximity triggers more behavioral contagion, leading to substantially more anti-social behavior overall (model 4, -23.891, $p<0.01$). Our results also show that the contagion of anti-social behavior is amplified with higher social proximity. This is true for the impact of high proximity compared to both unknown proximity (model 6, -57.158, $p<0.001$) and low proximity (model 6, -57.158 vs. -12.677, $p<0.01$). In addition, exposure to anti-social peer behavior begets more anti-social behavior, in particular for the initially anti-social individuals. For the frequency of behavioral contagion (model 2) our results indicate that initially anti-social individuals over-proportionally react to anti-social peers (50.025, $p<0.01$), also in comparison to the initially pro-social individuals (50.025 vs. 0.285, $p<0.01$).

All remaining controls are not robustly significant across our specifications and do not affect the main results. In sum, the regression results support our non-parametric results presented before. Both analyses indicate that anti-social behavior is more contagious especially when driven by social-proximity, and that initially anti-social individuals are

Table 2: Regression Analysis (Logit and OLS)

	Logit Specifications			OLS Specifications		
	DV: Frequency of Contagion (%)			*DV: Magnitude of Contagion (%)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Proximity *(Base Level: Unknown Proximity)*						
High Proximity	23.759***	16.252***	7.084**	-23.891***	-22.022***	4.071
	(16.461)	(10.107)	(6.375)	(7.668)	(7.286)	(6.790)
Low Proximity	4.357**	3.812**	3.920	-6.432	-5.397	2.294
	(2.611)	(2.039)	(3.374)	(4.150)	(3.550)	(4.747)
Observed Anti-Social Behavior	4.960***	1.473	2.327	-18.477***	-13.014**	0.134
	(2.582)	(1.045)	(2.081)	(4.205)	(6.514)	(4.208)
Initial Behavior *(Base Level: No Change of Fair Split)*						
Pro-Social	1.958	3.414	2.437*	7.401	6.766	3.799
	(1.206)	(2.707)	(1.249)	(6.166)	(6.997)	(4.955)
Anti-Social	1.536	0.201	1.836	-0.147	5.555	3.734
	(0.843)	(0.197)	(0.923)	(4.251)	(5.630)	(4.397)
Initial Behavioral Gap	0.997	1.014	0.997	0.010	-0.001	0.030
	(0.005)	(0.009)	(0.004)	(0.050)	(0.037)	(0.031)
Gender	1.671	1.034	0.960	-7.692	-7.615*	-7.845**
	(1.038)	(0.471)	(0.422)	(6.770)	(4.439)	(3.931)
Greed		0.858	0.948		1.265	-0.200
		(0.179)	(0.198)		(2.108)	(1.826)
Risk		0.759	0.782		0.796	1.734
		(0.168)	(0.168)		(1.835)	(1.655)
Interaction 1 *(Initial Behavior x Observed Behavior)*						
Pro-Social & Anti-Social		0.285			0.147	
		(0.350)			(11.665)	
Anti-Social & Anti-Social		50.025***			-10.845	
		(68.909)			(8.642)	
Interaction 2 *(Observed Behavior x Social Proximity)*						
Anti-Social & High Prox.			8.566*			-57.158***
			(11.004)			(13.605)
Anti-Social & Low Prox.			0.859			-12.677*
			(0.903)			(6.929)
Constant				17.998	8.833	-1.960
				(17.536)	(6.030)	(5.581)
Dummies Questionnaire	Yes	No	No	Yes	No	No
Adjusted R^2				0.096	0.153	0.299
Observations	172	172	172	172	172	172

Notes: Odds ratios (Logit) and regression coefficients (OLS). Standard errors in parentheses. Significances *** 1%; ** 5% level. For Logit, the dependent variable is a dummy that takes on the value 1 if behavioral contagion occurred (= behavior was revised into the direction of the behavior observed). For OLS, the dependent variable is the magnitude in behavioral contagion in percent (= revised amount). 'Proximity' is the degree of social proximity observed by the active participant and takes the value 1 (2) for high (low) proximity. 'Observed anti-social behavior' takes on the value 1 if observed behavior was anti-social. 'Initial behavior' is the active participant's initial decision with value 1 (2) if the participant initially took money away from (gave money to) charity. 'Initial behavioral gap' is the difference between the active participant's initial behavior and the observed behavior (in ECU). 'Gender' is a dummy with value 1 if participant was a male. 'Greed' (Hexaco) and 'Risk' (SOEP) are standardized scores where higher values indicate more greediness and higher risk seeking. 'Dummies Questionnaire' are dummies for the items used in the initial personality questionnaire used to generate the proximity score.

more susceptible to behavioral contagion, where they over-proportionally react to anti-social behavior compared to their pro-social peers.

Overall, we find conclusive evidence for behavioral contagion and support for our hypotheses, in particular with respect to the role of social proximity among peers (***H₂***). Interestingly, while we also find general support for hypothesis (***H₁***), unlike the predictions from our theoretical model and existing literature, we find behavioral contagion to be asymmetric towards anti-social behavior. [14] It is important to note that these results cannot be explained by a general tendency to become more anti-social when given the opportunity to revise one's own behavior that is independent of the information provided about peer behavior or social proximity (Thöni and Gächter, 2015).[15]

4. Discussion and Conclusion

Although peer effects have been extensively studied in different strategic contexts both in the field and in the lab, this paper aims to improve our understanding of whether, to which extent, and through which channels individuals are influenced by their peers to engage in more pro- or anti-social behavior. A wealth of literature points at methodological challenges in studying peer effects, including engodeneity, self-selection, and reflection problems (Manski, 1993, 2000; Angrist, 2014). We capitalize on a novel laboratory design that allows us to study behavioral contagion among peers along different dimensions of pro- and anti-sociality as well as social proximity. To this end, we introduce a variant of a give-or-take donation game (List, 2007; Bardsley, 2008), in which participants are able to

[14] We address potential experimenter demand concerns with transparency during the experiment and do not withhold information from the participants with regards to either the matching algorithm or the treatment variations. Participants are told that there is a random but equal chance that they would end up in the unknown, low, or high proximity treatment after being told about the details of all potential treatments. This transparency is a strong argument against confounds stemming from demand effects. This approach enables participants to interpret the presented social proximity information as the result of chance, rather than as additional information willingly implemented by the experimenter to trigger any particular behavior (for a discussion, see Zizzo, 2010).

[15] We solidify this assertion by running a follow-up experiment on Amazon Mechanical Turk (mTurk) with n=100 in which participants were given the opportunity to revise their initial behavior without observing any peer information. In this sample, only 7% decided to change their initial behavior in all three directions. The absence of *change of heart* in response to a revision opportunity without peer exposure is in line with Thöni and Gächter (2015) and thus solidify our main findings. Detailed results are available upon request. To ensure high quality data collection on mTurk, we utilize a combination of CAPTCHAs and sophisticated screening questions to avoid pool contamination. We applied the following restrictions to the participant pool: participants had to be in the U.S., approval rate greater than 99%, and could participate only once.

give money to or take money from a charity and has since been used in other experimental settings (e.g., Bolton et al., 2019). The main ingredients of our design that address the aforementioned challenges are the inclusion of a revision option following the exposure to peer behavior, as well as active and passive roles in which observed behavior is systematically held fixed and combined with variations in social proximity across treatments. Within this non-strategic environment, we are able to reduce the mechanism of behavioral change to the channel information dissemination about behavior of one's reference group, which is a first step to the exisitence, spread, and change of social norms (Bicchieri, 2006).

Overall, we find that anti-social behavior is more contagious than prosocial behavior, social proximity amplifies anti-social behavior in particular, and that initially anti-social individuals are most susceptible to peer exposure. For the study of peer-effects our results suggest that individuals are frequently at the tipping point of behaving anti-socially. Interestingly, since the nature of our implemented revision option retained full anonymity and thus contained no signaling value, our findings suggest that behavioral contagion occurs even if it merely contains self-signaling value.

