

Social framing and cooperation: The roles and interaction of preferences and beliefs

Elizabeth Bernold*[†] Elisabeth Gsottbauer[‡] Kurt A. Ackermann[†]
Ryan O. Murphy[†]

October 23, 2014

Abstract

Evidence suggests that there are substantial and systematic differences in cooperation rates under varying framing conditions in social dilemmas. Several explanations of these differences have been presented. Some (e.g. McCusker and Carnevale (1995), van Dijk and Wilke (2000)) argue that social frames mainly cause subjects' preferences to change, while others (e.g. Dufwenberg et al (2011), Ellingsen et al (2012)) argue that frame-specific terminology, such as the name of the game, mainly affects subjects' beliefs about others' behavior, which in turn affects their own behavior. This paper advances the discussion concerning the role of frames in social dilemmas by simultaneously identifying the effects framing has on both *preferences* and *beliefs* with the same players and within the same experiment. The current experiment employs a design in which the same interaction was labeled differently, such that the interaction was referred to as a COMMUNITY GAME, a WALL STREET GAME, an ENVIRONMENT GAME, or simply as a GAME in the control condition. In all four experimental conditions, we measured the subjects' (i) social preferences, (ii) cooperative behavior and beliefs in a one-shot public goods game with the strategy method, (iii) cooperative behavior and beliefs in a ten-round iterated public goods game with random group rematching, and (iv) donation decisions to a naturally occurring public good. Overall, our results show that preferences, as well as beliefs, are both significant predictors of cooperation decisions, and that framing has significant effects on these two predictors' relative weights and also on aggregate cooperation rates. However, the impact of framing on cooperative behavior is complicated, and our results

*FehrAdvice & Partners AG, Bergstrasse 114, 8032 Zurich, Switzerland

[†]ETH Zurich, Chair of Decision Theory and Behavioral Game Theory, Clausiusstrasse 50, 8092 Zurich, Switzerland, elizabeth.bernold@dbgt.gess.ethz.ch, kurt.ackermann@dbgt.gess.ethz.ch, rmurphy@ethz.ch

[‡]ETH Zurich, Chair of Economics, Clausiusstrasse 37, 8092 Zurich, Switzerland, elisabeth.gsottbauer@econ.gess.ethz.ch

indicate that the magnitude and direction of framing effects may depend on diverse and subtle context-dependent mechanisms that are not yet fully understood.

Keywords: Cooperation, Framing, Public Goods Games, Social preferences, Social Value Orientation (SVO)

1 Introduction

Understanding when, and more fundamentally why, individuals cooperate has emerged as an active research area in both economics and psychology. In particular, attention has been paid to identifying the determinants of cooperation, and especially to understanding how the structure of a decision environment may systematically impact decision makers' propensities to contribute to a common good. Superficial features of the choice environment have been shown to have an effect on choice behavior, a phenomenon referred to as *framing effect*.¹ Paraphrasing Rabin (1998), two logically equivalent (but not transparently equivalent) statements of a decision problem demonstrate a framing effect when they lead decision makers to choose different options. A wave of studies have sought to measure the implications of framing effects within experiments by varying the context imposed on otherwise equivalent decisions. In a well known experiment, Liberman et al (2004) found that subjects' propensity to cooperate in a prisoner's dilemma was significantly enhanced or curbed by a small change in context. Specifically, by changing only the name of the game, Liberman et al. found that average rates of cooperation were shifted. This intriguing finding suggests that framing may provide a reliable and inexpensive way of promoting cooperation without resorting to extensive monitoring or costly punishment. However, the particular mechanisms of why framing effects occur remain somewhat unclear.

