

The most recent version of the working paper can always be downloaded following [this link](#).

Observability, Social Proximity, and the Erosion of Norm Compliance

Cristina Bicchieri¹, Eugen Dimant^{1,**}, Simon Gächter^{2,4,5}, Daniele Nosenzo^{2,3}

¹ *University of Pennsylvania and Behavioral Ethics Lab*

² *University of Nottingham*

³ *Luxembourg Institute of Socio-Economic Research (LISER)*

⁴ *CESifo, Munich*

⁵ *IZA, Bonn*

Abstract

Social norms are ubiquitous in social and economic life but the drivers of norm conformity are poorly understood. We study the specific ways in which others' norm compliance influences own norm compliance. Our context is a repeated non-strategic Take-or-Give donation experiment in which we show that giving is considered socially appropriate while taking is inappropriate. We find that observation of norm violations by anonymous others strongly erodes own norm compliance. Most importantly, erosion is halted when people have even minimal social proximity to those they observe; in this case, individuals also pay attention to norm followers. Our results highlight the importance of social proximity for the dynamics of social norm compliance.

Keywords: Norm Erosion, Norm Compliance, Social Norms, Social Identity

JEL: C92, D64, D9

*We would like to thank Johannes Abeler, Roland Bénabou, Gary Bolton, Damon Centola, Stefano Fiorin, Bernd Irlenbusch, Erin Krupka, Michel Maréchal, David Rand, Al Roth, Shaul Shalvi, Dirk Sliwka, Bertil Tungodden, and Elke Weber for helpful comments and suggestions. We also appreciate discussions with attendees at the 2017 Norms and Behavioral Change Workshop, 2017 Science of Philanthropy Initiative (SPI) Conference at the University of Chicago, 2018 North American Economic Science Association (ESA) Meeting, 2018 Cultural Evolution Society Conference, 2018 Maastricht Behavioral and Experimental Economics Symposium, Workshop on Experiments in Cooperation and Norms (WECAN), the 3rd Cultural Transmission and Social Norms Workshop at MIT, as well as seminar participants at the Behavioral Insights Team UK, the Choice Lab at Bergen, George Mason University, Kiel Institute for the World Economy, the Lab@DC Behavioral Research Team, London School of Economics, Stockholm School of Economics, Wharton Business School, and the Universities of Amsterdam, Cologne, Lyon, Oxford, Princeton, and Yale.

**Corresponding author

Email addresses: cb36@sas.upenn.edu (Cristina Bicchieri), edimant@sas.upenn.edu (Eugen Dimant), Simon.Gaechter@nottingham.ac.uk (Simon Gächter), Daniele.Nosenzo@liser.lu (Daniele Nosenzo)

This version: October 19, 2019

1. Introduction

Social norms guide our behavior and interactions in a variety of economically interesting domains. They constrain anti-social acts, like corruption (Fisman and Miguel, 2007; Muthukrishna et al., 2017), tax evasion (Fellner et al., 2013; Hallsworth et al., 2017; Bott et al., 2019), abuse of the welfare state (Lindbeck et al., 1999), cheating (Kocher et al., 2017; Bicchieri et al., 2019a), youth delinquency (Kling et al., 2005; Chyn, 2018), and discrimination (Barr et al., 2018). Social norms also support pro-social behaviors, such as collective actions (Ostrom, 2000), environmental conservation (Goldstein et al., 2008), voter turnout (Knack, 1992; Gerber et al., 2008), charitable giving and altruistic sharing (DellaVigna et al., 2012; Krupka and Weber, 2013; Krupka and Croson, 2016; Gächter et al., 2017; Bašić et al., 2019), and social preferences (Gächter et al., 2013; Kimbrough and Vostroknutov, 2016; Bicchieri et al., 2019b).

Despite progress in understanding the importance of social norms in many economically interesting settings, the economics literature lacks a detailed understanding of the factors that influence individuals' willingness to *comply* with social norms (for a recent discussion see Bicchieri and Dimant, 2019). This is important because norm compliance can quickly erode in response to changes in the social, physical or political landscape (e.g., Goldin and Katz, 2002; Keizer et al., 2008; Centola, 2018; Bursztyn et al., 2019).

In this paper, we focus on two potential sources of erosion of norm compliance: observations of other people's behavior and the degree of social proximity one has with those whose behaviour they observe. Our study is guided by the conceptual framework of social norms based on preferences and beliefs developed by Bicchieri (2006, 2016) and relies on a novel experimental design conducted in a controlled laboratory experiment. Expanding the tradition of List (2007) and Bardsley (2008), we introduce a novel repeated Take-or-Give (ToG) donation game in a non-strategic group setting in which decision-makers can either give money to a charity, take money from it, or retain the initial equal split. We ran two types of experiments: one that establishes the existence of a social norm in our setting and another that allows us to see behaviour in the ToG donation game over time under three different information conditions. The ToG game allows us to study both active norm-following (i.e., by giving money to a charity, or retaining the initial equal split) and active norm-breaking (by taking money from charity) in absence of strategic elements as present related games such as Public Goods Games or Ultimatum Games (e.g., Bohnet

and Zeckhauser, 2004; Neugebauer et al., 2009). We established the *existence* of a norm in our setting by applying the tools introduced by Bicchieri and Chavez (2010) to measure subjects' empirical expectations, normative expectations, and personal normative beliefs about donating to the charity. All three measures provide strong evidence that players consider charitable giving to be a social norm and taking to be a norm violation.

We study *behavior* in the ToG donation game with a separate group of participants. Subjects repeatedly play the ToG donation game by making independent decisions across 20 periods. Between treatments, we systematically vary whether players can or cannot observe whether other decision-makers have given or taken money from a charity in previous periods. In the baseline treatment, 'NoObservation', players receive no feedback about others' ToG decision. In the 'Observation' treatment, players are informed after each round about the ToG decisions of their group members. In our final treatment, 'ObservationSP', group members can observe others' ToG decisions after each round but also receive information about their social proximity to the other group members on one particular and fairly minimal dimension that does not permit to further identify players (fandom for a local sports team, explained in detail below). Note that in all our treatments our design implements the decisions of others in a *non-strategic* and *anonymous* way, so that there is no monetary or strategic impact from learning others' behaviour.

We find that in the NoObservation treatment, donation rates to the charity are stable over time. This is an important benchmark result because it shows that increasing experience with the ToG donation game does not *per se* erode norm compliance. By stark contrast, donations in Observation decline by about 20% compared to donations in NoObservation. This decline is due to the asymmetric impact of observing compliance with (and violation of) the norm of giving: individuals strongly reduce the amount they donate to charity when they observe that others take money from the charity, yet they do not increase the donated amount when they observe that others give to charity. Knowledge of social proximity strongly moderates ToG donation behavior. In ObservationSP, the average donation behavior is not significantly different from our NoObservation baseline. The reason is that in socially proximate groups, individuals respond to observed norm compliance as well as observed norm violations. In the presence of known social proximity, both forms of behavior are contagious and stabilize donations roughly at their initial level.

Finally, we show that there is heterogeneity in the extent to which subjects respond to feedback about others' behavior. The biggest changes in behavior come from subjects who are norm-compliant and give to charity before being exposed to others' negative be-

havior. In contrast, subjects who initially take from charity, and thus violate the norm, do not respond to positive or negative feedback about others' behavior. The response from subjects who initially abstain falls between these two groups. We do not find a subgroup of individuals for whom feedback about others' behavior has an aggregate positive effect.

Our paper makes five novel contributions to a deeper understanding of social norm compliance. First, while most of the existing literature focuses on showing the existence of norms, we use a framework that enables us to distinguish between the *existence* of norms and the extent to which they are *followed*. Instead of using a definition of norms that relies only on normative expectations (beliefs about what others think is appropriate behavior) (Gächter et al., 2013; Krupka and Weber, 2013; Fehr and Schurtenberger, 2018), we focus on a richer definition that includes empirical expectations (beliefs about others' behavior) and personal normative beliefs (personal beliefs about appropriate behavior) as well (Bicchieri, 2006, 2016). This definition allows us to better focus on the drivers and dynamics of norm compliance, where empirical expectations play a crucial role.

Second, we compare the dynamics of norm compliance in a setting where individuals receive continuous feedback on others' behavior (treatment Observation) to one in which no such feedback is provided (baseline NoObservation). Our results suggest that the participants' updated empirical expectations play a key role in the dynamics of norm compliance. This result is reminiscent of behavior in a repeated public good game, where contributions do not decline without feedback about others' contributions, but do decline with such feedback (Neugebauer et al., 2009). Crucially, our results reveal that observing others behavior has a causal influence on the observer's behavior in a *non-strategic* setting.

Third, we show that in the Observation treatment the response to observed behavior is asymmetric: while individuals respond strongly to others' norm violations, the response to examples of norm compliance is weak. As a consequence, we observe a marked erosion of norm compliance over time when information about others' behavior is freely available. This finding extends recent work on peer effects in norm compliance, which also shows (in static, one-shot settings, in contrast with our dynamic, repeated setting) that individuals mainly respond to examples of norm violations, much less to examples of norm compliance (e.g. Gino et al., 2009a; Dimant, 2019; Gächter et al., 2019).

Fourth, and most importantly, we show that the erosion of norm compliance strongly depends on the degree of known social proximity between individuals. When people are aware of social similarities with their peers (in ObservationSP), we observe strong responses not only to norm violations, but also to examples of norm compliance. This mitigates

norm erosion in comparison to less socially proximate groups. This result highlights the importance of the broader social context in driving norm compliance. It also illustrates that punishment of norm violations, frequently seen as crucial to sustain norms (e.g., Coleman, 1994; Sutter et al., 2010; Fehr and Schurtenberger, 2018), is not always needed: even minimal social proximity can prevent the erosion of norm compliance.

Our fifth novel contribution to the existing literature is the ability to study the heterogeneity of norm erosion. This not only helps to understand *who* is more susceptible to the effects of observation and proximity, but also *to what extent* their behavior changes and contributes to the erosion of norm compliance over time. Our experimental design is particularly suited for this analysis: we deliberately designed a one-shot individual give or take decision towards the charity prior to both announcing the repeated group stage and being randomly assigned and exposed to groups. This design choice enables us to differentiate between initial types of pro- and anti-social individuals and trace their behavioral change across periods and treatments in a non-strategic setting. Our results paint a gloomy picture and indicate that initially anti-social individuals are largely invariant to behavioral change, whereas initially pro-social individuals are strongly swayed and their norm compliance erodes quickly. The behavioral reaction of initial abstainers falls in between.

In sum, we show that the dynamic of norm compliance is mainly driven by subjects' empirical expectations. Social norms exist to mitigate the tension between individual and collective welfare, but selfish motives continue to be present. Even when a social norm exists, the congruence of normative and empirical expectations determines compliance, in that individuals prefer to follow a norm conditionally on *both* expectations being present (Bicchieri (2006, 2016)). Incongruence between these expectations, especially when empirical expectations point to norm violation, weakens the normative pull of the norm (Bicchieri and Xiao, 2009). This is the reason we observe an asymmetric response to norm violations versus compliance: observing violations bolsters selfish motives, and weakens normative commitments. Observing compliance has a lesser effect, especially when compliance is not universal. Our main result points to a mitigating effect of social proximity. Focusing individuals on social identity (even a minimal one) seems to induce sensitivity to *both* signals, positive and negative, as individuals pay greater attention to and tend to emulate the behavior of those who are similar to them. This is an important finding because it shows that studying social norm compliance only in fully anonymous and socially decontextualized settings overestimates the erosion of norm compliance.

2. Social Norms: A Conceptual Framework

The concept of social norms has been used in various disciplines, from economics to sociology to psychology (e.g., [Elster, 1989](#); [Coleman, 1994](#); [Ostrom, 2000](#); [Cialdini and Goldstein, 2004](#); [Burke and Young, 2011](#); [Fehr and Gächter, 2000](#); [Bicchieri, 2006](#); [Fehr and Schurtenberger, 2018](#)). The precise definition of what constitutes a social norm varies across (and within) these disciplines, although norms are often understood as behavioral patterns that are predominant within a group, are supported by a shared understanding of acceptable actions, and are sustained through that group’s social interactions (for instance [Ostrom, 2000](#)). One issue with most existing definitions of social norms, including the one outlined above, is that they are difficult to operationalize in empirical research, since it is not clear how to practically measure norms or assess their influence on behavior. Here we follow the definition of social norms proposed by [Bicchieri \(2006\)](#). Differently from other approaches, this definition is firmly grounded within a belief and preference framework and is therefore easier to operationalize empirically (for a discussion see [Bicchieri, 2016](#)).

Our starting point is the notion that a norm is a special type of behavioral rule that prescribes or proscribes a certain behavior to a specific group of people in a specific class of situations. We say that the rule satisfies the properties of a social norm if the individual prefers to follow the rule on the condition that (a) she believes that sufficiently many others also follow it (empirical expectation), and (b) she believes that sufficiently many others believe that one should follow the rule and/or may be willing to sanction transgressions (normative expectation) ([Bicchieri, 2006](#)). Finally, we say that a norm exists within a group of people if the rule is known to apply within that group and normative expectations are mutually consistent.

Note that our approach to social norms ultimately defines norm compliance in terms of two underlying primitive constructs: beliefs about what is commonly done and what is commonly approved of; and preferences to undertake a specific behavior that are conditional on such beliefs. A key advantage of defining a social norm in terms of preferences and beliefs is that we can independently measure and quantify these primitive constructs (and hence norms) using the methodology of experimental economics. Belief-elicitation protocols (e.g., for normative expectations as proposed by [Bicchieri and Chavez, 2010](#), or [Krupka and Weber, 2013](#)) can be used to measure whether individuals hold sufficiently high empirical or normative expectations, and hence to determine whether a majority believes that a norm applies to a given situation.

A second important advantage of our definition is that it makes it easier to differentiate between social behaviors that are regulated by social norms as opposed to behaviors regulated by other types of behavioral rules. A crucial distinguishing feature of behaviors that are regulated by social norms is that the motivation to undertake such behaviors is *conditional* on a person’s belief about what is commonly done *and* what is commonly approved of within that person’s reference network. The psychological foundation of such conditionality lies in the notion that humans naturally strive to obtain approval and avoid disapproval from others (e.g., Sugden, 1998, 2000; Brennan et al., 2004; Bicchieri, 2006, 2016). In contrast, the motivation to follow other types of behavioral rules may only be conditional on a person’s belief about what is commonly done, as in the case of “descriptive norms” (Cialdini, 2007), or unconditional on her beliefs, as in the case of “moral rules” (Bicchieri, 2006). Assessing beliefs and preferences (and whether preferences depend on beliefs) lets us clearly distinguish between types of social behaviors, which is an important step in designing interventions to change behavior.

Finally, a crucial advantage of our approach is that it allows to distinguish between the notion that a norm exists and the notion that it is followed. The existence of a norm depends on the mutual consistency of individuals’ normative expectations. Whether or not a norm is followed, however, depends on a number of factors, including the degree of congruence between normative and empirical expectations, as well as factors that determine the individual’s sensitivity to what the norm stands for. A norm may exist without being followed at a particular time (Bicchieri, 2006), but if transgressions become common, the norm may eventually disappear. For instance, consider two communities that hold similar normative expectations about the inappropriateness of littering: in both cases, individuals believe that littering is inappropriate, i.e. an anti-littering norm exists. In one community, this general social disapproval for littering is accompanied by evidence that littering is rare; empirical and normative expectations are congruent. In the other community, however, the disapproval for littering is accompanied by widespread evidence that littering is common; empirical and normative expectations are incongruent. If individuals in the latter community observe a sufficient number of transgressions, they may transgress too, since their compliance is conditional upon what others do. As a consequence, compliance (i.e. norm-following) may be lower in the latter community than in the former one, even if all members hold similar beliefs about what is socially appropriate. Situational factors may also play an important role in determining whether or not a norm is followed: e.g., privacy can reduce compliance (Allen, 1965), and punishment can enhance it (Fehr and Gächter,

2000; Sutter et al., 2010; Fehr and Schurtenberger, 2018, but see also Abbink et al. 2017).