While the results are surprising, we can reconcile these findings with existing theoretical and experimental literature. Pro-social behavior implies bearing costs in order to improve the well-being of others. Anti-social behavior often implies improving one's own well-being at the expense of a third party. Exemplary, this reasoning is coherent with findings on the asymmetry between positive and negative reciprocity (Offerman, 2002) and the increase and decrease of pro-social behavior in contributing to a public good (Croson and Shang, 2008; Chen et al., 2010). This assumption is also in line with the recent findings of Smerdon et al. (2016), Gächter et al. (2018), and Bicchieri et al. (2019a) on the role of bad behavior and the perseverance of bad norms. In addition, research also indicates that bad social norms can be particularly sticky and thus likely favor the asymmetry towards the contagion of anti-social behavior, as observed in our experiment (Kling et al., 2005; Dimant et al., 2015). Results by Charness et al. (2019) who theoretically and experimentally examine conformity in strategic sequential trust games are also in support of this finding and substantiate our results. Among other things, they find that individuals are more likely to imitate actions that are more self-interested than their own initial choice. Because a norm is a collective belief about expected behavior, such behavior is more informative if learned from someone who is more proximate (Bicchieri et al., 2019a).

In sum, our results highlight the importance of our contribution in this paper: Peer effects are not uniform, but rather strongly dependent on both the (anti-)sociality of observed behavior and the degree of social proximity to the observed peer. We treat these insights as novel in that they provide better understandings of peer effects and behavioral contagion. In real life, the appropriateness of anti-social behavior is thought to be more ambiguous than that of pro-social behavior, which is usually unambiguously appropriate. The nature of anti-social behavior implies the overstepping of social boundaries or the violation of laws. It can thus be expected that individuals often might want to engage in anti-social behavior, but seek social cues and signals to justify engaging in such behavior themselves: a crook among crooks is likely to be more appreciated than a crook among do-gooders. Our interpretation of these findings is that individuals look for (mental) excuses to be less selfish, but are more eager to embrace a peer signal that it is socially acceptable to be selfish (Bicchieri et al., 2019b). This, in turn, can facilitate the creation of destructive and socially inefficient norms, with costly enforcement mechanisms in place and without (Abbink et al., 2017; Bicchieri et al., 2018; Morsky and Akçay, 2019).

More generally, our findings are particularly consequential from a policy perspective in that they allow the identification of the target group that is most susceptible to behavioral change, which is particularly important for norm-based interventions (Madrian, 2014; Bicchieri and Dimant, 2019). Deviant behavior that benefits oneself at the expense of others is socially harmful and distorts welfare by leading to distortive, second-best solutions. At the same time, individual behavior does not occur in a vacuum but is the result of multilayered social interactions (Rogers et al., 2018). One concrete policy implication is that in addition to providing information about peer behavior like with norm-based policy instruments (Miller and Prentice, 2016; Hallsworth et al., 2017), making use of social proximity information as well as our insights on who is most susceptible to behavioral contagion can help to direct the influence of peer-effects towards the increase pro-social behavior (e.g., charitable giving) and reduction of anti-social behavior (e.g., smoking or criminal behavior). The latter is already done through the *Scared Straight* or the *G-MACC* programs where current or former criminals attempt to educate and deter at-risk youth with similar social background. We hope our results and methodological contribution inform more effective norm- and peer-based interventions and facilitate further research.

References

Abbink, K., Gangadharan, L., Handfield, T., and Thrasher, J. (2017). Peer punishment promotes enforcement of bad social norms. *Nature Communications*, 8(1):609.

Akerlof, G. A. (1997). Social distance and social decisions. *Econometrica: Journal of the Econometric Society*, pages 1005–1027.

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

Andreoni, J., Nikiforakis, N., and Siegenthaler, S. (2017). Social change and the conformity trap. Technical report, Mimeo.

Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30:98–108.

Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.

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

Benjamin, D. J., Choi, J. J., and Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4):1913–28.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300.

Bicchieri, C. (2006). *The grammar of society.* Cambridge University Press.

Bicchieri, C. and Dimant, E. (2019). Nudging with Care: The Risks and Benefits of Social Information. Working Paper Available at SSRN: https://ssrn.com/abstract=3319088.

Bicchieri, C., Dimant, E., Gaechter, S., and Nosenzo, D. (2019a). Social Proximity and the Evolution of Norm Compliance. Mimeo, University of Pennsylvania.

Bicchieri, C., Dimant, E., and Sonderegger, S. (2019b). It's Not A Lie If You Believe It. Lying and Belief Distortion Under Norm-Uncertainty. PPE Working Paper, University of Pennsylvania.

Bicchieri, C., Dimant, E., and Xiao, E. (2018). Deviant or Wrong? The Effects of Norm Information on the Efficacy of Punishment. PPE Working Paper 0016, University of Pennsylvania.

Bicchieri, C. and Xiao, E. (2009). Do the right thing: but only if others do so. *Journal of Behavioral Decision Making*, 22(2):191–208.

Bohnet, I. and Frey, B. S. (1999). Social distance and other-regarding behavior in dictator games: Comment. *American Economic Review*, 89(1):335–339.

Bolton, G., Dimant, E., and Schmidt, U. (2019). When a Nudge Backfires: Using Observation with Social and Economic Incentives to Promote Pro-Social Behavior. PPE Working Paper 0017, University of Pennsylvania.

Bradley, A., Lawrence, C., and Ferguson, E. (2018). Does observability affect prosociality? *Proceedings of the Royal Society B: Biological Sciences.*

Buckenmaier, J., Dimant, E., and Mittone, L. (2018). Effects of institutional history and leniency on collusive corruption and tax evasion. *Journal of Economic Behavior & Organization.*

Bursztyn, L., Ederer, F., Ferman, B., and Yuchtman, N. (2014). Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301.

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

Charness, G., Gneezy, U., and Halladay, B. (2016). Experimental methods: Pay one or pay all. *Journal of Economic Behavior & Organization*, 131:141–150.

Charness, G., Naef, M., and Sontuoso, A. (2019). Opportunistic conformism. *Journal of Economic Theory*, 180:100–134.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352.

Chen, Y., Harper, F. M., Konstan, J., and Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–98.

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

Cialdini, R. B., Reno, R. R., and Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6):1015.

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

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–32.

de Oliveira, A. C., Croson, R. T., and Eckel, C. (2015). One bad apple? heterogeneity and information in public good provision. *Experimental Economics*, 18(1):116–135.

Dimant, E. and Hyndman, K. B. (2019). Becoming friends or foes? how competitive environments shape altruistic preferences. Working Paper Available at SSRN: https://dx.doi.org/10.2139/ssrn.3326107.

Dimant, E., Krieger, T., and Redlin, M. (2015). A crook is a crook but is he still a crook abroad? on the effect of immigration on destination-country corruption. *German Economic Review*, 16(4):464–489.

Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *American Economic Review*, 101(5):1739–74.

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

Eckel, C. C. and Petrie, R. (2011). Face value. *American Economic Review*, 101(4):1497–1513.

Falk, A. and Fischbacher, U. (2002). "Crime" in the lab-detecting social interaction. *European Economic Review*, 46(4-5):859–869.

Falk, A. and Ichino, A. (2006). Clean evidence on peer effects. *Journal of Labor Economics*, 24(1):39–57.

Fatas, E., Heap, S. P. H., and Arjona, D. R. (2018). Preference conformism: An experiment. *European Economic Review*.

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

Fortin, B., Lacroix, G., and Villeval, M.-C. (2007). Tax evasion and social interactions. *Journal of Public Economics*, 91(11-12):2089–2112.

Frey, B. S. and Meier, S. (2004). Social comparisons and pro-social behavior: Testing conditional cooperation in a field experiment. *American Economic Review*, 94(5):1717–1722.