Public goods games with different labels, i.e. POSITIVE versus NEGATIVE and GIVE versus TAKE frames have been shown to lead to similar patterns as those reported by Liberman et al (2004). For example, Andreoni (1995) reports that in a public goods game based on the voluntary contribution mechanism, cooperation rates are considerably higher under a POSITIVE frame than under a NEGATIVE one. Yet overall, experiments have provided mixed evidence on the direction of a framing effect on cooperative outcomes. For example, Park (2000) replicates Andreoni's framing experiment and confirms that contribution rates are higher under a positive frame. Sonnemans et al (1998) find that subjects are more cooperative in a public good than public bad frame using a step-level public goods game. Ellingsen et al (2012) and Dreber et al (2012) use a one-shot prisoner's dilemma game, reporting that subjects cooperate more when the decision is labelled COMMUNITY GAME than when it is called WALL STREET or PROFIT

¹The term *framing* is often, perhaps typically, applied to risky decision contexts contrasting gains and losses (e.g., Tversky and Kahneman (1981)). In this paper, we use the term *framing* more broadly to refer to the names or descriptions that are applied to otherwise identical choice situations. This broader usage of the term is consistent with Rabin (1998).

GAME. Others report contrary effects. For example, Brewer and Kramer (1986) find greater contributions under a TAKE or NEGATIVE than GIVE or POSITIVE framing, while Cubitt et al (2011) find no evidence for a framing effect when using TAKE and GIVE strategy labels. Furthermore, Brandts and Schwielen (2009), using a series of one-shot public goods games, report that a COMMUNITY frame leads to lower contributions than a dilemma framed as STOCK EXCHANGE GAME. Similarly, Dufwenberg et al (2011) find that contributions are higher in a NEUTRAL than a COMMUNITY frame.²

These diverse results may be better understood and organized by identifying the relevant drivers of cooperation. Two dominant conjectures have emerged in the literature. First, McCusker and Carnevale (1995), van Dijk and Wilke (2000) and Weber et al (2004) suggest that preferences for cooperation themselves may change based on context. In this vein, Levitt and List (2007) note that “only weak evidence” of cross-situational consistency in social preferences has been reported in the literature, meaning that preferences seem to change according to frames. List (2007) cites results from a series of dictator games that suggest context dependency in preferences, implying that a change in preferences due to context itself could be a central driver of cooperation decisions, and hence explain observed changes in the rates of cooperation.

A second conjecture is that frames affect beliefs about others’ behavior and thereby change expectations of cooperation. This speculation tacitly assumes that individuals’ social preferences are (relatively) stable and therefore not the driver of the noted different cooperation rates. Dufwenberg et al (2011) report results from a one-shot public goods game and find that manipulating the game’s title and labeling the subjects’ strategies in different ways affects the subjects’ beliefs about others’ anticipated contributions. The authors conclude that frames serve as social cues to comparable real-life situations from which others’ prosocial behavior can be better anticipated. Ellingsen et al (2012) report a prisoner’s dilemma game called either the COMMUNITY or STOCK MARKET GAME. To investigate whether social framing enters beliefs or preferences, they observe subjects’ decisions in a sequential and also a simultaneous prisoner’s dilemma. While beliefs could only be manipulated by frames in the simultaneous game, and social preferences should affect behavior whether decisions are made simultaneously or sequentially, Ellingsen et al (2012) point to the lack of framing effect observed in the sequential game as evidence that frames influence beliefs but not preferences. In the same vein, Cubitt et al (2011) note equivalent emotional responses to a one-shot public goods game under different frames. Consistent with these results, Dreber et al (2012) found that preferences are robust to superficial changes in the decision context as evidenced by equivalent dictator game allocation decisions observed under two different framing conditions.

This paper presents an experimental test that is designed to be comprehensive and discern between these two potential explanations. The design allows for the measure and analysis of social preferences, beliefs, cooperation decisions (both one-shot and repeated),

²Dufwenberg et al (2011) point to a particularly negative connotation of the word “community” that was held by the particular German subject pool members as a potential explanation for the unexpected direction of the effect.