In our experiments, across three treatments, a group of anonymous individuals play a repeated non-strategic ToG donation game where they can take or give money from/to a charity. We systematically vary whether subjects can observe the donation behavior of other members of their group in previous periods (NoObservation and Observation treatments, respectively), and the degree of known social proximity between group members (ObservationSP treatment). In a separate norm-elicitation experiment, we present direct evidence that a social norm that prescribes giving to and proscribes taking from the charity exists. Our data show that individuals' normative expectations are congruent with their own personal normative beliefs. This rules out that we observe an instance of 'pluralistic ignorance', whereby people privately reject a norm but believe that most others will follow it (see Smerdon et al., 2019 for pluralistic ignorance in the context of social norms).

With a separate group of individuals in our ToG donation game, we study how norm compliance is affected by feedback about what other group members have done in previous periods (by comparing Observation and NoObservation treatments). In our framework, feedback about the actions of others in previous periods constitutes information about what is commonly done in the experimental game. We expect this information to influence a subject's empirical expectations and hence her behavior, since norm-following partly depends on what others commonly do. Thus, we expect subjects would stop following the norm once they have enough evidence of others violating it. Should we similarly expect subjects to follow the norm after observing that others are also following it?

Although our conceptual framework does not make any explicit prediction about the symmetry or asymmetry of empirical information effects, the existing evidence suggests that these effects may be asymmetric: subjects strongly reduce compliance with norms when they receive evidence that others are transgressing, but they only weakly increase compliance if they observe that others comply (e.g., Croson and Shang, 2008; Thöni and Gächter, 2015; Dimant, 2019; Gächter et al., 2019). One explanation for this asymmetry is that social norms involve a tension between selfish and social goals - as is the case for the norm of giving in our experiment. Since compliance with norms is costly from the point of view of individual material self-interest, individuals may look for ways to discount collective benefits and rationalize self-interested behavior. Thus, they may make a self-serving use of empirical information, by responding more strongly to it when it is in their interest to do so (see, e.g., Charness et al., 2019; Dimant et al., 2019). In line with this interpretation, there is evidence that individuals revise their normative expectations

after receiving information about others' behavior. That is, commonness of an observed behavior influences its acceptability status (Lindström et al., 2018). However, this updating of expectations seems to occur predominantly when individuals receive evidence of norm transgression rather than norm compliance (Gächter et al., 2019).

The above discussion paints a rather pessimistic picture about the dynamics of norm compliance within groups that interact repeatedly and where individuals can observe each other's past behavior. If examples of norm violations are indeed more contagious than examples of compliance, we would expect to see an inexorable decline in norm-following over time in the absence of mechanisms that can effectively contrast the negative influence of bad examples. Are there means to reduce the erosion of norm compliance? Previous research has explored the effects of making normative expectations salient (Cialdini and Goldstein, 2004; Schultz et al., 2007; Jachimowicz et al., 2018). However, it has been shown that the combination of incongruent empirical and normative information is detrimental to norm compliance when the empirical information points to norm violation (Bicchieri and Xiao, 2009). One hitherto unstudied possibility is the role of *social identification* in moderating the effect of observed norm deviance. Social identification is defined as “the individual's knowledge that he/she belongs to certain social groups together with some emotional or value significance to him/her of the group membership” (Tajfel, 1982). Existing research has found that individuals care about the behavior of a group if they identify with it, and perceived group identification depends upon recognizing similarities between the self and other group members (Perkins, 2002). According to social identity theory, identification with a group promotes behavioral conformity because when the individual self-categorizes as a group member, she seeks to undertake behaviors that are perceived as “what good group members do” and avoid behaviors that are perceived as inappropriate for good group members (Turner, 1985; Akerlof and Kranton, 2000; Chen and Li, 2009). These perceptions of appropriateness are informed by the behavior of other group members, and conforming to perceived appropriate behavior signals commitment to the group (Hogg, 1992). Recent experimental evidence suggests that self-identification with the group drives conformity even when norms are completely arbitrary (Pryor et al., 2019).

These considerations imply that an accentuation of similarities between self and other in-group members (what we call the “social proximity” between individuals) will lead the individual to give greater weight to empirical information about what other group members do, both because this information may be viewed as a direct source of what is perceived as “good group behavior” and because deviations from group behavior may signal low

commitment to the group and thus trigger social sanctions by other in-group members. In terms of our conceptual framework, we can model these mechanisms as an heightened conditionality of preferences on empirical expectations in settings where individuals interact with socially proximate others. That is, individuals will attach greater significance (and respond more strongly) to the behavior of others when the others are in-group members than when they are anonymous strangers. Moreover, we may expect the self-serving use of empirical information to be less prevalent among socially proximate individuals: since non-conformity signals low commitment to the group, it may be more costly for the individual to rationalize self-interested behavior by ignoring evidence of norm compliance when this evidence comes from the behavior of in-group members.

Thus, we expect a different dynamic of norm compliance to apply within groups that observe the past behavior of socially proximate others (ObservationSP treatment), as compared to groups that observe the past behavior of others who are not known to be socially proximate based on our measure (NoObservation and Observation treatments). In the Observation case, the existing evidence about asymmetric empirical information effects suggests the decline of norm-compliance. In the ObservationSP case, we expect this process of erosion to be halted or substantially reduced, since individuals will respond to both examples of norm violation and norm compliance.

3. Experiments

We conduct three experiments (one of which is a robustness check and is explained in detail in the Appendix), for which we collected data from a total of 1,306 participants. All experiments rely on a Take-or-Give (ToG) donation game described in Section 3.1. In the first experiment, described in Section 3.2, we use the [Bicchieri and Chavez \(2010\)](#) norm-elicitation procedure to establish whether a social norm exists in the ToG game. We ask participants to describe their normative and empirical expectations, as well as their personal normative beliefs, in order to identify which norm, if any, applies to the setting.

In the second experiment (Section 3.3), carried out with a different set of participants, we measure how participants behave in the ToG game. The design of this behavioral experiment consists of several treatments where we vary whether subjects can observe other participants' behavior as well as the degree of known social proximity between themselves and other participants. We can thus measure the extent to which participants comply with or violate the norm of giving that applies, and how norm compliance is affected by

receiving information about the behavior of others and their social proximity.

3.1. *The Take-or-Give Donation Game*

In the ToG game each subject makes a donation decision towards a charity. There are three possible charities available and the subject chooses one of the following to be paired with: Doctors Without Borders, World Wildlife Fund, or UNICEF. The game starts with the subject and the charity both provisionally endowed with 100 ECUs (with 10 ECUs = \$1). The action space of subject i is represented by $x_i \in [-100, 100]$, where x_1 is an integer representing the monetary amount the i^{th} subject decides to take from or give to the charity, with the value 0 indicating no change to the initial equal split in endowments. The game is thus a variant of the dictator game that includes take options (e.g., List, 2007; Bardsley, 2008; Korenok et al., 2014), but where the recipient is a charity instead of another subject (as in Eckel and Grossman, 1996; Exley, 2015; Grossman and Eckel, 2015; Bolton et al., 2019; Dimant, 2019). A non-strategic setup is necessary to study the changes in norm compliance in absence of strategic elements that are present in related games such as public good games or ultimatum games (e.g., Bohnet and Zeckhauser, 2004).

3.2. *Norm-Elicitation Experiments*

We designed three set of experiments to elicit subjects’ (1) personal, (2) normative and (3) empirical expectations about behavior in the ToG game. These experiments were conducted with US residents on the online platform Amazon Mechanical Turk (MTurk) using the software Qualtrics.¹

To elicit personal normative beliefs and normative expectations, we used the procedure introduced by Bicchieri and Chavez (2010). We recruited two different groups of subjects and described the ToG donation game to them. We asked the first group of subjects ($N = 107$) to report what they thought one “should do” (i.e., their personal normative beliefs) in the ToG game: take from the charity; give to the charity; or abstain from changing the initial endowments (to make the task manageable, we did not ask about all possible levels of giving and taking, but only about the three options that constitute

¹We ensure high quality data collection by utilizing a combination of CAPTCHAs and sophisticated screening questions to avoid pool contamination. We applied the following restrictions to the subject pool: subjects had to be in the U.S. and approval rating > 99%. We used online tools to test IPs for low quality respondents. Recent evidence suggests that this contributes to the robustness, generalizability, and reproducibility of laboratory findings on MTurk (Arechar et al., 2018; Snowberg and Yariv, 2018),

natural partitions of the action space: take, give, or do nothing). We then asked the second group of subjects ($N = 105$) to predict what the majority of subjects in the first group had indicated that one “should do” in the experiment (i.e., normative expectations), again choosing among three possible responses: take, give, or abstain. This second elicitation was incentivized: subjects were paid a monetary reward of appropriate size relative to the effort and duration of the experiment (\$0.25) if their normative expectations matched the personal normative beliefs of the first group of subjects.² As we show in Section 4.1, these elicitations clearly indicate that a norm of giving to charity exists in the ToG game. The norm prescribes giving and proscribes taking in the sense that giving is viewed as what one ought to do in the game. Based on these results, we refer to the act of giving to the charity as norm-compliant behavior, and taking from the charity as norm-violating behavior.

Finally, we also conducted another belief-elicitation experiment, with a different group of subjects ($N = 100$), where we measured their empirical expectations. Subjects first read a description of the ToG game and were then told that a number of participants had previously made ToG choices in the game. They then had to predict which of the three options (give, take, abstain) the majority of those previous participants chose. Again, subjects were paid a small monetary reward (\$0.25) if they predicted correctly, and nothing otherwise. We interpret this elicited guess as a measure of subjects’ empirical expectations, which we can use to predict initial compliance with the norm in the ToG game, prior to the exposure to feedback about the behavior of others in the game.

3.3. Behavioral Experiment

We next study how participants actually behave in the game. We conducted a behavioral experiment where we recruited 842 participants (university students) to make choices in the ToG donation game, none of which have participated in any of the previous norm-elicitation experiments. Participants were on average 22 years old and 70 percent were female. The experiment was conducted in zTree (Fischbacher, 2007) and run in 52 sessions in the Behavioral Ethics Lab at the University of Pennsylvania across three between-subject treatments, described in more detail below. We have three treatments across which we vary the amount of information that participants receive in each period of Part III about the

²We also used an alternative method to measure normative expectations, based on the procedure introduced by Krupka and Weber (2013). We present this additional method in Online Appendix I. Results are very similar across the two methods. However, an advantage of the method presented in the main text is that it can uncover pluralistic ignorance.

ToG donation choices of other members of their group. In each treatment, every participant went through the same four parts of the experiment (plus a payment phase). The four parts as illustrated in Figure 1 were introduced sequentially. Participants were aware that the experiment consisted of multiple parts, but the details of each new part were only revealed upon completion of the previous parts. Throughout the experiment all decisions were collected anonymously, i.e., none of the subjects' decisions could be linked to the respective individual (see Appendix II and III for the instructions).

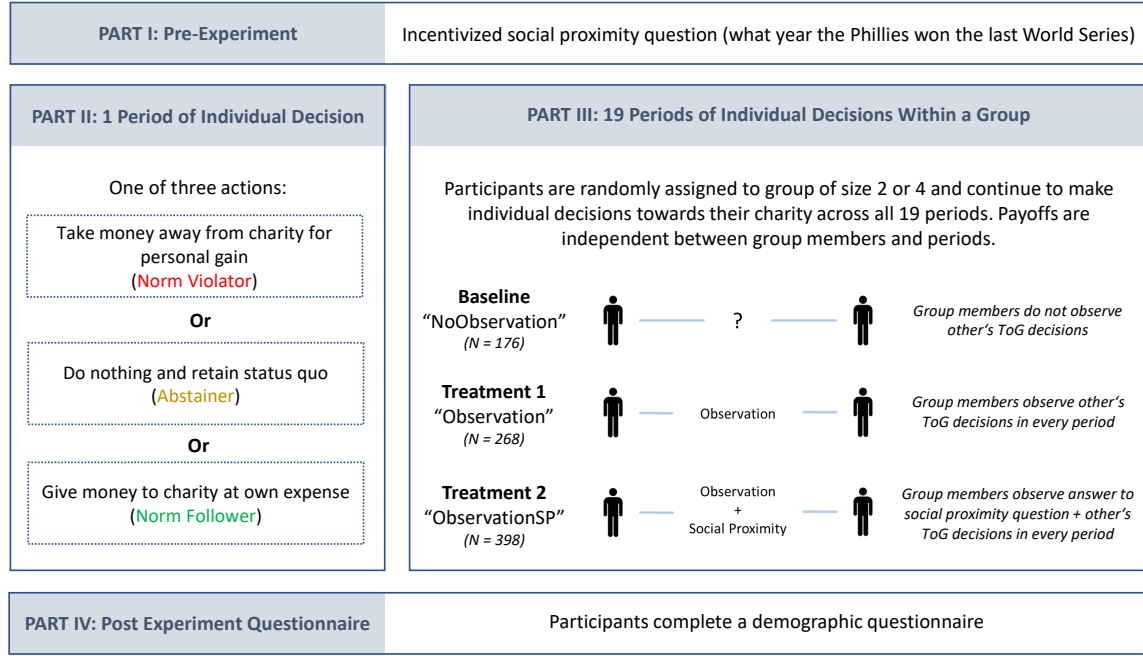


Figure 1: Experimental Design.

Part I – Social Proximity Question

At the beginning of the experiment, all participants were asked one incentivized knowledge question (the year in which Philadelphia's baseball team won its last World Series). Participants were paid \$1 at the end of the experiment if they provided the correct answer (2008) and \$0 otherwise. Overall, about 27% of our participants knew the correct answer to the question, indicating that getting it right would indeed send a salient signal. In the ObservationSP treatment (details see below), the answers to the knowledge question were used to prime social proximity: at the end of Part II, participants were matched in groups and informed about how each member of their group had answered the question (and thus

about the extent to which they could be considered supporters of the Philadelphia team).³ In the other treatments, participants were also asked the knowledge question, but were not informed about the answers of the other group members. In these treatments, participants learned whether their answer to the question was correct or not (and hence their payment for it) at the end of the experiment, while in the ObservationSP treatment, this information was given between Part II and Part III (see below).

Part II - Individual Decision (one period)

All participants made one ToG decision. As explained in Section 3.1, participants first selected their preferred charity from a list and then chose how much to take or give (from 0 ECUs up to +/-100 ECUs) from/to this charity. It was public knowledge that each individual's decision did not affect another participant's payoff or the payoff of a charity besides the individual's own. This removes any type of interdependency and strategic considerations between participants' payoffs.

Part III - Individual Decisions (nineteen periods)

Following the individual decision in Part II, for periods 2 to 20 participants were randomly assigned to a group that consisted of either two ($N = 354$) or four participants ($N = 488$). The group allocation remained constant throughout the experiment and was solely done for the purpose of whose decisions subjects could observe. We varied group size because existing evidence suggests that group size affects the pressure to conform with group norms (see Bond, 2005). Each participant continued to make the same type of individual decision as in Part II towards the charity for another 19 periods. In each period, both the subject and the charity were endowed with 100 ECUs each in order to avoid any path-dependent carryover effects from previous periods; payoffs were also strictly individual and hence independent between group members. Our treatments varied as follows:

In our baseline (NoObservation), participants were placed in groups of either two or

³In order to ensure that our social proximity signal is sufficiently strong, we decided to use a knowledge question about sports rather than an arbitrary grouping following the minimal-group paradigm literature (Tajfel and Turner, 1979). Historically, Philadelphia has had a strong sports identity that was facilitated by the success of the 76ers (NBA), Flyers (NHL), Phillies (MLB), and, most recently, the Eagles (NFL). At the time of running the study, the most recent big sports success was the 2008 World Series win of the Phillies. Arguably, a correct reply indicates exposure to and knowledge about the sports (and potentially, in addition, being a fan of the Phillies). One advantage of using a knowledge question as opposed to a pure preference question (e.g., "Are you a fan of Philadelphia sports teams?") is the avoidance of a social desirability bias and demand effects.

four members, but no behavior was observable by or revealed to any other participant. We study the natural evolution of compliance with the norm of giving across the 20 periods of Part II and III of the experiment.