Gächter, S., Molleman, L., and Nosenzo, D. (2018). Peer effects in norm compliance. Technical report, Mimeo.

Gächter, S., Nosenzo, D., and Sefton, M. (2013). Peer effects in pro-social behavior: Social norms or social preferences? *Journal of the European Economic Association*, 11(3):548–573.

Gino, F., Ayal, S., and Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20(3):393–398.

Gino, F. and Moore, D. A. (2007). Effects of task difficulty on use of advice. *Journal of Behavioral Decision Making*, 20(1):21–35.

Gould, E. D. and Kaplan, T. R. (2011). Learning unethical practices from a co-worker: the peer effect of jose canseco. *Labour Economics*, 18(3):338–348.

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

Gross, J., Leib, M., Offerman, T., and Shalvi, S. (2018). Ethical free riding: When honest people find dishonest partners. *Psychological science*, page 0956797618796480.

Grout, P. A., Mitraille, S., and Sonderegger, S. (2015). The costs and benefits of coordinating with a different group. *Journal of Economic Theory*, 160:517–535.

Hallsworth, M., List, J. A., Metcalfe, R. D., and Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, 148:14–31.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972):1480–1484.

Herbst, D. and Mas, A. (2015). Peer effects on worker output in the laboratory generalize to the field. *Science*, 350(6260):545–549.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010). Matching and sorting in online dating. *American Economic Review*, 100(1):130–63.

Hugh-Jones, D., Ool, J., et al. (2017). Where do fairness preferences come from? norm transmission in a teen friendship network. Technical report, School of Economics, University of East Anglia, Norwich, UK.

Ichino, A. and Maggi, G. (2000). Work environment and individual background: Explaining regional shirking differentials in a large italian firm. *The Quarterly Journal of Economics*, 115(3):1057–1090.

Jones, S. R. (1984). *The economics of conformism.* Blackwell.

Kling, J. R., Ludwig, J., and Katz, L. F. (2005). Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment. *The Quarterly Journal of Economics*, 120(1):87–130.

Lefebvre, M., Pestieau, P., Riedl, A., and Villeval, M. C. (2015). Tax evasion and social information: an experiment in belgium, france, and the netherlands. *International Tax and Public Finance*, 22(3):401–425.

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

Madrian, B. C. (2014). Applying insights from behavioral economics to policy design. *Annu. Rev. Econ.*, 6(1):663–688.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, 14(3):115–136.

Mas, A. and Moretti, E. (2009). Peers at work. *American Economic Review*, 99(1):112–45.

Meer, J. (2011). Brother, can you spare a dime? peer pressure in charitable solicitation. *Journal of Public Economics*, 95(7-8):926–941.

Miller, D. T. and Prentice, D. A. (2016). Changing norms to change behavior. *Annual Review of Psychology*, 67:339–361.

Morsky, B. and Akçay, E. (2019). Evolution of social norms and correlated equilibria. *Proceedings of the National Academy of Sciences*, page 201817095.

Offerman, T. (2002). Hurting hurts more than helping helps. *European Economic Review*, 46(8):1423–1437.

Rogers, T., Goldstein, N. J., and Fox, C. R. (2018). social mobilization. *Annual review of psychology*, 69:357–381.

Shang, J. and Croson, R. (2008). Field experiments in charitable contribution: The impact of social influence on the voluntary provision of public goods. *The Economic Journal*, 119(540):1422–1439.

Smerdon, D., Offerman, T., and Gneezy, U. (2016). Everybody's doing it: On the emergence and persistence of bad social norms. Technical report, Tinbergen Institute Discussion Paper.

Smith, S., Windmeijer, F., and Wright, E. (2015). Peer effects in charitable giving: Evidence from the (running) field. *The Economic Journal*, 125(585):1053–1071.

Thöni, C. and Gächter, S. (2015). Peer effects and social preferences in voluntary cooperation: A theoretical and experimental analysis. *Journal of Economic Psychology*, 48:72–88.

Wheeler, L. (1966). Toward a theory of behavioral contagion. *Psychological Review*, 73(2):179.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

Appendix

Section A: Theory

Intuitively, the behavior of socially closer people exhibits a stronger influence on one's decision making than the behavior of random strangers. The purpose of this section is to outline a simple extension of the quadratic utility model by Akerlof (1997). The intuition behind Akerlof's approach is that individual decision-makers face a trade-off between their own behavioral preferences and the preferences of peers. Thus, the individual's utility is subject to a relative evaluation of her own behavior and the behavior of peers, which is a function of the proximity between the pair.

Our experimental design is a two-stage give-or-take donation game with a revision option, in which participants are paired with a charity. Both the individual and the charity start with an initial endowment $I_c = I_i = I \in R^+$ of equal size. At each stage $t \in [1, 2]$, with t_1 (t_2) corresponding to Stage 1 - Action (Stage 3 - Reaction) in our experiment, each individual i faces the choice, x_t^i, of either (a) donating part or all of her own money to the charity, (b) retaining the equal split, or (c) taking away part or all of the charity's money and adding it to her own income. We will refer to (a) and (c) as pro-social or anti-social behavior, respectively. Naturally, the individual's decision is of the form $x_t^i \in [-I, +I]$. The only difference between both stages is the information set that the individual possesses about her peer's behavior. In t_2 (after the observation of a peer's behavior), the individual is given the opportunity to revise her initial decision, if desired. The experiment concludes after this revision decision, which the participant is not aware of at the time of the decision.[16]

Let $\alpha_t^{ij} \in (0, 1)$ depict the social proximity of an individual at time t, and θ^i describe the individual's inherent attitude towards pro-/anti-social behavior. That is, θ^i represents the individual i's preference to give or take a particular monetary amount within the boundaries of one's income in a given situation, thus being defined as $\theta^i \in [-I, +I]$. In the spirit of Akerlof (1997), individual utility is declining with increasing distance between one's own behavior and the peer's behavior and can be represented as the following trade-off:

$$\max_{x_t^i} U_t^i = I - (1 - \alpha_t^{ij})(x_t^i - \theta^i)^2 - \alpha_t^{ij}(x_t^i - x^j)^2 \tag{1}$$

[16] In order to make the decisions in t_1 and t_2 as comparable as possible, we deliberately avoided providing participants with detailed information about the exact number of stages in the experiment.

Following our experimental design, peer behavior is not observed at t_1. Hence, in the absence of any other information except own preferences, an individual will mould his expectations of the preferences of others on his own, $x_t^i - x_t^j$, which is in line with literature on the value of information (such as Grout et al. 2015). From this follows directly that the behavior in stage 1 simply reflects one's inherent preference $x_t^i = \theta^i$.

Revision After Peer Observation

In order to account for Manski's (1993) reflection problem in studying the magnitude of behavioral contagion in t_2, one has to hold individual j's behavior (passive peer) from t_1 constant while giving individual i (active peer) the ability to revise her initial decision. In our setup, active players receive a signal about one peer's behavior between t_1 and t_2. Consequently, individuals now face a trade-off decision in which deviation from the individual inherent characteristic θ^i has to be weighed against deviating from the observed peer behavior x_j^i. Individuals maximize:

$$\frac{\partial U_t^i}{\partial x_t^i} = 2\alpha_t^{ij}(x_1^j - x_t^i) - 2(\alpha_t^{ij} - 1)(\theta^i - x_t^i) \Leftrightarrow x^* = x_t^i = \theta^i - \alpha_t^{ij}\theta^i + \alpha_t^{ij}x_1^j \qquad (2)$$

We define a *Contagion Gap* as the difference in behavior between the observer and observee after revision compared to the difference in behavior prior to peer observation, formally $|x_1^i - x_1^j| - |x_2^i - x_1^j|$. We distinguish two cases.