and a donation to a real world cause, all measured at the individual level. We assess to what extent individuals' endogenous preferences and beliefs, considered concurrently, can predict cooperative decisions in a public goods game and also real world charitable donations. We hypothesize that cooperation rates will generally reflect those observed by Liberman et al (2004), namely that more cooperation would be observed in the COMMUNITY frame, and less in the WALL STREET frame, as compared to the NEUTRAL frame. Additionally we consider a new manipulation called the ENVIRONMENT frame and investigate it as yet another context. One natural application of the "name of the game" effect is in the domain of environmental decision making, where a low cost framing manipulation could be useful in nudging people toward collectively beneficial choices and pro-environmental behavior. We expect that the ENVIRONMENT frame will have a positive effect on cooperation levels generally similar to the effect we expect from the COMMUNITY frame. The experimental design allows for the evaluation of the relative predictive capacities of preferences and beliefs in different frames, and thus can inform us about the extent to which framing affects preferences, beliefs, or both at the same time. The disentanglement of these decision making processes has not been possible in previous studies and is one of this paper's major contributions.

2 Procedure and Experimental Design

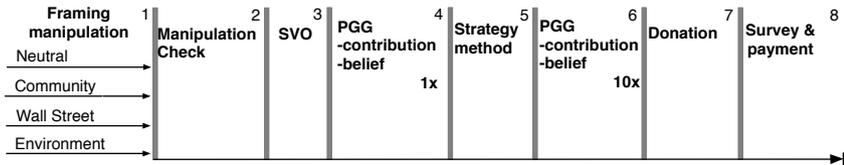
The experiment applied a between-subjects design with four framing treatments: NEUTRAL, COMMUNITY, WALL STREET, and ENVIRONMENT. The frame was implemented as the title of the experiment and was presented when explaining the decision tasks to the participants at the beginning of an experimental session. The name of the game was displayed on the front page of the instruction sheets, and was also displayed in the heading of every sub-section in the instructions. Apart from this framing information, the experimental procedure was identical across the four conditions.

To verify that participants had attended to the experimental frame, each subject was required to input the "name of the session" into the computer before the decision making phase of the experiment started. This served as a manipulation check and a way to verify that subjects had at least a minimal level of awareness of the particular label that was used for the experimental session and thus may induce framing. For example, in the COMMUNITY treatment, each subject had to type the words "community game" into a text box before being permitted to proceed with the experiment.

The experiment consisted of four main tasks which were conducted in the following order: a social preference measure, a one-shot public goods game using a strategy table, a 10 round repeated public goods game, and finally a donation decision to a real world public good (see Figure 1 for the overall experimental design). All choice tasks were incentivized, and it was clearly stated in the instructions that the different parts of the experiment were independent of each other. Written instructions were provided on paper, and were also conveyed over a loudspeaker from a recording to establish

common knowledge. Feedback on earnings and the outcomes of each choice task were only provided after completion of the entire experiment, except for the experimental block with the repeated public goods game, in which results from each round were available to players.

Figure 1: Experimental Sequence



Each of the four frames were implemented in two sessions to mitigate sampling errors. Thus there were eight experimental sessions in total, all of which were conducted at the ETH Decision Science Laboratory (DeSciL) in December 2013. The experiment was computerized and controlled via z-Tree (Fischbacher, 2007). The research was conducted in English (which is a common practice at the DeSciL), and participants were screened for their language comprehension ability prior to their participation.

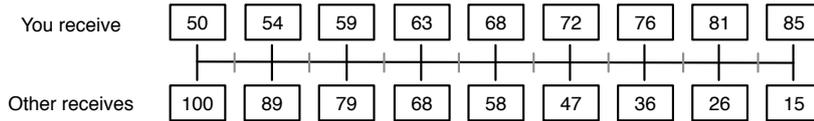
The 188 participants who voluntarily took part in the experiment were recruited from the Swiss Federal Institute of Technology (ETH) and the University of Zurich, and each subject participated in only one session. Each session included either 20, 24, or 28 participants based on show-up rates, and each session lasted for about 2 hours. None of the subjects had participated in a public goods experiment previously. During the course of the experiment, monetary units were referred to as points. The exchange rate of 8 points to 1 Swiss franc (CHF) was used throughout the experiment and this was common knowledge. Participants earned on average CHF 33.63 (approximately USD 37) in addition to their CHF 10 (USD 11) show up payment.