Our first treatment (Observation) is identical to the baseline condition except that in each period of Part III, participants received information about the ToG decisions of the other group members in the previous periods of the experiment, including their decisions in Part II of the experiment. This feedback information was presented in the form of a history table to allow participants to track the ToG decisions of each participant in their group over the whole duration of the experiment. At no point during the experiment were participants able to observe the ToG decisions of another participant who was not in their group. The comparison between Observation and NoObservation allows us to assess the effect that repeated exposure to empirical information has on compliance to the giving norm in the ToG donation game.

Finally, our second treatment (ObservationSP) is identical to Observation, except that in addition to observing group members' ToG decisions, participants also received information about their social proximity to their group members through the answers to the knowledge question in Part I. This information was displayed right below each of the group members' history tables in the form of an "answered the question correctly/incorrectly" indicator. We use this treatment to assess how the ToG donation decisions with respect to the giving norm is affected by the social proximity of other group members.

Part IV - Post-Experimental Questionnaire

We elicited general participant information, such as age and gender, alongside individual behavioral attributes, such as risk attitudes.

Payment

All participants were paid a show-up fee of \$10 in addition to their earnings from decisions made in the experiment. At the end of the experiment, we randomly selected one subject in each session and implemented exactly one (randomly chosen) decision of this subject from the 20 decisions he/she had made during Part I, II and III of the experiment. We used this randomly chosen decision to compute the subject's own payoff as well as the charity's payoff. The randomly selected participant also determined the charity which the money was transferred to/from. The decision of every other subject in the same session did not count towards his/her own and the charity's payoff. Instead, each of these subjects

received exactly 100 ECUs regardless of their allocation decision towards their charity in any of the periods. Participants were informed of this payment procedure at the beginning of the experiment. We chose this type of ‘pay one’ payoff structure to further eliminate from the design any type of interdependencies between subjects, not just in terms of their material payoffs, but also in terms of potential social payoffs — e.g. concerns about the charity, strategic substitution effects (‘I can take because you give and balance out the payoffs’), among others. This setup is in line with suggestions made by [Charness et al. \(2016\)](#) and retains incentive compatibility as theoretically argued by [Azrieli et al. \(2018\)](#). The experiment lasted about 45 minutes, yielding an average hourly pay of \$18.

3.4. Behavioral Hypotheses

In this section, we present hypotheses derived from the conceptual framework outlined in Section 2. First, we conjecture that in our experimental setting a norm of giving to charities exists. A necessary condition for the existence of a norm in the ToG donation game is that people’s personal normative beliefs and their normative expectations differ across the three options of taking, abstaining, and giving (if not, all three options would be normatively equally appealing, and hence there would be no norm). It is plausible to expect that fewer people will think that taking rather than abstaining or giving is their personal, normatively most appealing option; the same is likely true for normative expectations. Thus, we formulate our first hypothesis as:

H₁: *A norm of giving exists in the ToG donation game. Taking from the charity is considered the personally normatively least appealing and also least expected option.*

Having established that a norm of giving exists in the ToG donation game and that a majority of people think that both giving is expected *and* personally normatively desirable, we then study how the availability of empirical information affects compliance with it. As discussed in Section 2, we expect an asymmetric effect of empirical information on norm compliance: evidence of norm-violating ToG decisions will have a stronger impact than evidence of norm compliance. Consequently, we expect a decline in compliance with the giving norm across the 19 periods of Part III in Observation, compared to NoObservation where information is not available. Thus, our second hypothesis is:

H₂: *In Observation, individuals react more strongly to examples of norm violation (taking) than examples of norm compliance (giving). Over time, this leads to a decline of compliance with the norm of giving compared to NoObservation.*

As discussed in Section 2, we conjecture that the asymmetric effect of empirical information (and the subsequent decline in norm compliance) is due to a self-serving interpretation of the information available to the individual as they may ignore evidence of norm compliance because conforming with it is materially costly. However, we also expect social proximity to reduce this self-serving use of empirical information: ignoring evidence of norm compliance may be costly when this evidence comes from the behavior of socially proximate group members, since non-conformity signals low commitment to the group. Thus, we expect the decline of norm-compliance to be reduced (or completely halted) among socially proximate groups. This conjecture implies that we should expect a smaller decline in norm compliance in ObservationSP compared to Observation. Moreover, we expect that this effect is driven by the most socially proximate groups in ObservationSP - i.e. those groups where all members have given the same answer to the sport team’s social proximity question of Part I of the experiment.

***H₃:** Social proximity reduces the asymmetric effect of empirical information: subjects in ObservationSP will respond to examples of both norm compliance and violation, thus reducing the erosion of norm compliance. This effect will be driven by subjects in ObservationSP that belong to socially proximate groups.*

4. Results

We start by presenting evidence from the norm-elicitation experiments that establish the existence of a norm of giving in the ToG game. We then present the results of the behavioral experiment.

4.1. Norm-Elicitation Experiments: Existence of a Norm of Giving to Charity

Figure 1 shows the results of the norm-elicitation experiments. The figure presents data on three types of beliefs. First, it presents subjects’ personal normative beliefs. These are shown by the dashed line in Figure 2, representing the percentages of subjects believing that one ought to take, give, or abstain from changing endowments in the ToG game. The figure shows that a majority of subjects believes that giving constitutes prescribed behavior (51% indicated so), while only 17% of subjects believe that one ought to take.

The thick dark line in Figure 2 shows subjects’ normative expectations (second-order beliefs about what ought to be done in the game). The line shows that second-order beliefs track closely the corresponding (first-order) personal beliefs of prescribed behavior: 54%

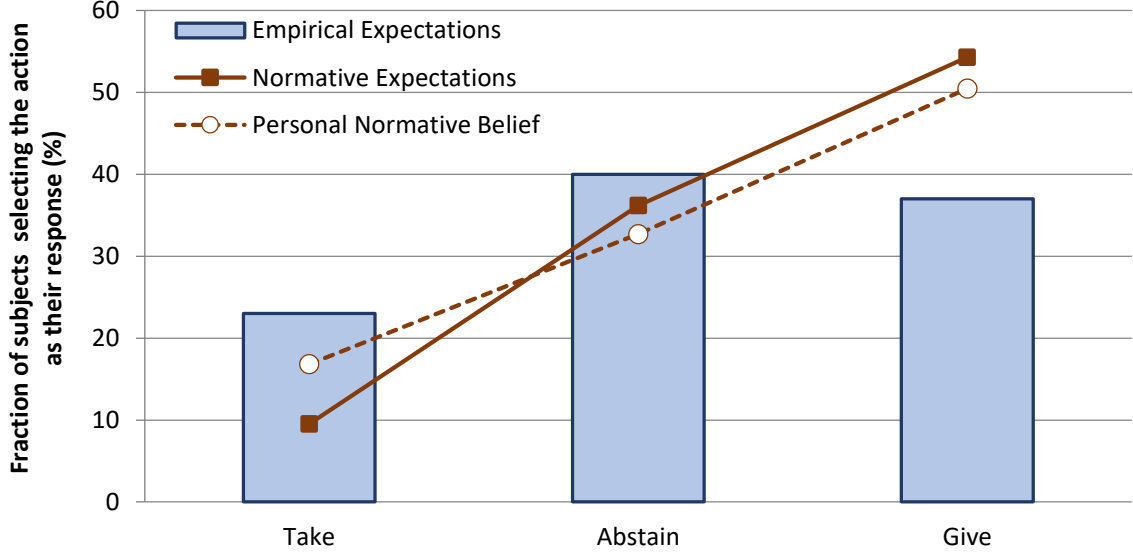


Figure 2: Personal normative beliefs, normative expectations and empirical expectations in the ToG game.

of subjects believe that a majority of others indicated that giving is what one ought to do in the game, and only 10% believe that the majority of others thought that one ought to take. A $\tilde{\chi}^2$ test shows that the distributions of personal normative beliefs and normative expectations are not significantly different from one another ($\tilde{\chi}^2(2) = 2.471$, $p = 0.290$). Together, the personal normative beliefs and normative expectations indicate that there are mutually consistent expectations among a majority of subjects that giving constitutes prescribed behavior in the experiment. This result confirms Hypothesis 1 and shows that a norm that prescribes giving and proscribes taking exists in the ToG donation game.

Finally, the blue bars in Figure 2 represent subjects' empirical expectations. This data represents subjects' empirical expectations in the absence of any feedback information about behavior of others in ToG game. We find that 37% of subjects expect a majority of participants in the game to give to charity, and another 40% to leave initial endowments unchanged. Only 23% of subjects expect others to take in the game. Thus, although normative expectations clearly indicate that giving is what one should do, subjects' initial empirical expectations indicate that subjects expect most others to abstain from taking, rather than to actively give to charity. More generally, empirical expectations appear more heterogeneous than normative expectations and personal normative beliefs, with a sizeable fraction of subjects anticipating others to actually take from charity; a $\tilde{\chi}^2$ test shows

significant differences between the distributions of empirical and normative expectations ($\tilde{\chi}^2(2) = 9.311$, $p = 0.010$).

Overall, the norm-elicitation experiment indicates that a norm of giving exists in the ToG game, but also that there is a difference between what individuals believe one ought to do and what they believe would actually occur in the ToG game. Since norm compliance depends both on normative and empirical expectations and our results show that these expectations are not completely congruent, we may expect to observe some deviations from the norm of giving in the behavioral experiment.

4.2. Behavioral Experiment

4.2.1. Behavior in Part II - Initial Compliance With the Norm of Giving

We observe the donation behavior of 842 subjects across 20 periods of the ToG game. The first period of the game (Part II of the experiment) is exactly the same in all of our treatments: subjects make a ToG decision without receiving any information about the behavior of others or their social proximity. In the Observation and ObservationSP treatments, participants did not know that their decisions in Part II would be displayed to other group members in Part III. We can use these initial donation decisions to establish the extent to which the norm of giving is followed when subjects have to rely on their homegrown expectations about what ought to be done and what is commonly done.

Table 1 displays the average ToG donation in Part II of the experiment across the three treatments. Due to lack of group size differences we pooled the data across group sizes (details are in the Online Appendix). The table also displays the percentage of participants in each treatment who gave to, took from, or abstained from changing initial endowments. We label participants who gave to charity “norm followers”, those who took money from charity “norm violators”, and those who neither gave nor took as “abstainers”.

Participants on average took between 18.9 and 25.2 ECUs from the charity. In all treatments, the majority of participants took from the charity (frequencies range between 42% and 45%, depending on the treatment). Only a minority of participants gave to the charity (between 16% and 19% depending on the treatment). As expected per proper treatment randomization, we find that behavior in Part II does not differ across treatments (Kruskal-Wallis test, $df = 2$, $p = 0.703$).

Our data show that there is a significant gap between subjects’ initial behavior and the normative expectations measured in our norm-elicitation experiment. The latter clearly indicates that a norm of giving exists in the game, but the behavior in the experiment

	Average amount given/taken	Percentage of initial takers (norm violators)	Percentage of initial equal split (abstainers)	Percentage of initial givers (norm followers)
NoObservation	-18.9 (57.1)	45%	36%	19%
Observation	-25.2 (55.9)	44%	40%	16%
ObservationSP	-21.8 (51.1)	42%	41%	19%

Table 1: ToG decisions in Part II of the experiment. Standard deviations in parentheses.

shows that compliance with the norm is quite low. This is not entirely surprising in light of our measure of initial empirical expectations (see Figure 2), which suggests that subjects anticipate a fair amount of taking and abstention from giving in the game. In fact, the actual level of taking and abstention in the behavioral experiment is almost double what was anticipated by subjects in the norm-elicitation experiment. Taken together, these results point to the importance of differentiating between the existence of a norm and the norm being followed in a population.

4.2.2. Behavior in Part III - The Dynamics of Norm Compliance

We now explore how compliance is affected when subjects are given information about the behavior of others in the game, which allows them to revise their empirical expectations downwards and towards the actual behavior in the game. Figure 3 (left panel) shows the evolution of average donations over the 19 periods of Part III across our treatments. Figure 3 (right panel) shows ToG donations averaged across all 19 periods. In both cases, the figures plot average donations normalized relative to behavior in Part II (which is displayed as period 0 in Figure 3). A positive value indicates that, relative to the average in Part II, the amount given to the charity has increased, and a negative value that the amount taken from the charity has increased.

In NoObservation, where participants did not receive any additional information about behavior of other group members, donations in Part III are similar to donations in Part II. On average, across the 19 periods of Part III, participants took 4.9 additional ECUs from the charity relative to Part II. Using two-sided Wilcoxon sign-rank tests with group averages as observations, we cannot reject the null hypothesis that average donations in Part III of

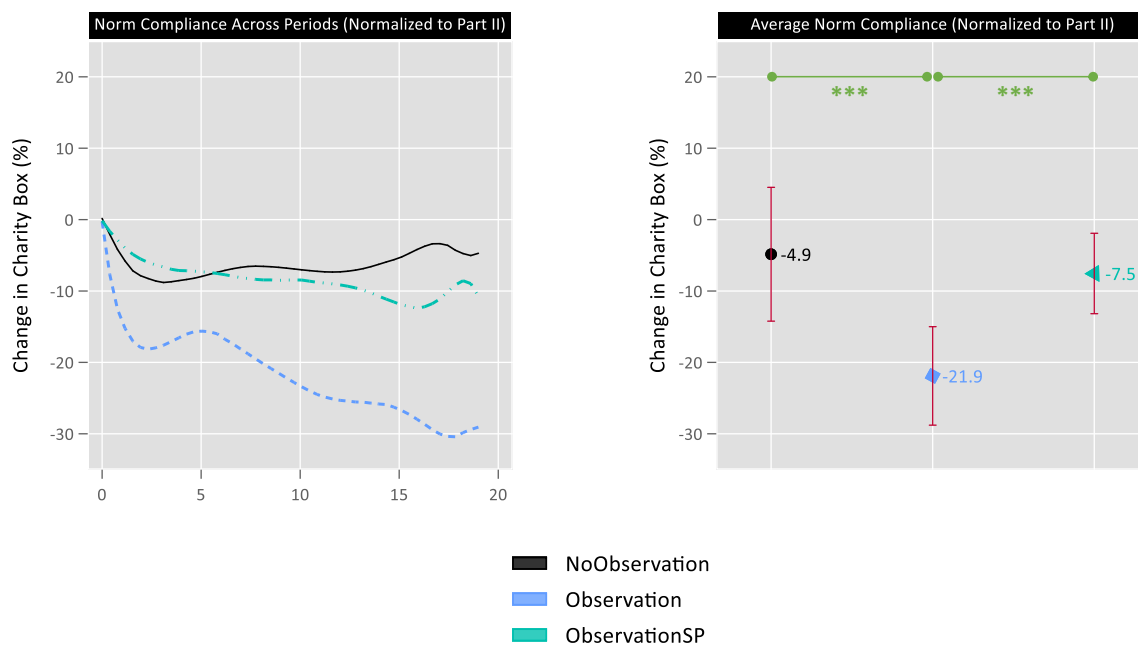


Figure 3: Left panel: Subject’s behavior across treatments and periods in Part III, normalized to average behavior in Part II (period 0 in the graph). For purposes of readability, lines represent kernel-weighted local polynomial smoothing of degree 5. Right panel: Average subject’s donations across treatments, averaged over the 19 periods of Part III. Whiskers indicate 95% confidence intervals. Stars indicate significant differences using two-sided Mann-Whitney U-tests and group averages as observations, with p-values adjusted for a false discovery rate (Benjamini and Hochberg, 1995) at the conventional levels of $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$.