First, suppose $\theta^i > x^j$. This implies for the contagion gap:

$$|\theta^i - x_1^j| - |x^* - x_1^j| = \theta^i - x^* = \theta^i - \theta^i + \alpha_t^{ij}\theta^i - \alpha_t^{ij}x_1^j = \alpha_t^{ij}(\theta^i - x_1^j) > 0. \qquad (3)$$

Second, suppose $\theta^i < x^j$. Then

$$|\theta^i - x_1^j| - |x^* - x_1^j| = x^* - \theta^i = \theta^i - \alpha_t^{ij}\theta^i + \alpha_t^{ij}x_1^j - \theta^i = \alpha_t^{ij}(x_1^j - \theta^i) > 0. \qquad (4)$$

Thus, in both cases we have $|x_1^i - x_1^j| > |x_2^i - x_1^j|$, that is, the contagion gap is reduced. We can easily infer that the larger (smaller) the α, the closer (further away) x_2^i is to (from) x_1^j and hence the smaller (larger) the contagion gap.[17]

[17] Observe that, unlike in Akerlof's (1997) general conformity model, this approach generates a unique prediction due to restrictions put on the social proximity parameter in the linear reaction function of the form $0 < \alpha_t^{ij} < 1$.

Section B: Additional Figures and Robustness Checks

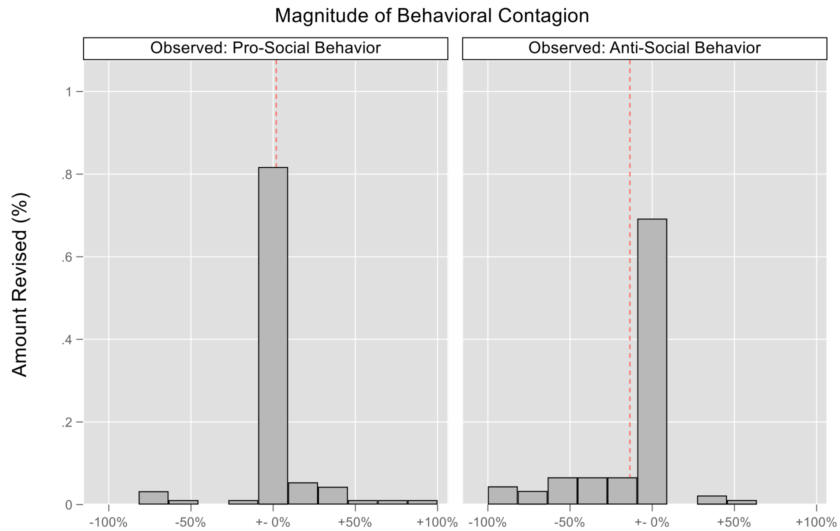


Figure A1: Histogram representation of behavior in Figure 3

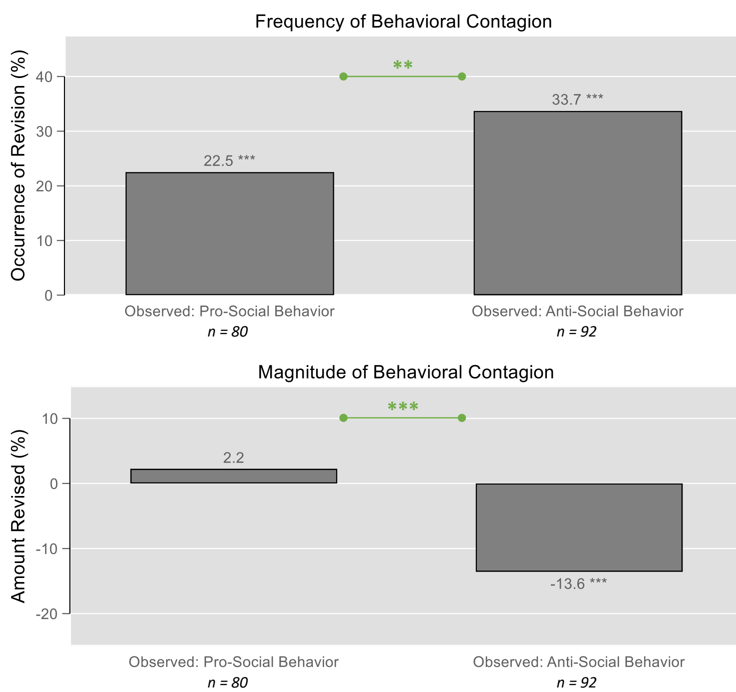


Figure A2: Robustness of Figure 3 (exclusion of participants who observed same behavior as their own)

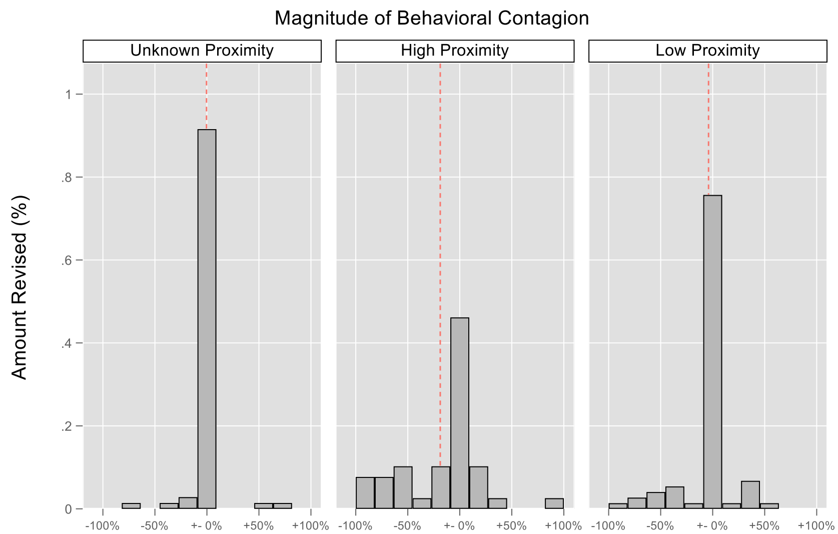


Figure A4: Histogram representation of behavior in Figure 4

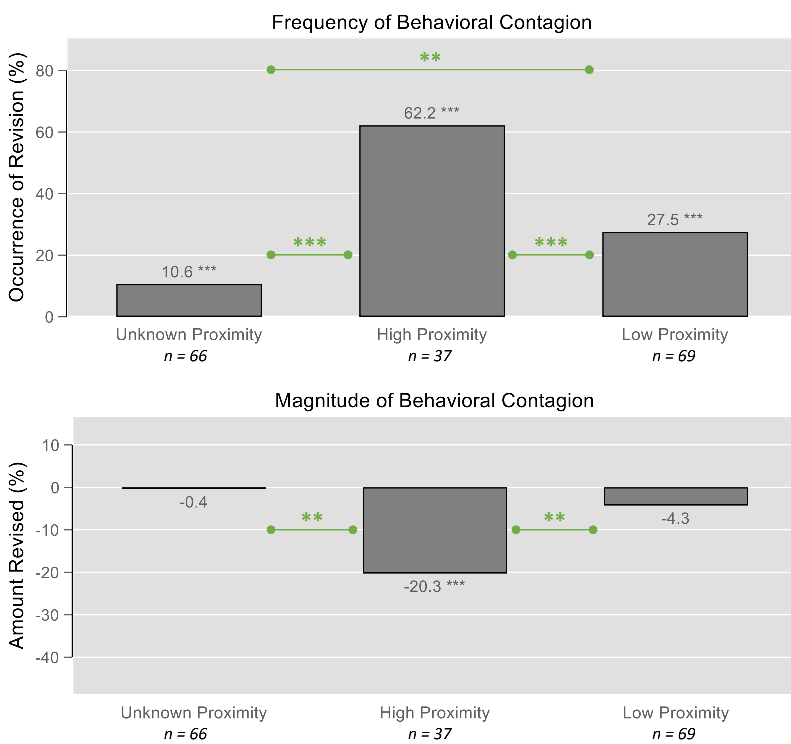


Figure A5: Robustness of Figure 4 (exclusion of participants who observed same behavior as their own)