2.1 Social Preferences

Subjects' social preferences were assessed with the SVO Slider Measure (Murphy et al, 2011) as implemented in terms of a z-Tree module (Crosetto et al, 2012). The measure consists of 15 joint payoff allocation questions. For each item, a subject selects the most preferred allocation of monetary resources between herself and another person out of 9 options in total (see Figure 2 for an example). The measure yields a single index of an individual's social preference on a continuous scale spanning competitiveness (i.e. maximizing advantageous inequality), individualism (i.e. narrow self-interest), prosociality (i.e. maximizing efficiency), and altruism (i.e. maximizing the other person's payoff). Subjects were informed that one of their 15 joint allocation decisions would be selected at random to become relevant for their final payment, such that their choice in one randomly selected item would affect their own payoff, as well as the payoff of some other

randomly selected subject, but they were not informed which decision would be paid until the end of the experiment, of course. Self/other pairings were neither disclosed nor reciprocal (i.e., Subject A's *other* amount was allocated to Subject B, but Subject B's *other* amount was not allocated to Subject A), eliminating the opportunity for strategic decision making by subjects in this task. All of this was common knowledge.

Figure 2: Example SVO Slider Item



2.2 One-Shot Public Goods Game

After completing the social preference measurement, subjects were informed about the details of next task, which was a standard linear one-shot public goods game with belief elicitation and a strategy table (Selten, 1967; Fischbacher et al, 2001). Each subject was randomly matched with three other subjects to form a group of four and was provided with an endowment of 20 points of which subjects could then transfer any proportion to the group account. This transfer decision we refer to as the *unconditional contribution*. The total contribution to the group account was multiplied by 1.6 and the product distributed evenly among the four group members. Concurrent to their transfer decision, subjects were required to report the average amount they believed the other three members of their group would transfer. The elicitation of their beliefs was incentivized using a proper scoring rule.³ After making the unconditional contribution and belief inputs, subjects completed a strategy table in which they indicated their conditional contributions, namely, how much they would contribute given each possible average integer contribution of group members over the interval 0-20 inclusive. For payment, one of the four group members was randomly selected as the conditional contributor whose relevant contribution was the respective entry in their strategy table, while the unconditional contributions became relevant for the other three group members. Subjects also completed a quiz to demonstrate their understanding of the task before it began.

2.3 Repeated Public Goods Game

Next, subjects played 10 rounds of a repeated public goods game with identical characteristics as the previous game (the group size was 4, multiplier was 1.6, each subject

³A proper scoring rule yields incentives to accurately report beliefs (Gächter and Renner, 2010). In this experiment, a correctly reported belief yielded the subject 40 extra points; a belief that deviated by 1 point earned 32 additional points; a deviation by 2 points earned 24 additional points, and zero points otherwise.

received an endowment of 20 points each round). Subjects were randomly rematched into new groups of four players each round (i.e. stranger matching). In addition to selecting the number of points they would transfer to the group account in a particular round, participants also indicated the average number of points they believed their other group members would contribute in that round, again as a concurrent and incentivized beliefs elicitation task.

2.4 Public Good Donation and Survey

Following the tenth and final round of the repeated public goods game, participants viewed their earnings from all decisions and could choose to transfer a portion of their earnings to a real world carbon offset project, i.e., the sustainably-designed Monte Rose Hut in the Swiss Alps.⁴ After making their donation decision, subjects completed a questionnaire with standard socio-economic demographic questions, field of study questions, environmental preferences measures, cognitive reflection test (Frederick, 2005) and a free-answer question about their impressions of the experiment. Then, they were paid their earnings and excused from the laboratory privately one at a time.

3 Results

Out of the 188 subjects in total, ten subjects were excluded from all subsequent analyses and the presented results since they showed intransitive choice patterns in the SVO Slider Measure, which is an indicator of random responding. Hence the analyses are based on the remaining sample of $n = 178$ subjects. Note that no substantive differences in results emerge when the intransitive subjects are included or excluded from the subsequent analyses.