NoObservation are identical to donations in Part II ($z = -0.921$, $p = 0.357$). In the absence of new empirical information, we do not observe a significant change in compliance with the norm of giving across the 19 periods of Part III. In contrast, when participants received information about other group members’ behavior, the average amount donated to the charity dropped relative to donations measured in Part II. Using two-sided Wilcoxon sign-rank tests with group averages as observations, we detect significant differences between Part II and Part III donations for Observation ($z = -5.487$, $p < 0.001$) and ObservationSP ($z = -2.370$, $p = 0.018$). However, the decline in compliance appears to be much stronger in Observation, where participants took on average an additional 21.9 ECUs relative to Part II, compared to ObservationSP, where participants took on average only an additional 7.5 ECUs from the charity. A Kruskal-Wallis test confirms that there are significant differences between the three treatments in Part III of the experiment ($df = 2$, $p < 0.001$). We perform bilateral treatment comparisons using two-sided Mann-Whitney U-tests and group averages

as observations, with p-values adjusted for a false discovery rate (Benjamini and Hochberg, 1995). Across all periods of Part III, we find that the average donations are significantly different between Observation and NoObservation ($z = -3.732$, $p < 0.001$) and also between Observation and ObservationSP ($z = 3.524$, $p < 0.001$). No significant differences are detected between NoObservation and ObservationSP ($z = 0.877$, $p = 0.381$).⁴

We test the robustness of our findings by performing multivariate regression analysis that include controls for observable differences across treatment groups. Table 2, column 1 reports results of random-effects GLS regressions where the dependent variable is the amount given to the charity by subject i in period t of Part III of the experiment. The independent variables are treatment dummies (Observation and ObservationSP, using NoObservation as base category), a dummy for group size (using group size 2 as base category), a variable measuring the amount contributed by the participant in Part II (to control for a participant’s underlying predisposition to give or take from the charity), a period variable, and various other controls (including the correct answer to the social proximity knowledge question and socio-demographic controls collected in the post-experimental questionnaire).

The regression in column 1 of Table 2 corroborates the results reported above. Relative to the NoObservation treatment, participants in the Observation treatment took significantly more ECUs from the charity; the difference is significant at the 1% level. Moreover, a test of equality of coefficients between Observation and ObservationSP confirms that participants took more from the charity in Observation ($\tilde{\chi}^2(1) = 19.30$, $p < 0.001$). We do not observe, however, any significant difference between NoObservation and ObservationSP. We also do not detect any differences in behavior in Part III of the experiment across participants who were in groups of size 2 or 4. We do not observe any general effect of group size on taking behavior; a separate regression in which we interact group size with treatments does not yield any significant differences either.

Our results so far show that when participants only receive feedback about the behavior of other group members, they mainly respond to behavior that is in violation of the norm of giving. In other words, we observe a spread of norm violations, in line with our second hypothesis about the effects of empirical information, H_2 . In contrast, when participants

⁴Importantly, our main result is not driven by the disproportional existence of taking-behavior in Part II. In fact, as illustrated in Figure A.2 in the Online Appendix, we do observe a substantial decay in norm compliance even for the sub-groups in which such behavior was in the minority and thus a ‘not-taking’ norm was present. This additional robustness check emphasizes the generalizability of the observed dynamics.

<i>DV: Amount Change in Charity Box</i>	(1)	(2)
Treatment		
(Base level: NoObservation)		
<i>Observation</i>	-18.02 ^{***}	-4.26
	(3.51)	(3.95)
<i>ObservationSP</i>	-3.67	0.60
	(3.09)	(3.15)
Groupsize = 4	-2.81	-1.73
	(2.64)	(2.36)
Initial Individual Donation (Part II)	0.67 ^{***}	0.67 ^{***}
	(0.02)	(0.02)
Round	-0.35 ^{***}	-0.25 ^{***}
	(0.08)	(0.07)
Neg. Avg Don. t-1		0.05
		(0.03)
Neg. Avg Don. t-1 * Observation		-0.29 ^{***}
		(0.05)
Neg. Avg Don. t-1 * ObservationSP		-0.17 ^{***}
		(0.05)
Pos. Avg Don. t-1		-0.01
		(0.04)
Pos. Avg Don. t-1 * Observation		-0.02
		(0.10)
Pos. Avg Don. t-1 * ObservationSP		0.21 ^{***}
		(0.07)
Zero Avg Don. t-1		-1.18
		(1.45)
Zero Avg Don. t-1 * Observation		0.69
		(3.20)
Zero Avg Don. t-1 * ObservationSP		2.60
		(2.18)
Constant	-36.95 ^{***}	-39.85 ^{***}
	(7.20)	(7.22)
Controls	Yes	Yes
N.	15998	15998
N. Clusters	299	299
R ²	0.47	0.49

Table 2: Random-effects GLS regressions. Dependent variable is donation of subject i in period t of Part III of the experiment. Controls are: gender, age, whether or not the proximity knowledge question has been answered correctly, a measure of self-control taken from [Tangney et al. \(2004\)](#), a variable measuring the (self-reported) liking for charities, a measure for the ability to anticipate future consequences of current behavior taken from [Strathman et al. \(1994\)](#), and a self-reported measure of risk preferences taken from [Dohmen et al. \(2011\)](#). We always use the same set of controls in all of our regressions. Because our analysis does not focus on the role of these controls and in order to allow for better readability, the estimates of the controls are not reported separately in Table 3, but are available upon request. Robust standard errors reported in parentheses (standard errors clustered at the group level). Stars indicate significant differences at the conventional levels of $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$.

receive both feedback about the behavior of other group members and information about their social proximity, the decline in compliance is substantially reduced. Behavior is at a level comparable to NoObservation, which supports our hypothesis H₃.

Why does norm compliance not significantly decline when information about social proximity is available? At the heart of H₂ and H₃ is the idea that, in the absence of knowledge of social proximity, participants mainly respond to examples of norm violations. When instead group members have knowledge of social proximity, participants also respond to examples of norm compliance in addition to norm violation, leading to a zero net effect.

To better understand the drivers of the decline in norm compliance across our treatments, in column 2 of Table 2 we augment the random-effects GLS regression of column 1 with variables capturing the different types of information that participants were exposed to in the 19 periods of Part III, as well as interactions between these variables and the treatment dummies. For the purpose of our analysis, we follow [Fischbacher and Gächter \(2010\)](#) and introduce variables lagged by one period that capture learning within groups.⁵

We distinguish between three types of information. First, we use the variable “*Neg. Avg Don. t-1*” to capture the effect of receiving feedback that other group members on average took money from the charity in the previous period (i.e. participants observed, on average, an example of norm violation). Thus, this variable is equal to the absolute value of the (lagged) average amount transferred to the charity by other group members if this amount is strictly negative, and it is equal to 0 otherwise. Second, we use the variable “*Pos. Avg Don. t-1*” to capture the effect of receiving feedback that the other group members on average gave money to the charity in the previous period (i.e. an example of norm compliance). This variable is equal to the (lagged) average amount transferred to the charity by other group members if this amount is strictly positive, and 0 otherwise. Finally, to capture the effect of observing an average zero transfer to the charity in the previous period, we use the dummy variable “*Zero Avg Don. t-1*”, which takes value 1 if the amount transferred was zero, and 0 otherwise. We then interact each of these variables with our treatment dummies.

The interactions between the lagged donation variables and the Observation dummy capture the differential impact of empirical information in Observation relative to NoOb-

⁵The presented results for t₋₁ are robust to also controlling for higher lags of t₋₂ and t₋₃. Results are available upon request.

ervation.⁶ The estimates reveal that observing norm-compliant behavior, or receiving feedback that on average other group members left the charity’s endowment untouched, does not statistically significantly affect a participant’s decision to give or take money to/from the charity. In contrast, observing norm-violating behavior has a strong and negative effect on donations. We estimate that receiving feedback that other group members take 1 ECU from the charity reduces the amount donated on average by $-0.29 + 0.05 = -0.24$ ECUs (see Table 2, column 2). This effect is significant at the 1% level.⁷ An F-test shows that the effects of examples of violation are statistically significantly larger than the effects of examples of compliance ($\tilde{\chi}^2(1) = 18.41$, $p < 0.001$).

The interactions between the lagged donation variables and the ObservationSP dummy capture the impact of feedback about others’ behavior when participants could also observe a signal of social proximity of the other group members. As in Observation, feedback that on average other group members abstained from changing the charity’s endowment does not affect a participant’s donation. Also, as in Observation, observing norm violations has a negative impact on donations, with an estimated reduction of donations of $-0.17 + 0.05 = -0.12$ ECUs for each ECU that others took from the charity (see Table 2). This effect is significant at the 1% level.⁸ However, differently from the Observation treatment, participants in ObservationSP also responded to examples of norm compliance. Receiving feedback that other group members give on average 1 ECU to the charity increases the amount donated by $0.21 - 0.01 = 0.20$ ECUs, which is also significant at the 1% level.⁹ An F-test does not reject the null that the effects of observed compliance and violation are of similar magnitude ($\tilde{\chi}^2(1) = 0.21$, $p = 0.649$).

This analysis reveals that the fundamental difference between the treatments with observation lies in the fact that, without information about social proximity, participants respond significantly only to examples of norm violation by reducing their donated amount.

⁶Given our interactions, the main coefficients of the lagged donation variables (*Neg. Avg Don. t-1*, *Pos. Avg Don. t-1*, and *Zero Avg Don. t-1*) measure the effects of the average amount transferred by other group members in NoObservation. This is a placebo test for the effects of social information since participants in NoObservation did not actually receive feedback about the amounts transferred by other group members during the experiment. As expected, the estimates show that in NoObservation information about others’ behaviors did not have any significant effect on donations.

⁷We reject: *Neg. Avg Don. t-1 + Neg. Avg Don. t-1 * Observation = 0* ($\tilde{\chi}^2(1) = 34.40$, $p < 0.001$).

⁸We reject: *Neg. Avg Don. t-1 + Neg. Avg Don. t-1 * ObservationSP = 0* ($\tilde{\chi}^2(1) = 13.23$, $p < 0.001$).

⁹We reject: *Pos. Avg Don. t-1 + Pos. Avg Don. t-1 * ObservationSP = 0* ($\tilde{\chi}^2(1) = 8.02$, $p = 0.005$) and both interactions (-0.02 and 0.21) are significantly different from each other ($\tilde{\chi}^2(1) = 30.70$, $p < 0.001$).

However, being informed about the social proximity of the other group members induces responses to examples of norm compliance as well as norm violation. Participants in the ObservationSP treatment reduce the amount donated when they observe others taking from the charity, but they also increase the amount donated when they observe that others give money to the charity. The effects of observing others taking and giving to charity are roughly similar in magnitude. As a consequence, the net effect of others' behaviors on donations in this treatment is not significantly different from zero.

Overall, these results suggest that knowledge of social proximity can moderate the decline in norm compliance that would otherwise happen in interactions among anonymous strangers. In line with H_3 , an implication of this finding is that the moderating effects of social proximity may be stronger among groups in which all members are similar in the relevant social dimension. We examine this conjecture in the regressions of Table 3 for ObservationSP in more detail. The regressions report the effects of observation of others' behavior (disaggregated in the same way as in Table 2) on donations in ObservationSP. In column 1, we only use observations from groups where either all group members had answered the knowledge question of Part I correctly (thus revealing that they were all likely to be fans of Philadelphia's baseball team), or where all group members answered it incorrectly (revealing that none of them were likely to be fans of Philadelphia's baseball team). In column 2, we instead use observations from groups where some group members had answered the question correctly, whereas others had answered it incorrectly. Hence, we expect social proximity to be higher among the former group than the latter. Based on this, and on our conjecture about the relation between proximity and observation, we expect the effect of proximity to be stronger among the more socially proximate groups.

The regression results confirm our conjecture. Among participants of socially proximate groups (column 1), we observe that donations are affected by examples of both norm compliance and violations. In contrast, in non-proximate groups (column 2), participants responded to examples of norm violations but did not significantly respond to examples of norm compliance, just as in Observation. Moreover, although not significantly different, the point estimates indicate the magnitude of the effects of violation and compliance are larger among proximate versus non-proximate groups (0.16 vs. 0.09 for examples of violations; 0.21 vs. 0.16 for examples of compliance). This result supports the conjecture that the mechanism behind the differential effects of information between Observation and

<i>DV: Amount Change in Charity Box (only in ObservationSP condition)</i>	(1) Proximate Groups	(2) Non-Proximate Groups
Negative Avg. Donation ϵ_{t-1}	-0.16*** (0.06)	-0.09*** (0.03)
Positive Avg. Donation ϵ_{t-1}	0.21*** (0.07)	0.16 (0.11)
Zero Avg. Donation ϵ_{t-1}	-2.24 (2.35)	3.83 (2.44)
Groupsize = 4	-4.52** (2.05)	-1.19 (2.67)
Initial Individual Donation (Part II)	0.66*** (0.05)	0.65*** (0.06)
Period	-0.16 (0.14)	-0.39*** (0.14)
Constant	-15.43 (15.98)	-58.46*** (16.94)
Controls	Yes	Yes
N.	3648	3914
N. Clusters	78	70
R ²	0.55	0.37

Table 3: Random-effects GLS regressions. Dependent variable is donation of subject i in period t of Part III of the experiment of ObservationSP since proximity was observable only in this treatment. Robust standard errors reported in parentheses (standard errors clustered at the group level). Column 1 uses observations from groups where all group members answered the knowledge question of Part I in the same way (either correctly or incorrectly). Column II uses observations from groups where group members gave different answers to the knowledge question. Controls are the same as in Table 2; see note to Table 2. Stars indicate significant differences at the conventional levels of * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

ObservationSP is indeed related to the degree of known social proximity.¹⁰

4.2.3. Heterogeneity in Norm Erosion

A final question that we address is whether initial dispositions towards norm compliance mediate the responsiveness to observation of others' behavior. From a policy perspective, this analysis is interesting because it not only helps to understand who is more susceptible to the effects of observation, but to what extent their behavior changes and contributes

¹⁰One could further distinguish between groups where all members gave a correct answer and groups where all members gave an incorrect answer. However, we have only a very small number of groups where all members gave a correct answer to the knowledge question. Because we do not have sufficient statistical power to draw reliable conclusions, we merely present an exploratory graphical representation in the Online Appendix Figures A.3 and A.4. The suggestive evidence is in line with our theoretical discussion.

to the erosion of norm compliance. Based on the donation behavior in Part II of the experiment, we divided our participants into three groups: norm followers, violators, and abstainers. As previously shown in Table 1, between 16% – 19% of participants can be classified as norm followers, 42% – 45% as norm violators, and 36% – 41% as abstainers.

To examine how each group of subjects is differently affected by observation, we consider their behavior across three distinct phases of the experiment: Part II, where there was no information about others’ behavior, and periods 1-10 and 11-19 of Part III, where they continuously received empirical information about others’ ToG decisions in their group. Figure 4 plots, separately for each type of subject and treatment, the distribution of donations made in these three phases of the experiment depicted in blue, green, and red, respectively. The vertical dotted lines in each panel indicate the average donation behavior and the colors correspond to the respective phase (for a different breakdown of the data see Figure A.6 in the Online Appendix).