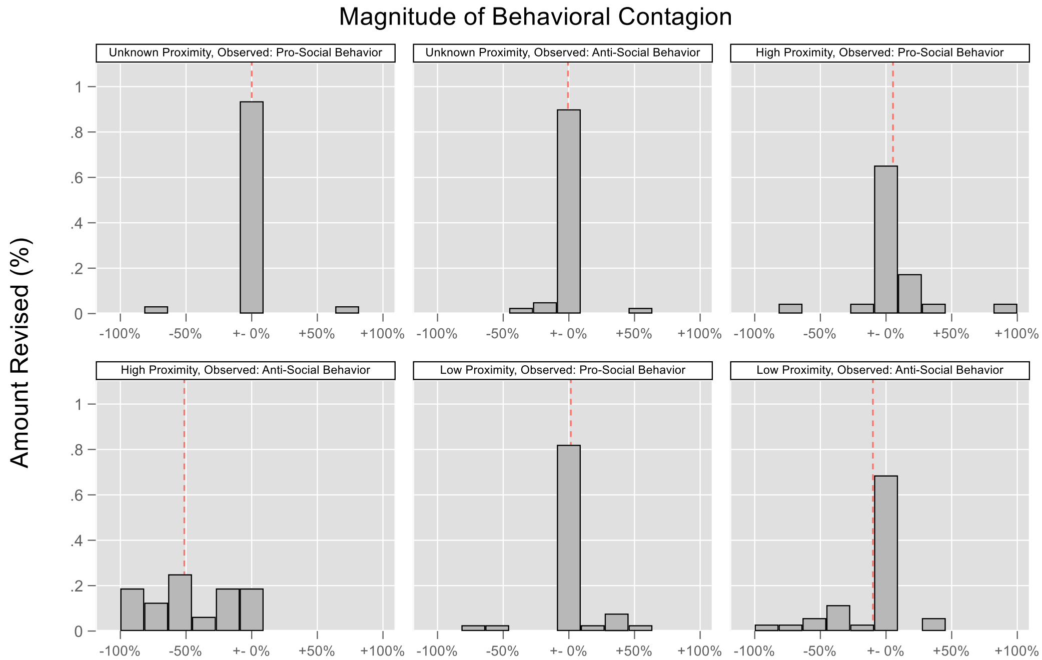


Figure A6: Histogram representation of behavior in Figure 5

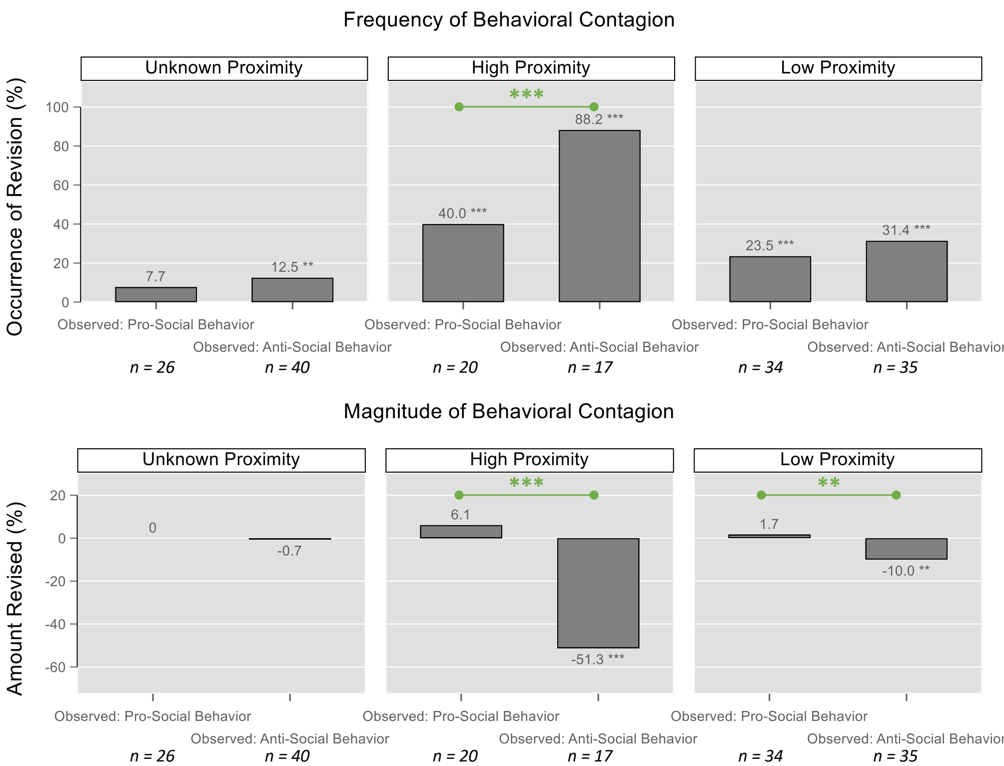


Figure A7: Robustness of Figure 5 (exclusion of participants who observed same behavior as their own)

Section C: Additional Material

C1: Instructions

General Information on the Experiment

- First of all, we would like to thank you very much for participating in this experiment. Please read the instructions carefully. The experiment will last for about 45-60 minutes.

- During the entire experiment, no communication is allowed. If there is something you do not understand or if you have any questions, now or at some point during the experiment, please raise your hand and remain seated. One of our colleagues will come to you and answer your question.

- During the experiment, you have the possibility to earn money. The amount you will receive at the end of the session depends on how many Experimental Currency Units (ECU) you earn during the experiment.

- At the end of the experiment, the amount of ECU that you have earned will be converted into real money at an exchange rate of 20 ECU = 1 Euro.

- All decisions you make during this experiment will remain anonymous. None of the participants gets to know the identity of other participants in the experiment and decisions cannot be linked to a specific participant. Moreover, you will be paid anonymously at the end of the experiment.

Order of Events:

- The experiment consists of a list of statements that you will receive at the beginning and further decisions. Explanations and information related to these decisions will be given as the experiment progresses. You will make these decisions once.

- Both you and the charitable organization of your choice (i.e. an officially registered charity organization) will be provisionally assigned a monetary amount of 300 ECU each.

- During the experiment you will have to decide on whether you want to:
  1. take a part or all of the money from the charitable organization.

2. leave the division of the sum of money as it is.
3. give a part or all of your money to the charitable organization.

- In case you decide to take money from the charitable organization, the respective amount of money will be transferred to your individual cash account and exactly the same amount will be deducted from the cash account of the charitable organization.

- Should you decide to give money to the charitable organization of your choice, the respective amount of money will be deducted from your individual cash account and given to the charity. The experimenter will double all ECUs remaining in the charity's account at the end of the experiment.

- Your decision remains anonymous and neither the other participants of the experiment nor the experimenters have the possibility to assign your choices to your identity.

- At the end of the experiment, one participant will be chosen at random and his or her choice will be implemented and count towards the charity (i.e. that choice will be relevant for the payment). The receipt of this donation will be published on the homepage of the BaER-Lab (http://www.baer-lab.org) in a timely manner. All other participants will receive 150 ECU (including the show-up fee) at the end of the experiment.

The total payoff of the participants:

- **In case you are the randomly chosen participant**
  300 ECU +/- (the amount of money that has been given to/taken from the cash account of the charitable organization)

- **In case you are not the randomly chosen participant**
  150 ECU

The total payoff of the charitable organization:
Amount of money in the cash account of the charitable organization of the randomly chosen participant

- At the end of the experiment, the relevant information on the payment will be made visible to each participant on his or her screen.
- After the actual experiment concludes, we will ask you to fill out a questionnaire. Please fill out the questionnaire carefully and truthfully.