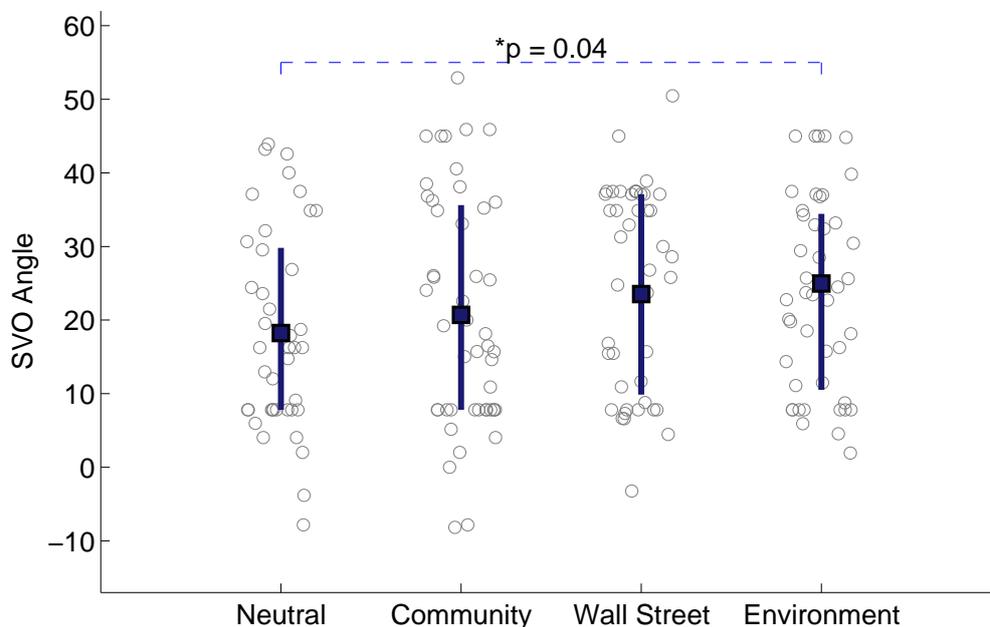
3.1 Social preferences (SVO)

From a subject's pattern of choices in the Social Value Orientation (SVO) Slider Measure, a continuous score reflecting the weight the subject attaches to the outcome for the other in relation to his or her own outcome can be computed. The score is often represented in terms of an angular degree, such that an angle of 0° corresponds to narrow self-interest and an angle of 45° indicates perfect prosociality, meaning that a decision maker attaches the same weight to the outcome for another person as the decision maker attaches to

⁴The participants saw a screen with a picture of the building complex and the following text, "You now have the option to transfer some portion of your earnings to an environmental protection project in Switzerland. Many of your daily activities create CO₂ emissions, contributing to atmospheric degradation and climate change. Your contribution to the Monte Rosa SAC Hut (pictured to the right) allows you to offset some of your "carbon footprint," neutralizing your environmental impact. The ETH guarantees that your contribution will support carbon offsets at the Monte Rosa SAC Hut."

his or her own outcome, which is isomorphic to a coefficient of 1 in a simple joint utility model (see Murphy et al (2011); Murphy and Ackermann (2014b) for details). A negative angle indicates spitefulness, meaning that a person would be willing to pay to decrease the other person's payoff. The observations and interquartile ranges of the SVO angles' distributions are displayed in Figure 3. Apart from a statistically significant but small difference in the distribution of social preferences between the ENVIRONMENT and NEUTRAL frames (Kolmogorov-Smirnov test, $p = 0.04$), no significant differences between treatments were observed. This result indicates that social preferences were largely unaffected by the experimental framing.

Figure 3: Means and interquartile ranges of subjects' social preferences organized by treatment over the jittered individual scores.