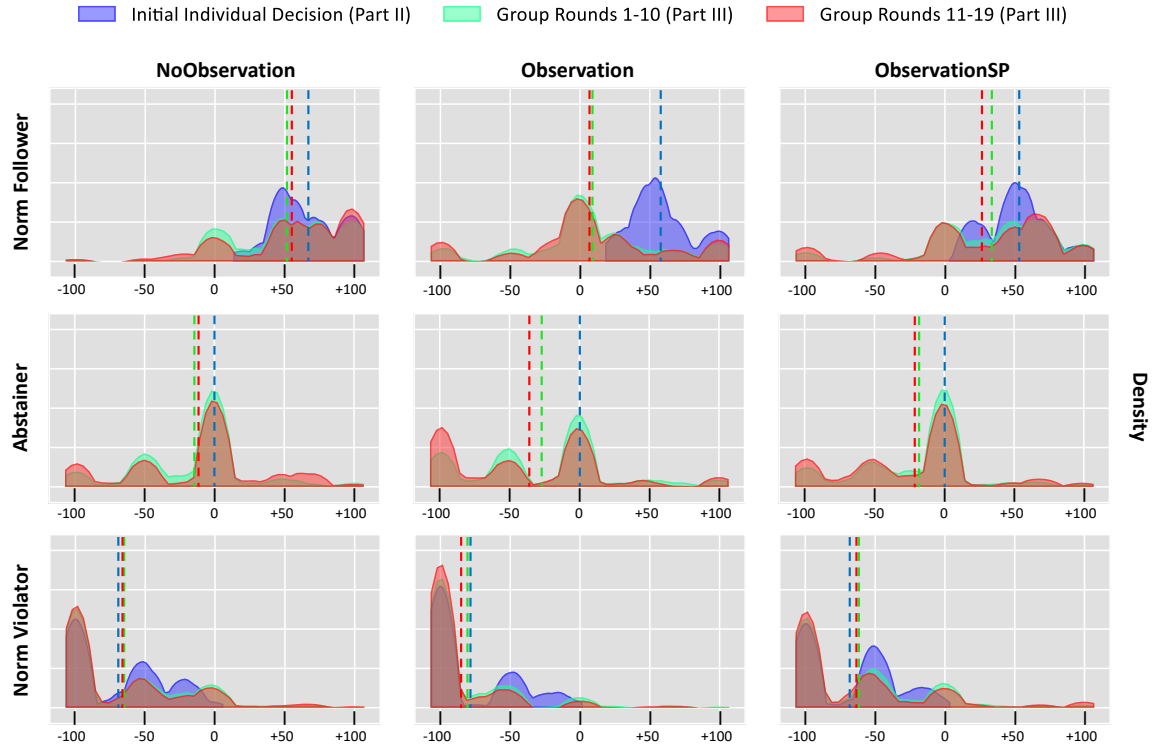


Figure 4: Shown are the distribution of donations (density) disaggregated by group of subject (norm violators, abstainers, norm followers, as defined by Part II initial behavior) and treatment. Vertical lines represent averages. Colors correspond to the respective phases. Blue: donations in Part II. Green: average donations across periods 1-10 of Part III. Red: average donations across periods 11-19 of Part III.

We have a number of interesting results. In NoObservation (leftmost column of Figure 4), most groups only make small changes to their behavior between Part II and III of the experiment, which is to be expected given that participants neither observe nor are being observed by others. Norm followers give on average 65.3 ECUs to charity in Part II and 52.2 ECUs in Part III, a decrease of 13.1 ECUs that is significant at the 5 percent level ($p = 0.013$). Abstainers, by definition, give 0 to the charity in Phase II of the experiment. In Phase III their average donation drops to -13.0 ECUs, and the difference is statistically significant ($p = 0.005$). Finally, norm violators take on average 69.3 ECUs from charity in Part II, while the average amount taken across the 19 periods of Part III is 64.9 ECUs. This change of 4.4 ECUs is not statistically significant ($p = 0.140$). The change in behavior of norm followers and abstainers is larger than that of norm violators (both $p < 0.005$), while we find no difference between norm followers and abstainers ($p = 0.994$).¹¹

In Observation (middle column of Figure 4), donations drop more dramatically. Again we find the largest changes in behavior for norm followers and abstainers, while norm violators only minimally adjust their choices. Norm violators' donations drop from -78.9 ECUs in Part II to -82.7 ECUs in Part III (a statistically insignificant decrease of -3.8 ECUs, $p = 0.220$). Norm followers, instead, reduce their donations from +58.8 ECUs in Part II to +7.6 ECUs in Part III (a drop of -51.2 ECUs), and abstainers reduce their donations by -30.8 ECUs. Both effects are highly significant ($p < 0.001$), and different from the change in behavior of violators (both $p < 0.001$). We also find that the drop in donations of norm followers is significantly larger than that of abstainers ($p = 0.011$).

We find similar effects in ObservationSP (rightmost column of Figure 4), albeit smaller in magnitude. For norm followers, donations drop from +51.9 ECUs in Part II to +30.2 ECUs in Part III, a change of -21.7 ECUs that is highly significant ($p < 0.001$). Abstainers significantly reduce their donation by -20.1 ECUs ($p < 0.001$). For norm violators, donations go from -67.9 ECUs in Part II to -61.0 ECUs in Part III. This change of 6.9 ECUs is statistically significant ($p = 0.009$). Again, we find that the change in donations

¹¹To assess the statistical significance of the changes in behavior between Part II and III, we use OLS regressions, conducted separately for each treatment. The dependent variable measures, for each subject, the difference between the donation made in Part II and the average donation made in Part III (thus, the subject is the unit of observation in this analysis). We regress this on group dummies, and test whether, for each group, the estimated coefficient is significantly different from zero. The standard errors are clustered at the group level to account for potential interdependencies between subjects of the same group. We report two-sided p-values adjusted for multiple-testing using the FDR correction (Benjamini and Hochberg, 1995).

of followers and abstainers is larger than that of violators ($p < 0.001$), but find no difference between followers and abstainers ($p = 0.764$). Initial norm followers display the largest changes in behavior between Part II and III of the experiment, followed by initial abstainers, and then initial norm violators. The reason for this is straightforward: based on Figure 3, norm followers experienced the largest discrepancy between their own behavior and the behavior of their group members. In contrast, norm violators observed the smallest difference. This is in line with our framework presented in Section 2: we expect norm followers to be those who would have to make the largest (downwards) revisions to their empirical expectations after receiving information about how other members of their group actually behaved. Since they have a conditional preference for following the norm, their behavior is strongly influenced by what they observe. Our assertion that, in light of observing others' behavior, the observed behavioral change – and subsequent erosion of norm compliance – can be attributed the updating of normative expectations rather than to the updating of personal normative beliefs is further corroborated by the findings of [Bicchieri et al. \(2019a\)](#) and [Gächter et al. \(2019\)](#).

5. Discussion and Conclusion

Social norms are a fundamental component of social and economic life. Therefore it is important to study conditions under which norm compliance obtains. In this paper we focused on how observing others' norm compliance or norm violation influences own norm compliance. To investigate norm compliance, we adopted the theory of social norms developed by [Bicchieri \(2006, 2016\)](#) and designed a novel, non-strategic Take-or-Give donation game, where people could give to charity, take from it, or abstain from changing the initial allocation between self and the charity. We first established that most people think taking from the charity is socially inappropriate, whereas giving to the charity is the social norm. Our behavioral results reveal a strong asymmetry in norm compliance: observing an anonymous other violating the norm, that is, taking from the charity, increased the likelihood that people violate the norm as well; observing norm compliance (donating to the charity), however, did not increase donations to the charity. In sum, observing norm violations by anonymous others eroded norm compliance in a way that was not compensated by observing norm compliance. No erosion occurred when others' behavior was unobservable.

While this asymmetry in reactions paints a bleak picture about norm compliance when anonymous others can be observed, in most real-world interactions people have clues about

their social proximity to the people they observe. Our experiment accounted for this fact by designing a treatment where a very minimal cue about social proximity, fanhood of a sports team, had a very strong effect in reducing the asymmetry in reactions to observing norm violations and norm compliance, respectively. Under social proximity, people not only paid attention to norm violations, but were also willing to emulate norm compliant behavior, thus halting norm erosion.

Our results on the importance of social proximity for norm compliance are in line with a large literature that shows significant differences in behavior towards socially proximate as opposed to socially distant others (“in-groups” vs. “out-groups”). For instance, individuals are more altruistic, trusting, and trustworthy towards in-groups rather than out-groups (McEvily et al., 2006; Balliet et al., 2014); more likely to comply with requests by an in-group rather than an out-group member (e.g., Burger et al., 2004); and more likely to take an in-group’s advice into account when judging one’s own actions compared to an out-group’s (Gino et al., 2009b). Our contribution to this literature is to show that social proximity can also stabilize norm compliance, by reducing the asymmetry in reactions to observations of norm violations and norm compliance, respectively. Punishment, often seen as an important stabilizer of social norms, may not be necessary when social proximity induces emulation of norm compliance.

Our results have two key implications for understanding the dynamics of norm compliance. First, in order to predict norm compliance, it is essential to measure not only what people regard as appropriate behavior in a given situation (e.g., by using the norm-elicitation technique by Krupka and Weber, 2013), but also to gauge the extent to which people believe others will actually engage in that behavior. Building on Bicchieri and Chavez (2010), in this paper we propose a set of belief-elicitation measures to accomplish this. Second, our results point to the importance of integrating the broader social context in the study of social norm compliance. Most existing experimental research studies norms in abstract, anonymous, and context-neutral decision settings. While using contextually neutral decision environments is one of the hallmarks of experimental control, our study shows that this comes at the cost of missing important insights into the drivers of norm compliance. For example, our results show that one would draw substantially different conclusions about the effects of observation on norm compliance depending on whether interactions occur among strangers or socially proximate individuals.

Our results have implications for behavioral change interventions (for a discussion, see Bicchieri and Dimant, 2019). Most social encounters in natural and online environments

involve observation of shared characteristics, thereby inducing social categorization and group identity. Information about others' behavior will have a more balanced effect in natural settings where signals of social proximity are available, compared to our Observation treatment that deliberately suppresses any signal of social proximity. Effective policy interventions should take into consideration that people are more likely to pay attention to norm compliance by similar others, but not by anonymous or otherwise generic others.

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. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.
- Allen, V. L. (1965). Conformity and the role of deviant. *Journal of Personality*, 33(4):584–597.
- Arechar, A. A., Gächter, S., and Molleman, L. (2018). Conducting interactive experiments online. *Experimental Economics*, 21(1):99–131.
- Azrieli, Y., Chambers, C. P., and Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4):1472–1503.
- Balliet, D., Wu, J., and De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, 140(6):1556.
- Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, 11(2):122–133.
- Barr, A., Lane, T., and Nosenzo, D. (2018). On the social inappropriateness of discrimination. *Journal of Public Economics*, 164:153–164.
- Bašić, Z., Falk, A., and Kosse, F. (2019). The development of egalitarian norm enforcement in childhood and adolescence. *Journal of Economic Behavior & Organization*.
- 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: The nature and dynamics of social norms*. Cambridge University Press.
- Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.
- Bicchieri, C. and Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making*, 23(2):161–178.
- Bicchieri, C. and Dimant, E. (2019). Nudging with care: The risks and benefits of social information. *Public Choice*.
- Bicchieri, C., Dimant, E., and Sonderegger, S. (2019a). It’s not a lie if you believe it: On norms, lying, and self-serving belief distortion. Working Paper Available at SSRN: <https://dx.doi.org/10.2139/ssrn.3326146>.
- Bicchieri, C., Dimant, E., and Xiao, E. (2019b). Deviant or wrong? The effects of norm information on the efficacy of punishment. Working Paper Available at SSRN: <https://dx.doi.org/10.2139/ssrn.3294371>.
- 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 Zeckhauser, R. (2004). Social comparisons in ultimatum bargaining. *Scandinavian Journal of Economics*, 106(3):495–510.
- Bolton, G., Dimant, E., and Schmidt, U. (2019). When a nudge backfires: Using observation with social and economic incentives to promote pro-social behavior. Working Paper Available at SSRN: <https://dx.doi.org/10.2139/ssrn.3294375>.
- Bond, R. (2005). Group size and conformity. *Group Processes & Intergroup Relations*, 8(4):331–354.
- Bott, K. M., Cappelen, A. W., Sorensen, E., and Tungodden, B. (2019). You’ve got mail: A randomised field experiment on tax evasion. *Management Science*.
- Brennan, G., Pettit, P., et al. (2004). *The economy of esteem: An essay on civil and political society*. OUP Oxford.
- Burger, J. M., Messian, N., Patel, S., del Prado, A., and Anderson, C. (2004). What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin*, 30(1):35–43.
- Burke, M. A. and Young, H. P. (2011). Social norms. In *Handbook of Social Economics*, volume 1, pages 311–338. Elsevier.
- Bursztyn, L., Egorov, G., and Fiorin, S. (2019). From extreme to mainstream: How social norms unravel. NBER Working Paper.
- Centola, D. (2018). *How behavior spreads: The science of complex contagions*, volume 3. Princeton University Press.
- 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.
- Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1):431–57.
- Chyn, E. (2018). Moved to opportunity: The long-run effects of public housing demolition on children. *American Economic Review*, 108(10):3028–56.
- Cialdini, R. B. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika*, 72(2):263.
- Cialdini, R. B. and Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55:591–621.
- Coleman, J. S. (1994). *Foundations of social theory*. Harvard University Press.
- Croson, R. and Shang, J. Y. (2008). The impact of downward social information on contribution decisions. *Experimental Economics*, 11(3):221–233.
- DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics*, 127(1):1–56.
- Dimant, E. (2019). Contagion of pro-and anti-social behavior among peers and the role of social proximity. *Journal of Economic Psychology*, 73:66–88.

- Dimant, E., Gerben, A. v. K., and Shalvi, S. (2019). Requiem for a nudge: Framing effects in nudging honesty. Working Paper Available at SSRN: <https://dx.doi.org/10.2139/ssrn.3416399>.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.
- Eckel, C. C. and Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, 16(2):181–191.
- Elster, J. (1989). *The cement of society: A survey of social order*. Cambridge University Press.
- Exley, C. L. (2015). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, 83(2):587–628.
- Fehr, E. and Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14(3):159–181.
- Fehr, E. and Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, 2(7):458.
- Fellner, G., Sausgruber, R., and Traxler, C. (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association*, 11(3):634–660.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.
- Fischbacher, U. and Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–56.
- Fisman, R. and Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy*, 115(6):1020–1048.
- Gächter, S., Gerhards, L., and Nosenzo, D. (2017). The importance of peers for compliance with norms of fair sharing. *European Economic Review*, 97:72–86.
- Gächter, S., Molleman, L., and Nosenzo, D. (2019). The behavioral logic of rule following and social norm compliance. Unpublished Manuscript.
- 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.
- Gerber, A. S., Green, D. P., and Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, 102(1):33–48.
- Gino, F., Ayal, S., and Ariely, D. (2009a). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science*, 20(3):393–398.
- Gino, F., Shang, J., and Croson, R. (2009b). The impact of information from similar or different advisors on judgment. *Organizational Behavior and Human Decision Processes*, 108(2):287–302.
- Goldin, C. and Katz, L. F. (2002). The power of the pill: Oral contraceptives and women’s career and marriage decisions. *Journal of Political Economy*, 110(4):730–770.

- Goldstein, N. J., Cialdini, R. B., and Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35(3):472–482.
- Grossman, P. J. and Eckel, C. C. (2015). Giving versus taking for a cause. *Economics Letters*, 132:28–30.
- 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.
- Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. Harvester Wheatsheaf.
- Jachimowicz, J. M., Hauser, O. P., O’Brien, J. D., Sherman, E., and Galinsky, A. D. (2018). The critical role of second-order normative beliefs in predicting energy conservation. *Nature Human Behaviour*, 2(10):757.
- Keizer, K., Lindenberg, S., and Steg, L. (2008). The spreading of disorder. *Science*, 322(5908):1681–1685.
- Kimbrough, E. O. and Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, 14(3):608–638.
- 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.
- Knack, S. (1992). Civic norms, social sanctions, and voter turnout. *Rationality and Society*, 4(2):133–156.
- Kocher, M. G., Schudy, S., and Spantig, L. (2017). I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups. *Management Science*, 64(9):3995–4008.
- Korenok, O., Millner, E. L., and Razzolini, L. (2014). Taking, giving, and impure altruism in dictator games. *Experimental Economics*, 17(3):488–500.
- Krupka, E. L. and Croson, R. T. (2016). The differential impact of social norms cues on charitable contributions. *Journal of Economic Behavior & Organization*, 128:149–158.
- Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.
- Lindbeck, A., Nyberg, S., and Weibull, J. W. (1999). Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics*, 114(1):1–35.
- Lindström, B., Jangard, S., Selbing, I., and Olsson, A. (2018). The role of a “common is moral” heuristic in the stability and change of moral norms. *Journal of Experimental Psychology: General*, 147(2):228.
- List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493.
- McEvily, B., Weber, R. A., Bicchieri, C., and Ho, V. T. (2006). Can groups be trusted? an experimental study of trust in collective entities. *Handbook of Trust Research*, pages 52–67.