C2: Screenshots of Decision Screens (indicating what participants have seen on the screen)

1. **List of statements: generates the proximity measure in all treatments**

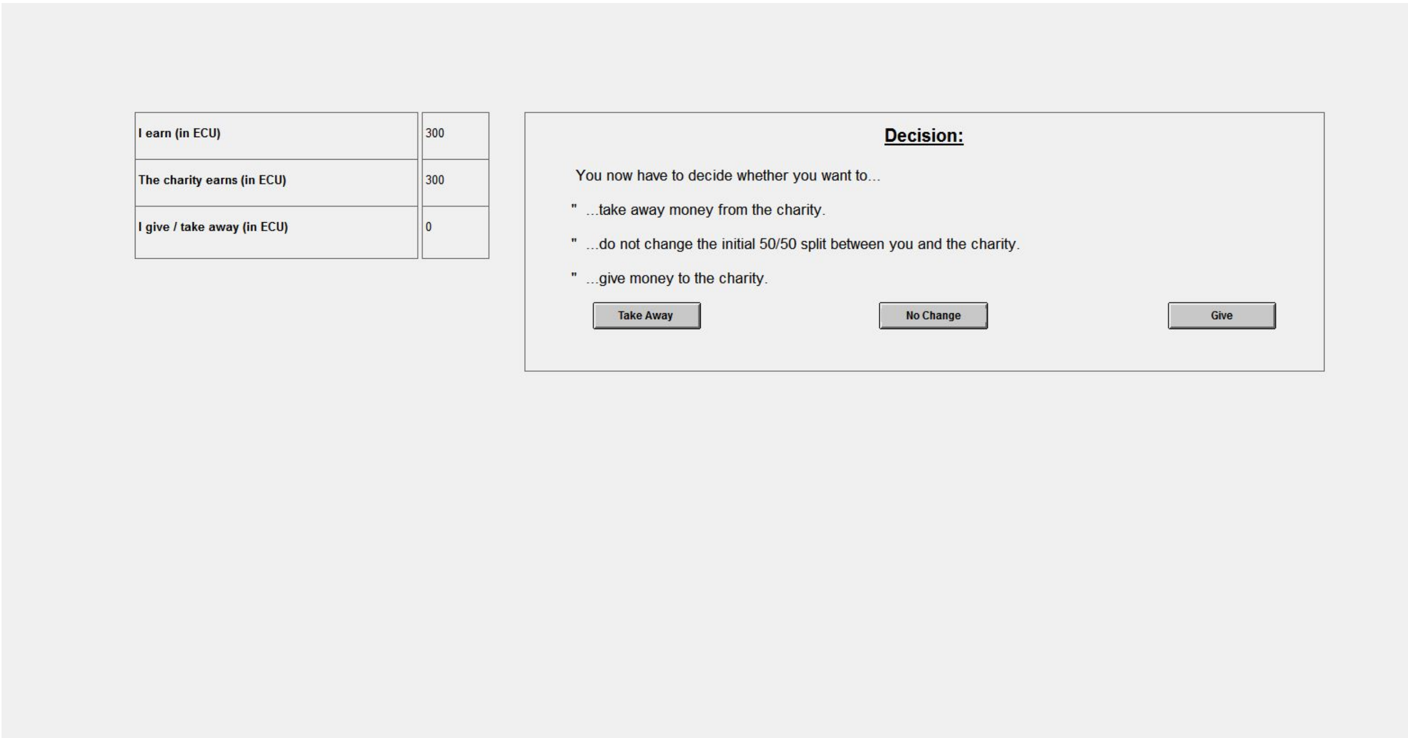
Information

- In this phase you receive a list that consists of **25 statements**.
- Please take your time to read carefully through all the statements. Please choose only those statements that apply to you.
- You can choose **as many statements** as you want to. The amount of statments chosen will **not** affect your payments in the experiment.

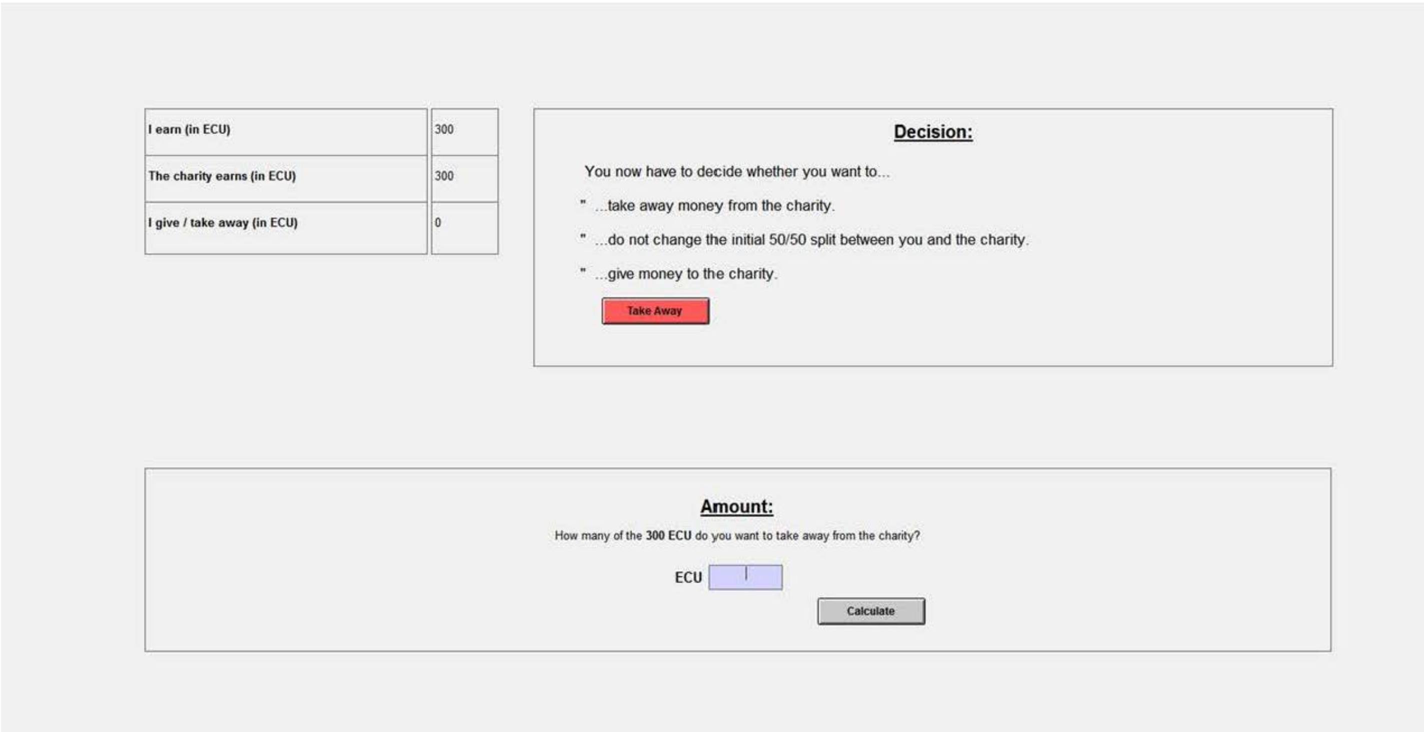
Statement 1:	I am a reliable person.	☐
Statement 2:	I am interested in politics and/or economics.	☐
Statement 3:	Money is important to me.	☐
Statement 4:	I am an honest and sincere person.	☐
Statement 5:	I am a cinephile.	☐
Statement 6:	I am interested in sports.	☐
Statement 7:	I am a religious person and faith is important to me.	☐
Statement 8:	I am fond of animals.	☐
Statement 9:	I am interested in art and/or cultures.	☐
Statement 10:	I am an active and adventurous person.	☐
Statement 11:	I am interested in cars and/or technology.	☐
Statement 12:	I am fond of children and family-oriented.	☐
Statement 13:	I am interested in foreign languages and/or countries.	☐
Statement 14:	I am a warmhearted and helpful person.	☐
Statement 15:	I am a tolerant person.	☐
Statement 16:	I like to gossip.	☐
Statement 17:	I am a faithful person.	☐
Statement 18:	I play an instrument.	☐
Statement 19:	I like to go out and dance.	☐
Statement 20:	I am a goal-oriented person.	☐
Statement 21:	I spend a lot of time in front of the TV.	☐
Statement 22:	I am a sociable person and like to be among people.	☐
Statement 23:	I like to play videogames.	☐
Statement 24:	I am a humorous and entertaining person.	☐
Statement 25:	I am a strong-willed person.	☐