3.2 One-Shot Public Goods Game

3.2.1 Beliefs

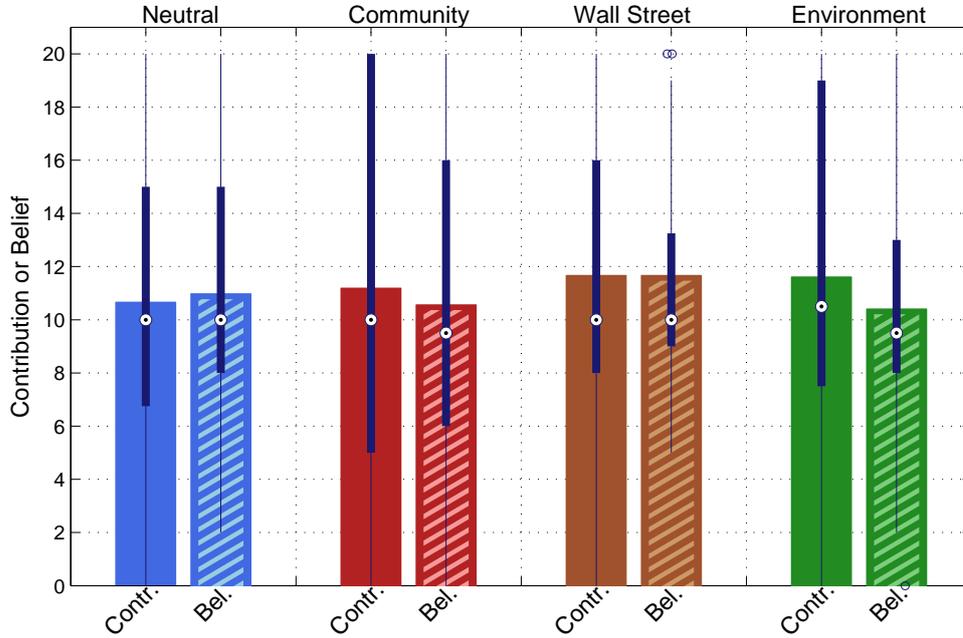
In the one-shot public goods game, subjects expected others to contribute an average of 10.56 points in the COMMUNITY frame, and 11.67 points in the WALL STREET frame (see Figure 4). The mean belief in the NEUTRAL frame was 10.98, and in the ENVIRONMENT frame it was 10.41. Overall, mean beliefs in all four treatments are roughly similar but

surprisingly subjects in the WALL STREET frame expected higher contributions from their group members as compared to subjects in the COMMUNITY frame on the aggregate (Kolmogorov-Smirnov test, $p = 0.04$). While the beliefs reported in the COMMUNITY, ENVIRONMENT and NEUTRAL frames ranged from a minimum of 0, 2, and 2 points respectively, no subject in the WALL STREET frame reported to believe that the other group members would contribute less than 5 points.

3.2.2 Unconditional Contributions

Figure 4 also displays the aggregate contribution levels in the one-shot public goods game. The average level of unconditional contributions was around 55 percent of the endowment in all frames (10.70 in NEUTRAL, 11.19 in COMMUNITY, 11.67 in WALL STREET, 11.61 in ENVIRONMENT frame), such that unconditional contributions do not differ significantly between treatments (two-sample Mann-Whitney-U test). The average level of unconditional contributions, around 55 percent of the endowment, is slightly higher than generally reported in the literature (around 40 percent, see Kocher et al (2008)). Overall, the variance in unconditional contributions within treatment is higher in the COMMUNITY frame as compared to the other three treatments, while lowest in the WALL STREET frame.

Figure 4: Means and interquartile ranges of unconditional contributions and beliefs in the one-shot game



3.2.3 Conditional Cooperation

Figure 5 shows the subjects' average conditional contributions from the strategy method. On average, there is a positive average contribution to the public good when others give zero across all framings, indicating strong cooperative tendencies. Subjects in the NEUTRAL treatment are closest to the perfectly conditionally cooperative (45°) line. Interestingly, WALL STREET participants conditionally contribute in excess of the contribution matching for others' contribution for levels 0 - 12 (while those in the other three treatments conditionally contribute less in these fields) and so appear relatively generous compared to the conditional contribution behavior in the other treatments.