- Muthukrishna, M., Francois, P., Pourahmadi, S., and Henrich, J. (2017). Corrupting cooperation and how anti-corruption strategies may backfire. *Nature Human Behaviour*, 1(7):0138.
- Neugebauer, T., Perote, J., Schmidt, U., and Loos, M. (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology*, 30(1):52–60.
- Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, 14(3):137–158.
- Perkins, H. W. (2002). Social norms and the prevention of alcohol misuse in collegiate contexts. *Journal of Studies on Alcohol, supplement*, (14):164–172.
- Pryor, C., Perfors, A., and Howe, P. D. (2019). Even arbitrary norms influence moral decision-making. *Nature Human Behaviour*, 3(1):57–62.
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5):429–434.
- Smerdon, D., Offerman, T., and Gneezy, U. (2019). Everybody’s doing it: On the emergence and persistence of bad social norms. *Experimental Economics*.
- Snowberg, E. and Yariv, L. (2018). Testing the waters: Behavior across participant pools. Working Paper 24781, National Bureau of Economic Research.
- Strathman, A., Gleicher, F., Boninger, D. S., and Edwards, C. S. (1994). The consideration of future consequences: weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology*, 66(4):742.
- Sugden, R. (1998). Normative expectations: the simultaneous evolution of institutions and norms. In Ben-Ner, A. and Putterman, L., editors, *Economics, Values, and Organization*, page 73–100. Cambridge University Press, Cambridge.
- Sugden, R. (2000). Team preferences. *Economics & Philosophy*, 16(2):175–204.
- Sutter, M., Haigner, S., and Kocher, M. G. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *The Review of Economic Studies*, 77(4):1540–1566.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33(1):1–39.
- Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations*, 33(47):56–65.
- Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High Self-Control Predicts Good Adjustment, Less Pathology, Better Grades, and Interpersonal Success. *Journal of Personality*, 72(2):271–324.
- 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.
- Turner, J. C. (1985). Social categorization and the self concept: A social cognitive theory of group behavior. In Lawler, E. J., editor, *Advances in Group Process*, volume 2, pages 77–122. JAI, Greenwich, CT.

Online Appendix to “Observability, Social Proximity, and the Erosion of Norm Compliance”

Cristina Bicchieri, Eugen Dimant, Simon Gächter, Daniele Nosenzo

Contents

- I. Robustness Checks
 - II. Laboratory Experiment
 - II.a. Instructions
 - II.b. Experimental Screenshots
 - III. Norm Elicitation on MTurk
 - III.a. Norm Elicitation Following [Bicchieri and Chavez \(2010\)](#)
 - III.b. Norm Elicitation Following [Krupka and Weber \(2013\)](#)
-

I. Robustness Checks

I.a. Complementary Normative Expectations Elicitation Procedure

As an additional approach to elicit normative expectations, we follow the recent technique introduced in the economics literature by [Krupka and Weber \(2013\)](#). Subjects ($N = 152$) were given a description of the ToG donation game and were asked to rate the social appropriateness of various possible behaviors in the game. In the original version of the procedure introduced by [Krupka and Weber \(2013\)](#), subjects are asked to rate each action available to the decision maker in the game. In our case, however, this would be too demanding since the game involves a very large action space. As in the tasks used in the main text, we therefore asked subjects to rate three types of behaviors that constitute natural partitions of the action space in the game: *take* money from the charity; *abstain* (leave initial endowments unchanged); or *give* money to the charity.

For each behavior, subjects were asked to report whether they perceived it as socially appropriate or socially inappropriate using a 4-point scale (very inappropriate; somewhat inappropriate; somewhat appropriate; very appropriate). Importantly, subjects were not asked to report their own perception about the appropriateness of the behavior, but rather their belief about the shared perception of appropriateness among individuals completing the same task as themselves. That is, subjects were asked to report their beliefs about what is commonly perceived as appropriate behavior in the game, rather than what they personally perceive as appropriate or inappropriate. We interpret this belief as measuring a subject’s normative expectation about behavior in the game. In order to incentivize the elicitation of normative expectations, subjects were paid \$0.25 if their rating of appropriateness for a behavior matched the modal response given by other participants.

We present the results of the elicitation in Figure A.1, where we plot, for each behavior, the average appropriateness rating given by subjects in the experiment (where a rating of -1 means "very inappropriate" and +1 "very appropriate"). For comparison, Figure A.1 also reproduces the results of our normative expectations elicitation presented in the main text (representing the distribution of second-order beliefs about what is considered normative behavior). Figure A.1 shows that, regardless of the method used to elicit normative expectations, one arrives at a similar conclusion about the prevailing social norm in the ToG donation game. In both cases, giving money to the charity is rated as the most appropriate behavior / what one ought to do, while taking money away from the charity is the least appropriate behavior.

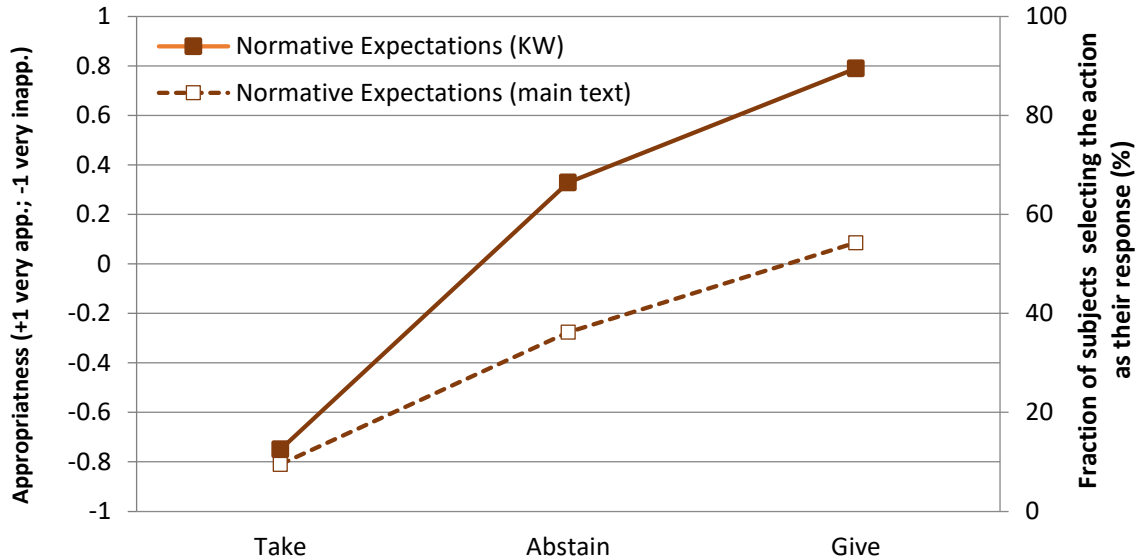


Figure A.1: Normative expectations elicited following [Krupka and Weber \(2013\)](#) (left axis) and our procedure presented in the main text (right axis).

I.c. Additional Analyses of Norm Compliance Erosion

Here we present an additional analysis to examine the erosion of norm compliance for those subgroups that have not started with a majority of taking behavior. Our findings strongly support our main results: compared to NoObservation, even when taking behavior is in the minority from the beginning, the erosion of norm compliance accelerates in the Observation condition ($p=0.0201$), while being substantially muted in the presence of social proximity in ObservationSP ($p=0.8731$ compared to NoObservation; $p<0.01$ compared to Observation).

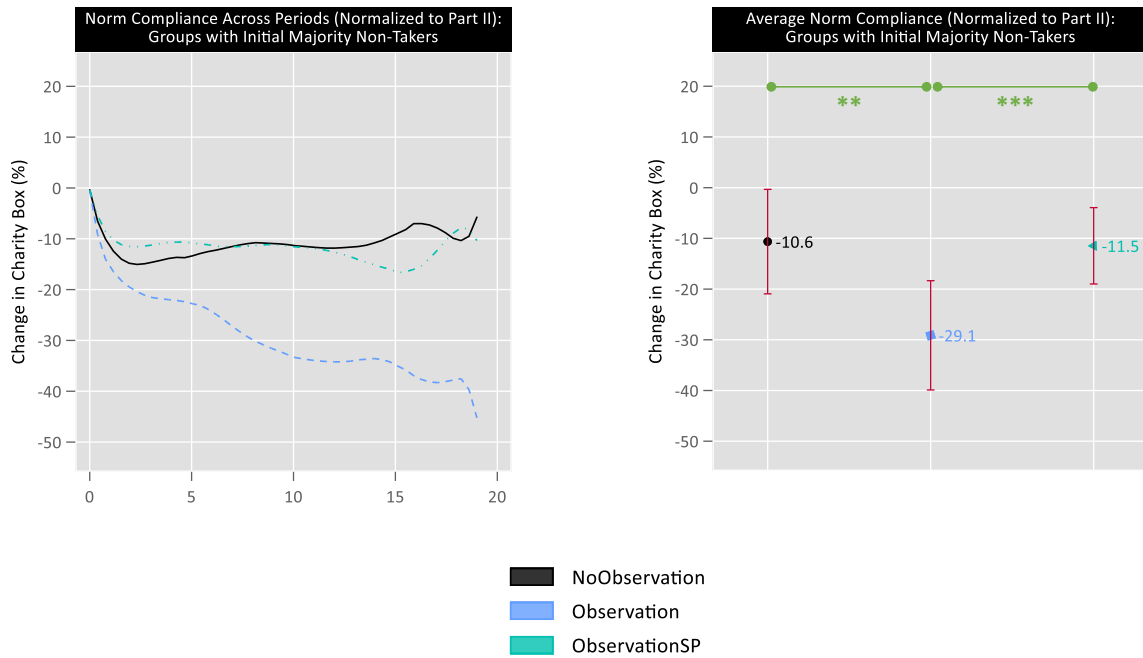


Figure A.2: Left panel: Subject’s behavior across treatments and periods in Part III for groups in which the initial average behavior was not taking, normalized to average behavior in Part II (period 0 in the graph). For purposes of readability, lines represent kernel-weighted local polynomial smoothing of degree 5. Right panel: Average subject’s donations across treatments, averaged over the 19 periods of Part III. Whiskers indicate 95% confidence intervals. Stars indicate significant differences using two-sided Mann-Whitney U-tests and group averages as observations, with p-values adjusted for a false discovery rate (Benjamini and Hochberg, 1995) at the conventional levels of * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

I.c. Group Proximity

Here we present graphical illustrations of the evolution of norm compliance in ObservationSP. Figure A.3 shows this for proximate groups, where all group members answered the proximity question in the same way (either all correctly or all incorrectly), and non-proximate groups (mixed answers within the same group). Figure A.4 breaks down the data for the proximate groups in one additional way: groups where all answers to the proximity questions were correct compared to groups where all answers were incorrect.

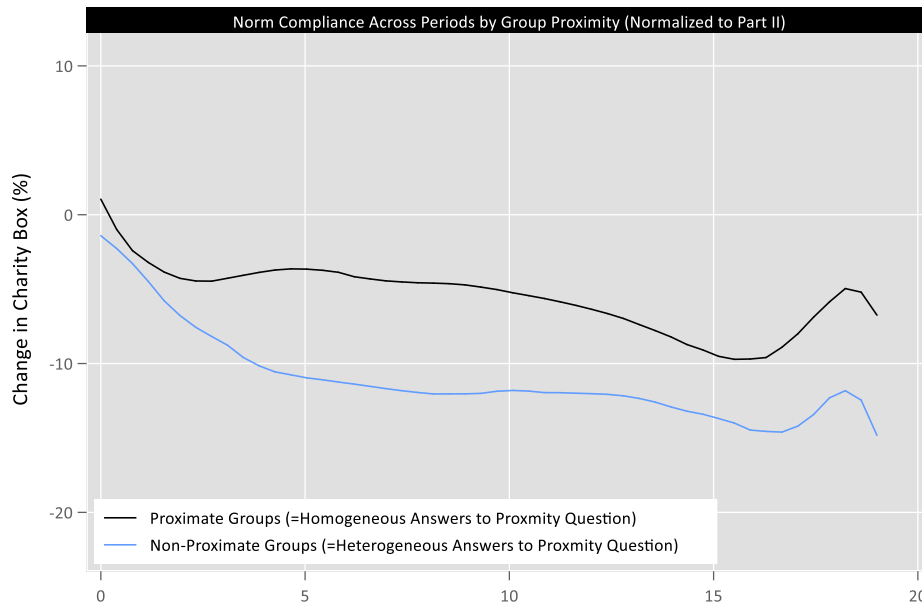


Figure A.3: Evolution of norm compliance (normalized to Part II behavior) in ObservationSP for groups answering the proximity question in the same way (proximate) or in different ways (non-proximate). Polynomial smoothing applied.

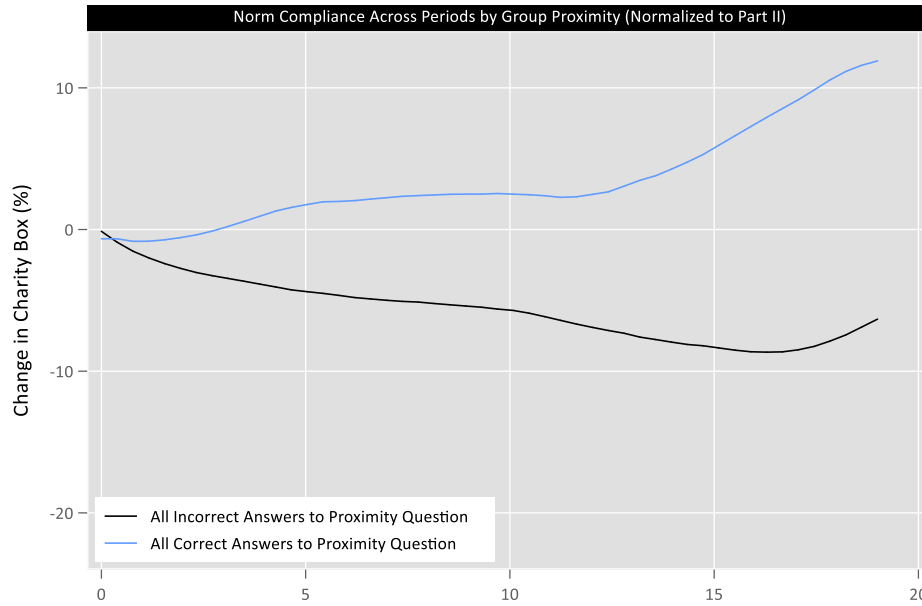


Figure A.4: Evolution of norm compliance (normalized to Part II behavior) in ObservationSP for groups answering the proximity question correctly or incorrectly. Polynomial smoothing applied.

I.d. Additional Heterogeneity Analysis

One additional way in which we can analyze the data from Section 4.2.3 is by looking at the stability of behavior across treatments. For this purpose, we use the same classification of initial behavior (initial norm violators, followers, abstainers) and examine the rate at which individuals switched from one category to the other across the different treatments. That is, conditional on the classification based on one’s initial behavior, the following analysis sheds light on whether the frequency at which individuals switch to a different classification (e.g. from norm follower to norm violator) depends on the availability of information about others’ behavior and proximity. The calculation is straightforward: for each period in the group stage, a participant’s behavior is coded as ‘1’ (‘0’) if the type of behavior (violator / follower / abstainer) was different (the same) as that participant’s behavior in the initial donation decision in Part II. For each individual, we then average this sequence of 1s and 0s across all 19 periods, resulting in a single value per individual.