Continue

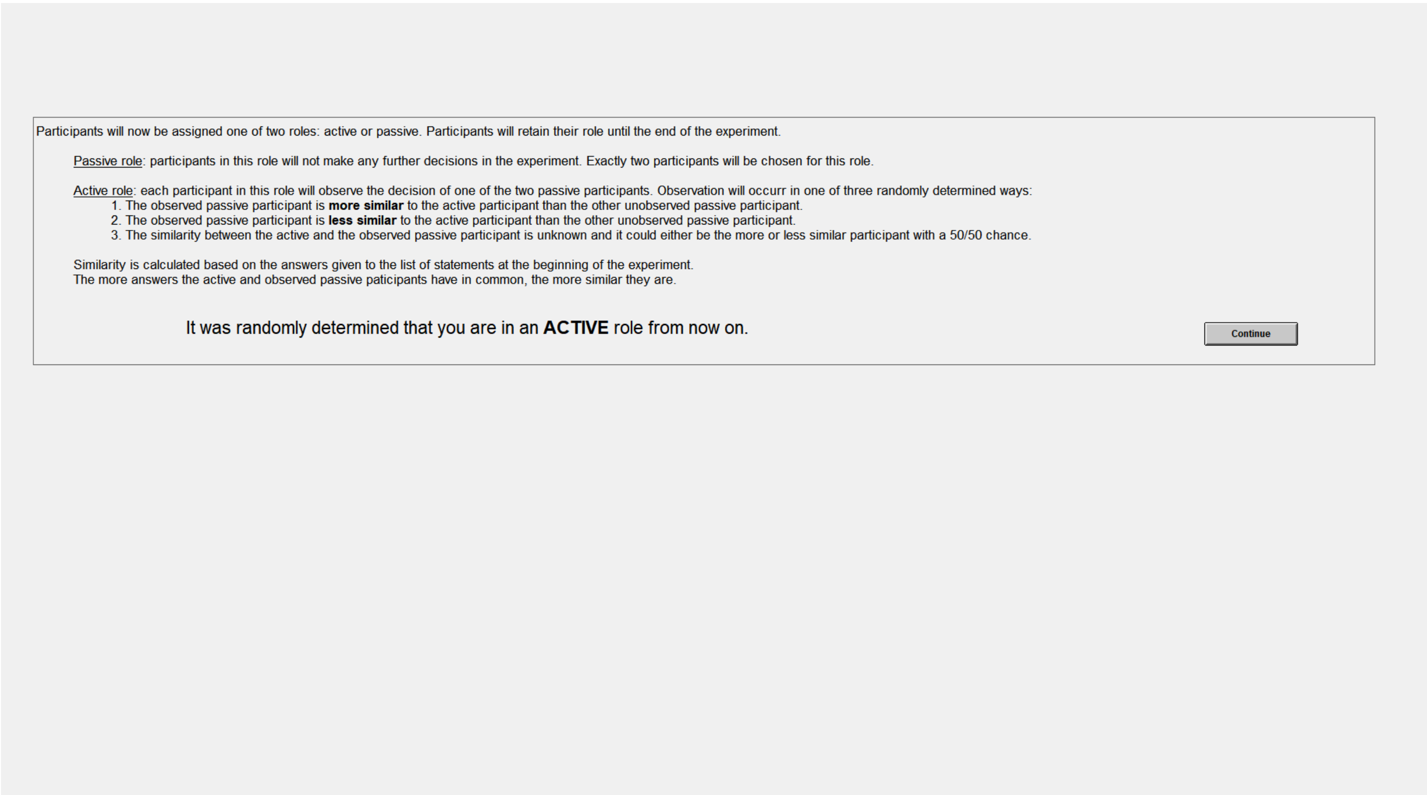
2. **First decision towards charity**



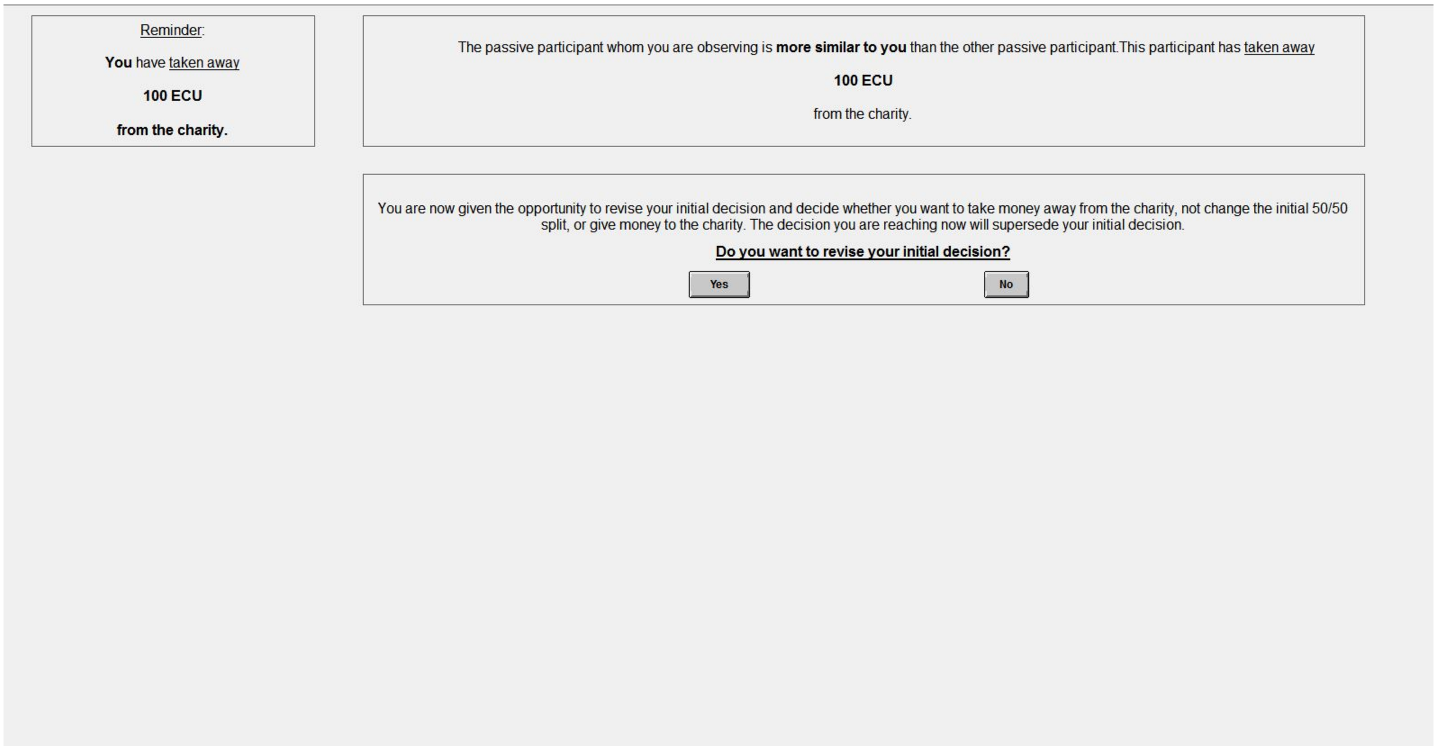
(Exemplary for the taking away decision)