Figure 5: Average conditional contributions

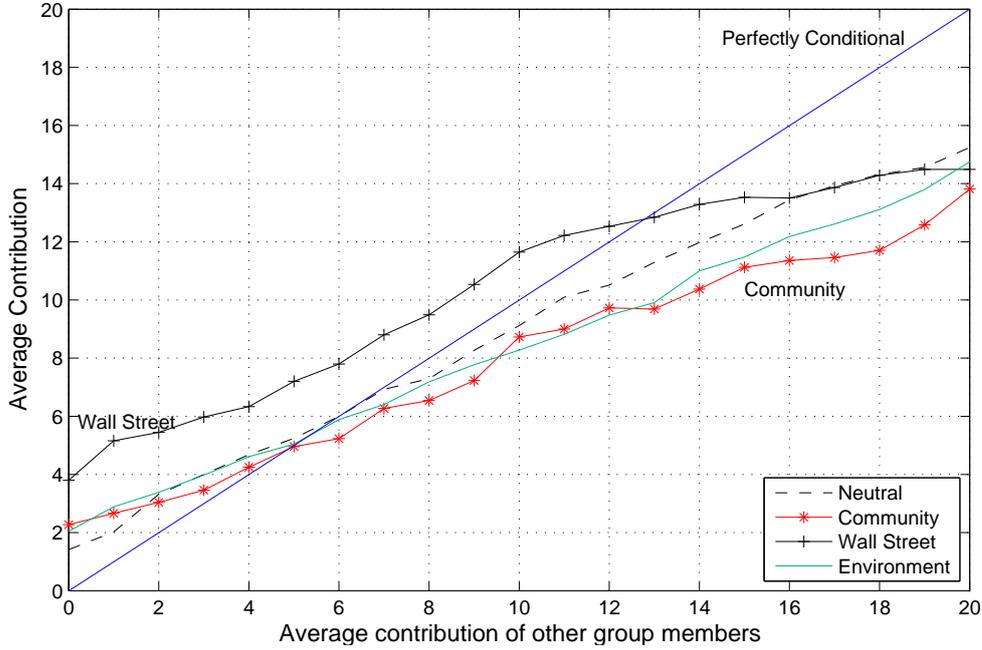


Figure 6 communicates the extent to which subjects' conditional contributions as indicated in the strategy table match their actual contributions based on their reported belief (so called *belief-based conditional contribution* or *BBC*). We find that behavior across the NEUTRAL, COMMUNITY and ENVIRONMENT frames is consistent in the sense that subjects contribute roughly the same as they indicated they would contribute in the strategy table given their beliefs about the average contribution of others. However, subjects in the WALL STREET treatment show considerable inconsistencies in this respect. That is, the correlation between subjects' one-shot contributions and their BBCs is only 0.50 in the WALL STREET frame, while it is greater than 0.75 in the other three treatments (see Figure 6 for a visualization).

We follow convention e.g., Fischbacher et al (2001), by categorizing subjects into four types on the basis of their choices in the strategy table: *Conditional cooperators*, *Hump-shaped contributors*, *Free riders*, and *Others* (i.e., subjects who show a different pattern of choices in the strategy table). Table 1 reports the distribution of cooperator types (and their average unconditional contribution and belief) for each treatment. Conditional cooperator is the modal category, comprising between 62 percent (WALL STREET) and 85 percent (ENVIRONMENT) of the sample. There were marginally significantly more conditional cooperators in the ENVIRONMENT than WALL STREET frame ($p = 0.05$); the differences between the other categories and frames were not significant ($p > 0.05$).