We present the results in Figure A.5. Our results show that individuals are substantially more likely to change behavior in Observation, which is highly statistically significantly different compared to NoObservation (test of proportions, $p < 0.01$) and weakly significantly different from ObservationSP (test of proportions, $p = 0.089$). The rate at which behavior changed in NoObservation and ObservationSP is not significantly different (test of proportions, $p = 0.12$). As indicated in the bottom panel, this is mainly driven by the initial norm followers and norm abstainers, whereas the likelihood that individuals who were initially violators change their type of behavior does not vary significantly across treatments (test of proportions, all $p > 0.16$).

In more detail, initial abstainers are substantially more likely to switch to a different behavior classification in Observation compared to both NoObservation ($p = 0.02$) and ObservationSP ($p < 0.01$), whereas again the behavior in the latter two conditions is not statistically significantly different from each other ($p = 0.83$). Similarly, initial followers are substantially more likely to switch behavior in Observation condition compared to both NoObservation ($p < 0.01$) and ObservationSP ($p = 0.011$). Here, behavioral change in ObservationSP is marginally more likely than in NoObservation ($p = 0.098$).

Overall, these results suggest that, with information about others’ behavior, but not others’ proximity, individuals are strongly influenced by feedback about how others behave and are substantially more likely to change their initial behavior. On the other hand, adding social proximity stabilizes behavioral change at levels similar to an environment devoid of information about others’ behavior. At the same time, initial norm violators seem to be the most resistant to switching behavior and remain consistent throughout the game.

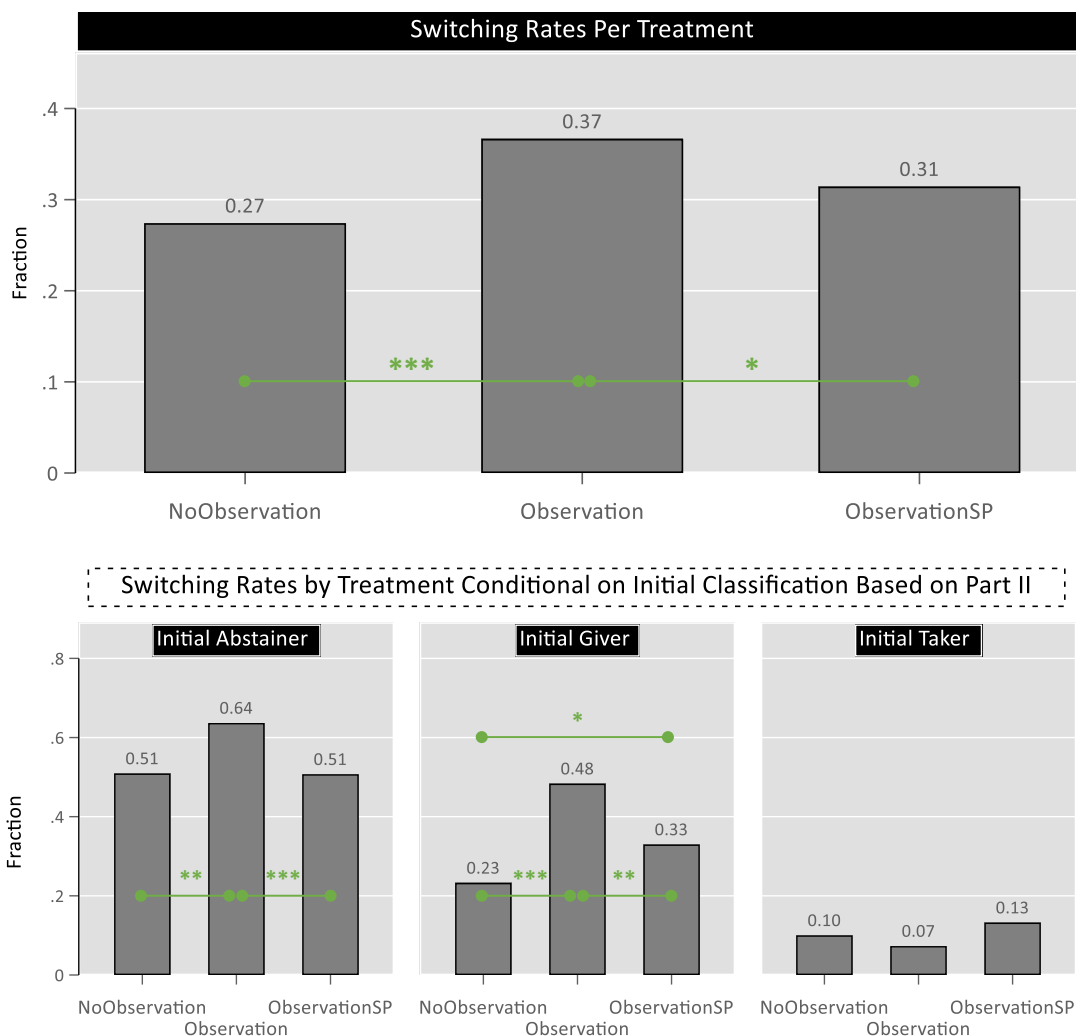


Figure A.5: Fraction of individuals switching their initial behavioral classification pooled across treatments (top panel) and conditional on the individual's initial behavior based on Part II (bottom panel). Stars indicate significant differences at the conventional levels of $*p < 0.1$, $**p < 0.05$, and $***p < 0.01$.

We can substantiate our results from Figure 4 by presenting a more detailed breakdown of changes in norm compliance across treatments and individual types. In line with the previously discussed results, the initial norm followers display the largest changes in behavior, followed by initial abstainers and initial norm violators. We present the results in Figure A.6.

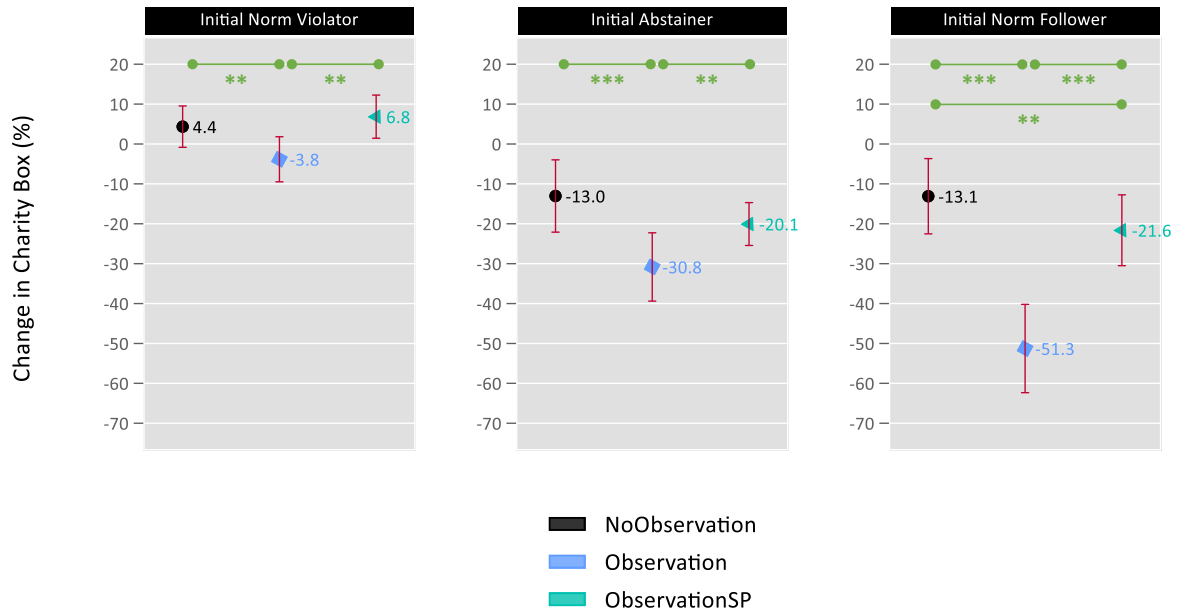


Figure A.6: Fraction of individuals switching their initial behavioral classification conditional on the individual's initial behavior based on Part II. Whiskers indicate 95% confidence intervals. Stars indicate significant differences at the conventional levels of * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$.

II. Laboratory Experiment

II.a. Instructions

Below, we present the instructions for all our treatment variations. Recall that the instructions for the first decision were the same across all treatments and only the instructions for the subsequent group stage differed. We highlight the text in green that was used instead of the preceding numbers and sentences between the NoObservation and the Observation and ObservationSP treatments. For the latter, the instructions were the same. Information about the social proximity was not pre-announced but simply displayed on the computer screen (see screenshots in Appendix III).

Instructions for First Decision

General

Thank you for coming! You have earned \$10 for showing up on time. The following instructions explain how you can potentially earn more money by making a number of decisions. To maximize your chances to earn more money, please read these instructions carefully! If you have a question at any time, please raise your hand, and an experimenter will assist you. For the purpose of the experiment, it is important that you do not talk or communicate in other ways with the other participants. Please turn off or silence your cell phone and all other electronic devices. You are asked to abide by these rules. If you do not abide, we will have to exclude you from this, and future, experiments (with the current experimenter) and you will not receive any compensation for the current experiment. Your decisions will remain anonymous to any other participants throughout the experiment. No participant will know who has made what decisions. Please do not talk to each other during the experiment. During the experiment, all amounts will be presented in ECU (Experimental Currency Units). At the end of the experiment, the ECU you have earned will be converted to Dollars as follows:

$$10 \text{ ECU} = 1 \text{ Dollar}$$

Roles in the Experiment

- There are two roles in the experiment:
 1. Decision Maker
 2. Receiver
- Each participant in the experiment plays in the role of the decision maker.
- The role of the receiver is represented by one of the following three charities:
 1. Doctors Without Borders
 2. World Wildlife Fund (WWF)

3. UNICEF

Order of Events

- You will make multiple decisions in this experiment. Explanations and information related to these decisions will be given at the relevant points throughout the experiment.
- Both you as well as the receiver in the form of a charitable organization will be provisionally assigned a monetary amount of 100 ECU each.
- The decision maker will have to decide whether to...
 - ... take a part or all of the money from the receiver.
 - ... leave the division of the sum of money as it is.
 - ... give a part or all of his/her money to the receiver.
- At the end of the experiment, one decision maker will be chosen at random, exactly one decision of this decision maker will be implemented, and count towards his/her own payoff as well as the receiver. The randomly selected participant at the end of the experiment will determine the receiving charity.
 - The decision of every other decision maker will not count towards his or her own and the receiver's payoff. Instead, each of these decision makers will receive exactly 100 ECU (their show-up fee) regardless of their allocation decision towards the receiver.
- The ECU amount the receiving charity is left with at the end will be doubled (meaning that the multiplier equals 2) and actually transferred to the respective charity. We will upload a receipt of the donation to our website (<http://www.sas.upenn.edu/ppe/>) where you can verify the validity of the donation.

Payoff Structure

- The total payoff of the participants:
 - **If you are the randomly chosen decision maker**
 - * 100 ECU (show-up fee) +/- the amount of money that was given to / taken from the receiver
 - **If you are not the randomly chosen decision maker**
 - * 100 ECU (show-up fee)
- The total payoff of the receiving charity:
 - (100 ECU +/- amount of money that was given to / taken from the charity by the randomly chosen decision maker)x2

Instructions for Group Stage (Distributed After Initial Decision)

We highlight the treatment differences in green

General

- Every decision maker is now randomly paired up with three other decision makers in this room. That is, each group consists of exactly 4 (2) decision makers.
- This grouping will not change over the course of the next 19 rounds, meaning you will remain paired up with the same 3 (1) decision makers from now until the end of the experiment. The experiment ends after round 20.
- Over the course of the next 19 rounds, every decision maker continues to make the same decisions as before. That is, the decision maker will have to decide whether to...
 - ... take a part or all of the money from the receiver.
 - ... leave the division of the sum of money as it is.
 - ... give a part or all of his / her money to the receiver.
- The decisions that are made over the course of the next 19 periods will remain anonymous. This means that individual behavior will not be displayed to any participant within your group or anyone else. (At the end of each round, every decision maker's decision will be revealed to every other decision maker inside the group. The decisions being made over the course of the next 19 rounds will be displayed within each group, but not to anyone else.)

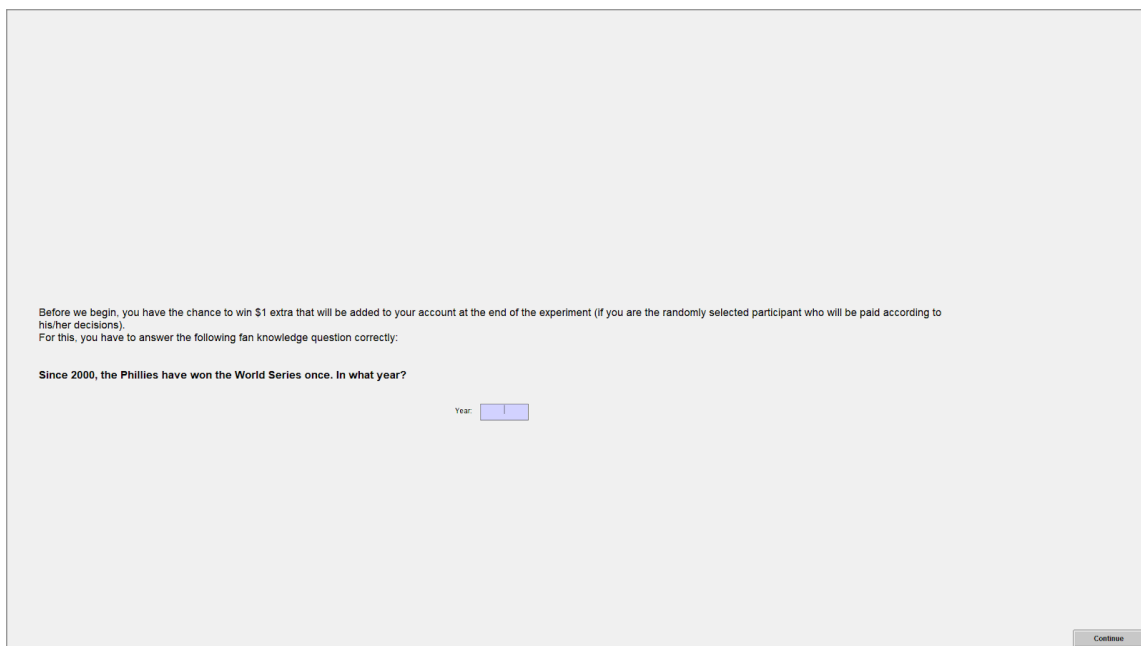
Payoff

- The payoff mechanism for this part of the experiment is exactly the same as for the first part of the experiment. That is (as a reminder):
 - At the end of the experiment, one decision maker will be chosen at random, exactly one decision of this decision maker will be implemented, and count towards his/her own payoff as well as the receiving donation.
 - The decisions of every other decision maker will not count towards his or her own and the receiver's payoff. Instead, each of these decision makers will receive exactly 100 ECU regardless of their allocation decision towards the receiver.
 - The final ECU amount left with the receiving charity will be doubled and actually transferred to the respective charity. We will upload a receipt of the donation to our PPE website (<http://www.sas.upenn.edu/ppe/>) where you can verify the validity of the donation.

- The decisions that are made over the course of the next 19 rounds will remain anonymous. This means that individual behavior will not be displayed to any participant within your group or anyone else. (At the end of each round, every decision maker's decision will be revealed to every other decision maker inside the group. The decisions being made over the course of the next 19 rounds will be displayed within each group, but not to anyone else.)
-

II.b. Experimental Screenshots

Here, we exemplarily present the screenshots for *ObservationSP* with group size 4. Screenshots for other treatments are available upon request. Merely Figure A5 changed across the different conditions: in *NoObservation*, participants did not see any information in other group members' history table. In the *Observation* condition, participants observed the behavior of all participants in their group for all periods. For *ObservationSP* (as displayed in Figure A5), participants observed the behavior of all participants in their group for all periods plus the information as to whether or not the respective participant answered the knowledge question correctly.