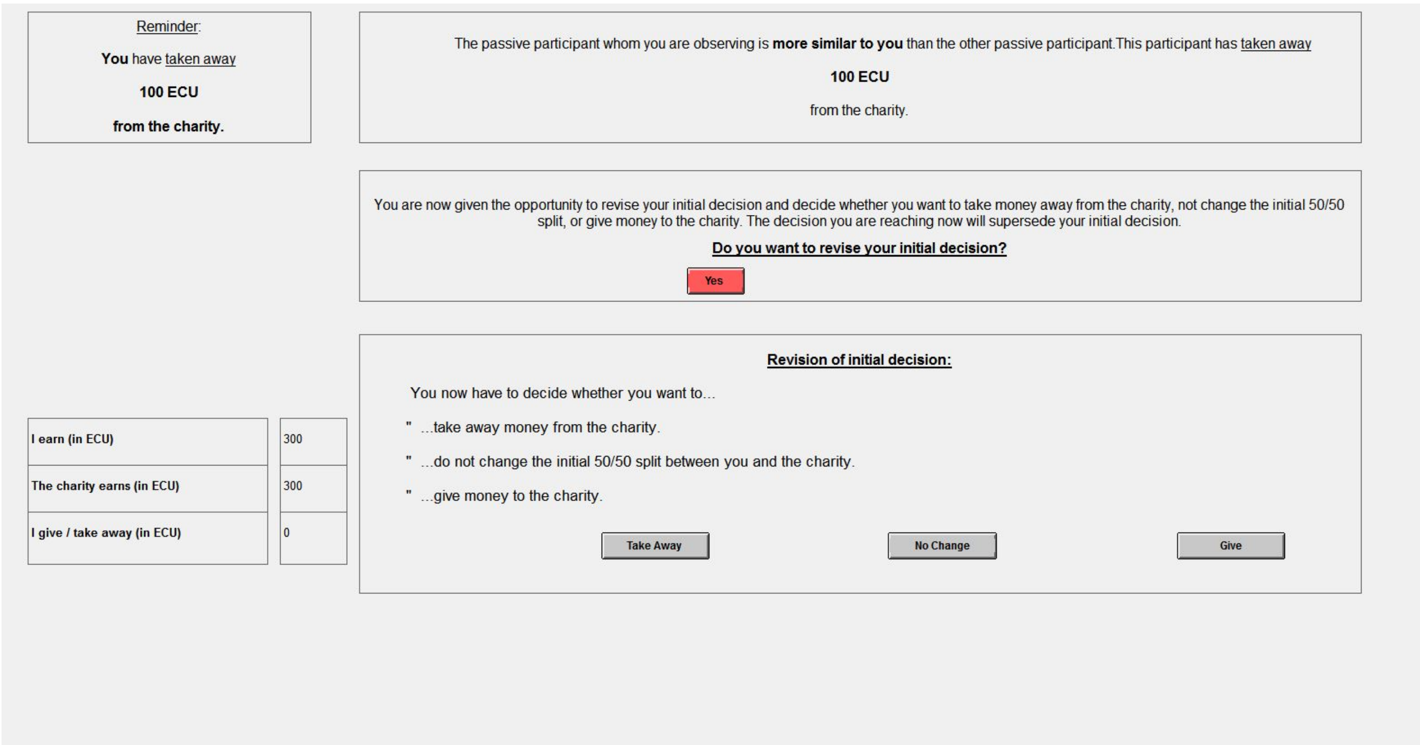
3. **Assignment to either the passive or active role and explanation of proximity**



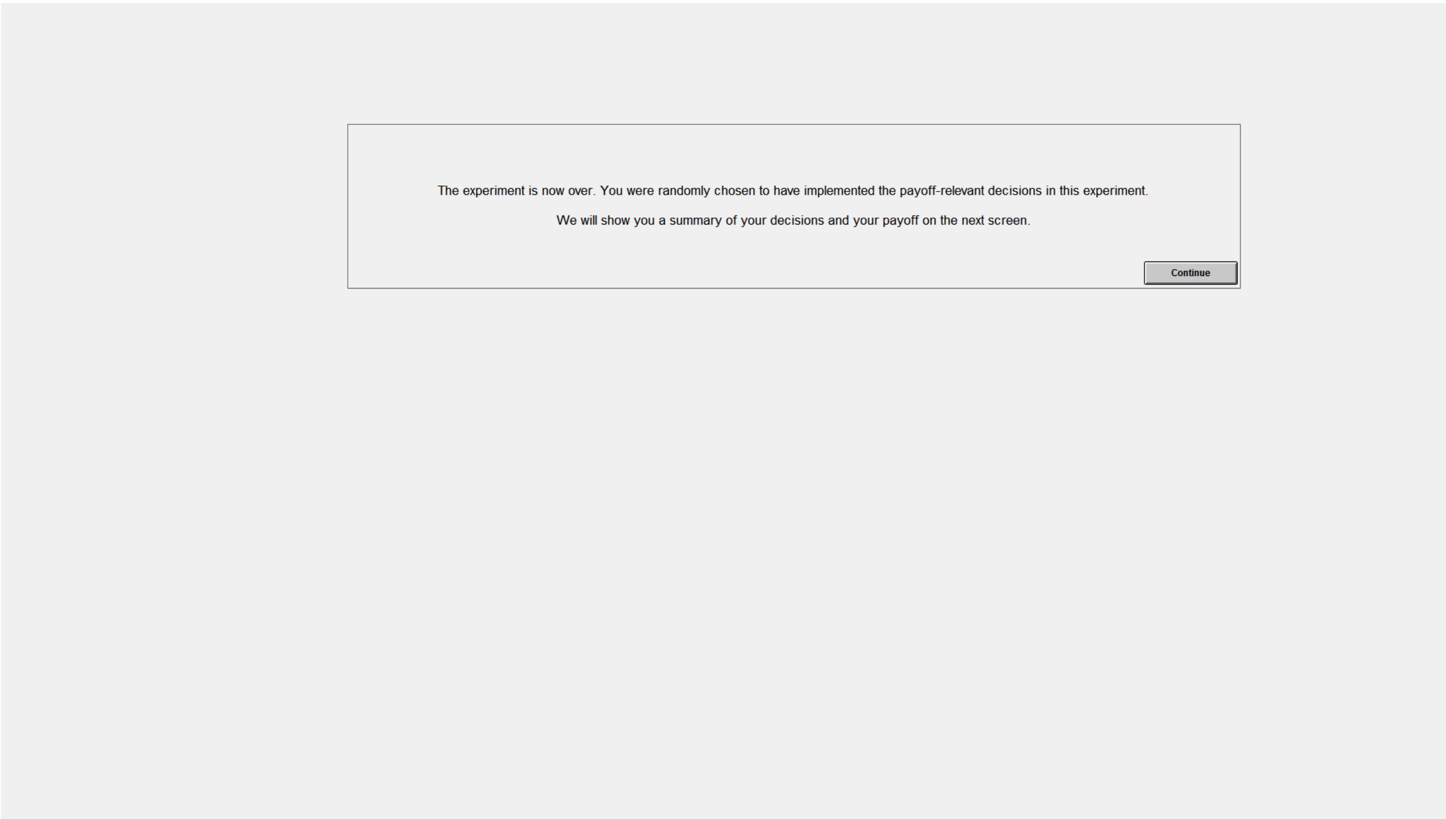
4. **Observation of one's peer and potential revision of one's initial decision**
Important: treatments vary by the information on the social proximity measure (first box): unknown similarity (Baseline), more similar (Treatment 1), less similar (Treatment 2).



(Exemplary for the taking away decision)



5. End of the experiment



6. Final payoff screen