Screen 1: Knowledge question (correct answer: 2008).

Decision:

You now have to decide whether you want to...

- ...take away money from the charity.
- ...not change the initial 50/50 split between you and the charity.
- ...give money to the charity.

Screen 2a: Decision towards charity (choosing the type of behavior).

I earn a total of (in ECU):	150
Change in my personal account (in ECU):	50
Change in the charity account (in ECU):	-50
Total earning of the charity (100 ECU +/- your decision) x 2, in ECU:	100

Decision:

You now have to decide whether you want to...

- ...take away money from the charity.
- ...not change the initial 50/50 split between you and the charity.
- ...give money to the charity.

Take

Amount:

How many of the 100 ECU do you want to take away from the charity?

In ECU:

Screen 2b: Decision towards charity (choosing the amount).

• Your Player ID is: 4
 • Explanation of values in column "Behavior": Took = Took away from charity; No Change = Did not change initial fair split; Gave = Gave money to charity

PlayerID: 1		
Period	Behavior	Amount
1	Took	50
2	Took	25
3	Took	50

PlayerID: 2		
Period	Behavior	Amount
1	Took	75
2	No change	0
3	Took	50

PlayerID: 3		
Period	Behavior	Amount
1	Gave	25
2	Gave	25
3	No change	0

PlayerID: 4		
Period	Behavior	Amount
1	No change	0
2	Took	33
3	Gave	33

Answer to knowledge question (correct answer = 2008):

correct correct incorrect incorrect

Continue

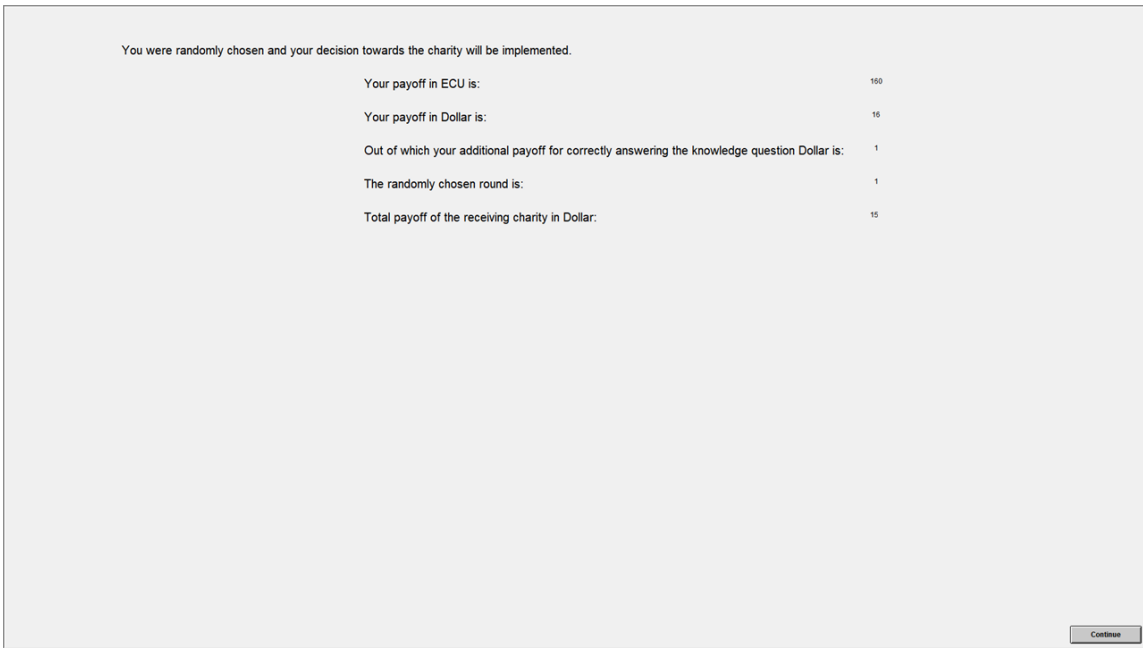
Screen 3: Observation stage (after 3 periods).

Please indicate which of the following three charities should receive the money (your choice will be implemented if you are the participant who is chosen at random at the end):

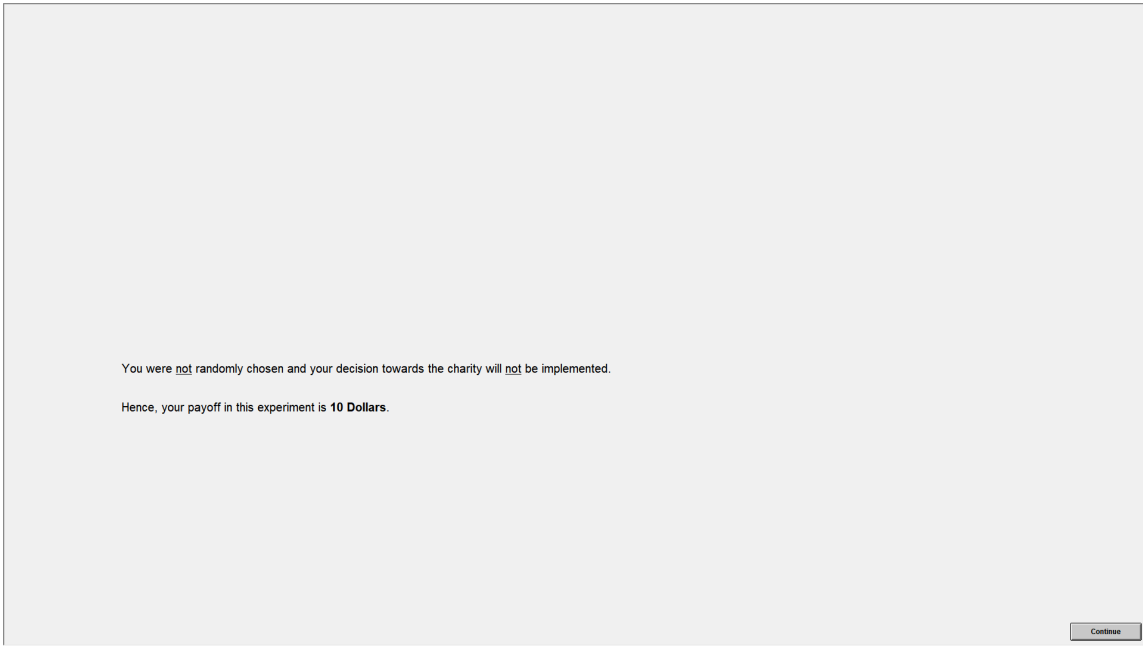
Doctors Without Borders
 United Nations Children's Fund (UNICEF)
 World Wide Fund for Nature (WWF)

Confirm

Screen 4: Observation stage (after 3 periods).



Screen 5: Final payoff screen (if participant was randomly chosen).



Screen 6: Final payoff screen (if participant was not randomly chosen).

III. Norm Elicitation on MTurk

Below, we present the screenshots of the norm elicitation experiments (all between-design) that we ran on MTurk. As indicated, the first three and the last 2 screens were the same across all elicitations. The other screens correspond to the respective norm elicitation treatment.

III.a. Norm Elicitation Following [Bicchieri and Chavez \(2010\)](#)

Consent Form (same for all elicitations)

University of Pennsylvania
Department of Philosophy, Politics, and Economics
Claudia Cohen Hall, Room 311
Philadelphia, PA 19104
Phone: (215)-898-3023
Fax: (215) 573-2231

Informed Consent/ Assent Form for Non-Pool Participants Earning Money

You are invited to take part in a study named *Guessing Task*. The purpose of this research study is to explore human decision-making. You will complete a series of computer tasks, each involving semantic as well as visual stimuli materials. If you agree to be in this study, you will need to make decisions and answer questions regarding the study materials. We will also ask you to provide demographic information. We will not ask for your name or any information that will make you identifiable. Overall, this study will take approximately 5-10 minutes.

For your participation in this study, you will receive a fixed payment of \$0.25. Additionally, you may receive a monetary bonus. The exact amount depends on your results in the experiment. The risks to participating are no greater than those encountered in everyday life. Your participation in this study is completely voluntary, and you may refuse to participate or withdraw from the study without penalty or loss of benefits to which you may otherwise be entitled. Compensation will be awarded upon completion of the entire study.

Results may include summary data, but you will never be identified. If you have any questions about this study, you may contact the Behavioral Ethics Lab at behavioralethicslab@gmail.com.

For any questions, concerns, suggestions, or complaints that are not being addressed by the researcher, please contact the Institutional Review Board at the University of Pennsylvania, 3624 Market Street, Suite 301 South Philadelphia, PA 19104-6006. Phone: (215) 898-2614.

Please feel free to print or save a copy of this consent form.

By continuing from this page you are indicating that you have read and understood this consent form and wish to continue your participation in this study.

Consent

●

General Instructions (same for all elicitations)

Thank you for choosing to participate in this study! It is important that you read all of the instructions carefully to maximize your earnings. This is a survey-based study where you will be asked to answer questions and complete simple tasks. At the end of the study, you will be asked to fill out a short questionnaire and demographic questions.

It is expected that this study will take approximately 5-10 minutes. There is no deception in this study. Everything you see or read is true.

You begin the survey with an endowment of \$0.25 (your fixed payment) and will have the opportunity to earn more based on your answers. You will be presented with a set of statements and asked to determine which actions you would be more likely to take after reading these statements. You will also be asked to guess what other participants have done previously. Guessing correctly will earn you an additional \$0.25 each.

You may stop participating in this study at any time; however, you will not receive your payment unless you complete the entire survey through to the end. You will receive a completion at the end of the survey.

Norm Elicitation 1 Following Bicchieri & Chavez (2010)

Scenario

In what follows, we describe some scenarios to you based on a series of prior actual experiments. **While you will be making hypothetical decisions, the scenarios below describe the original experimental setup truthfully.** In particular, this experiment involved real stakes, and the taking or donation decisions towards real charities were implemented and paid accordingly to both the participants and the respective charities.

Please read the subsequent statements carefully and answer the questions accordingly. Carefully reading the question will help you to make additional earning in this survey.

In this setup, **you and a charity** (either Doctors Without Borders, World Wildlife Fund (WWF), or UNICEF) **start with an equal endowment of \$10 each** (meaning that both you and the charity own \$10 at the beginning of the experiment).

You are in the role of a **decision maker**. As a decision maker, you would have to decide whether to...

- ...**take** a part or all of the money **from the charity** and add this amount to the decision maker's account.
- ...**leave the equal division** of the sum of money as it is.
- ...**give** a part or all of the decision maker's money **to the charity** and add this amount to the charity's account.

Three Belief Elicitations (Between-Design)

1. Personal Normative Beliefs (PNB) - Decision

Based on the scenario described above, please tell us what you think one should do?

- Leave the equal division** of the sum of money as it is.
- Give** a part or all of the money **to the charity**.
- Take** a part or all of the money **from the charity** and add this amount to the decision maker's account.



2. Normative Expectations (NE) - Decision

Please guess what you believe other participants who participated in the survey **said**. If you guess correctly what other participants said, you will receive an additional \$0.25.

Guess the correct statement: The **majority of participants in the survey said that one should...**

- Leave the equal division of the sum of money as it is.
- Take a part or all of the money from the charity.
- Give a part or all of the money to the charity.



3. Empirical Expectations (EE) – Decision

Please guess what you believe other participants who already participated in the experiment described above **did**. If you guess correctly what other participants who already participated did, you will receive an additional \$0.25.

Guess: **Based on the behavior of participants who previously played the game in the role of a decision maker with the charity, the...**

- Majority of participants left the equal division of the sum of money as it was.
- Majority of participants took a part or all of the money from the charity.
- Majority of participants gave a part or all the money to the charity.



Norm Elicitation 2 Following Krupka & Weber (2013)

In what follows, we describe some scenarios to you based on a series of actual prior experiments. The scenarios below describe the original experimental setup truthfully. In particular, this experiment involved real stakes and the taking or donation decisions towards real charities were implemented and paid accordingly to both the participants and the respective charities.

Please read the subsequent statements carefully and answer the questions accordingly. Carefully reading the question will help you to make additional money in this survey.

In this setup, you and a charity (either Doctors Without Borders, World Wildlife Fund (WWF), or UNICEF) start with an equal endowment of \$10 each (meaning that both you and the charity own \$10 at the beginning of the experiment).

You are in the role of a decision maker. As a decision maker, you would have to decide whether to...

- ...take a part or all of the money from the charity and add this amount to the decision maker's account.
- ...leave the equal division of the sum of money as it is.
- ...give a part or all of the decision maker's money to the charity and add this amount to the charity's account.

In what follows, you need to make payoff-relevant guesses.

Note: this survey will also be given to a number of other participants.

In order to earn additional money, you need to identify the answer that will be given the most often by the other participants. If you identify the most frequently given answers to the following statements, you will earn an additional \$0.25 each per correct statement.

Please rate the appropriateness of the following behavior based on what you believe the most frequent answer will be in this survey:

	Very Socially Inappropriate	Somewhat Socially Inappropriate	Somewhat Socially Appropriate	Very Socially Appropriate
Taking a part or all of the money from the charity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Leaving the equal division of the sum of money as it is	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Giving a part or all of the money to the charity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Post-Experimental Questionnaire (same for all elicitations)

What is your gender?

Female
 Male

What is the highest level of school you have completed or the highest degree you have received?

Less than high school degree
 High school graduate (high school diploma or equivalent including GED)
 Some college but no degree
 Associate degree in college (2-year)
 Bachelor's degree in college (4-year)
 Master's degree
 Doctoral degree
 Professional degree (JD, MD)

How old are you in years?

How do you see yourself:

Please indicate on the scale if are you generally a person who is fully prepared to take risks or do you try to avoid taking risks?

0 means: 'not at all willing to take risks'
10 means: 'very willing to take risks'

0 1 2 3 4 5 6 7 8 9 10

Risk Willingness

Below are a series of statements about you. Please read each statement and decide how much you agree or disagree with that statement.

Please answer every statement, even if you are not completely sure of your response.

If I knew that I could never get caught, I would be willing to steal a million dollars.

I would never accept a bribe, even if it were very large.

I'd be tempted to use counterfeit money, if I were sure I could get away with it.

Having a lot of money is not especially important to me.

I would get a lot of pleasure from owning expensive luxury goods.

I wouldn't use flattery to get a raise or promotion at work, even if I thought it would succeed.

If I want something from someone, I will laugh at that person's worst jokes.

I wouldn't pretend to like someone just to get that person to do favors for me.

What is your general opinion of charitable organizations?

'1' = I do not support the work of charitable organizations.
 '10' = I fully support the work of charitable organizations.

1 2 3 4 5 6 7 8 9 10

General Opinion of Charities

What is your general opinion of the United Nations International Children's Emergency Fund (UNICEF)?

'1' = I do not support the work of UNICEF.
 '10' = I fully support the work of UNICEF.

1 2 3 4 5 6 7 8 9 10

General Opinion of the UNICEF

What is your general opinion of the Doctors Without Borders?

'1' = I do not support the work of Doctors Without Borders.
 '10' = I fully support the work of Doctors Without Borders.

1 2 3 4 5 6 7 8 9 10

General Opinion of the Doctors Without Borders

What is your general opinion of the World Wildlife Fund (WWF)?

'1' = I do not support the work of the WWF.
 '10' = I fully support the work of the WWF.

1 2 3 4 5 6 7 8 9 10

General Opinion of the WWF

So that individual payments can be made, please enter your Amazon Mechanical Turk Worker ID. (This number can be found on your worker Dashboard)

Type carefully. If your Worker ID is entered incorrectly, additional earnings may not be received.

Worker ID:

After payments are made, your Worker ID will be permanently deleted from the experimenter's data.

Optional:
 If you have any comments please feel free to write them here, thank you!