Figure 6: Conditional vs. unconditional contributions

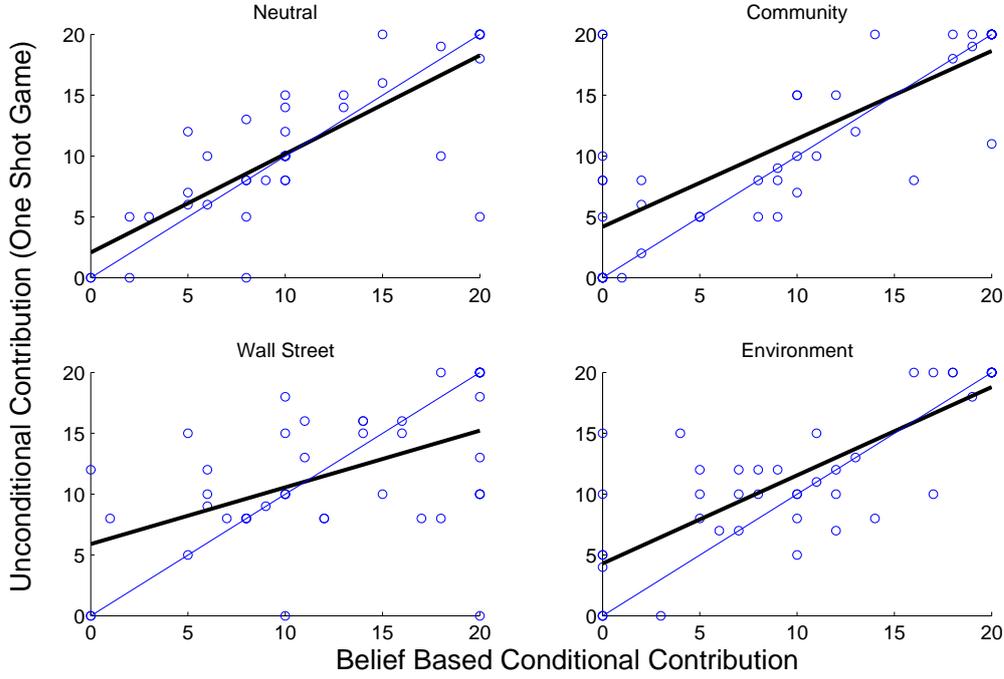


Table 1: Conditional contributor types and corresponding average unconditional contribution and belief (in points)

	Cond. Coop.			Hump Shaped			Free Rider			Others		
	%	Cont.	Belief	%	Cont.	Belief	%	Cont.	Belief	%	Cont.	Belief
Community	70.8	11.8	11.1	10.4	8.0	8.2	10.4	4.6	6.8	8.3	14.0	13.5
Environment	77.1	12.3	11.2	2.1	7.0	9.0	8.3	5.0	4.2	4.2	14.0	8.5
Neutral	72.7	11.8	11.5	9.1	6.0	7.8	4.5	0.0	3.0	6.8	12.0	14.7
Wall Street	60.4	12.8	11.9	14.6	10.9	11.0	6.2	4.0	10.0	12.5	11.0	12.2

3.2.4 Relationship between Preferences, Contributions and Beliefs

Unconditional contributions and beliefs are significantly correlated across all treatments ($p < 0.01$ for all frames), yet subjects' contributions are much better explained by their beliefs in the NEUTRAL, COMMUNITY and ENVIRONMENT frames (all correlation coefficients are at least 0.84) than in the WALL STREET frame. The Spearman rank correlation coefficients between contributions and beliefs are highest in the COMMUNITY frame (0.92 Spearman rank correlation) and lowest in the WALL STREET frame (0.63).

The relationship between social preferences and contributions also varies across frames.

The correlation between social preferences and contributions is large (0.66, $p < 0.01$) in the COMMUNITY and NEUTRAL frame (0.51, $p < 0.01$), and rather low in the ENVIRONMENT (0.26, $p = 0.088$) and WALL STREET frame (0.21, $p = 0.17$).

The regression analysis in Table 2 corroborates the findings from the correlation analysis, showing that beliefs and social preferences are significant predictors of contributions in the NEUTRAL, COMMUNITY and ENVIRONMENT frames, while predictive power is substantially weaker in the WALL STREET frame, where social preferences lack significant predictive capacity altogether. Social preferences and beliefs explain at least 75 percent of the variance in contribution levels across all frames, except for the WALL STREET frame, in which only 40 percent of the variance in contributions is explained. This relatively lower explanatory power of preferences and beliefs may either indicate that the subjects' behavior is noisier in the WALL STREET frame compared to the other frames, or it may indicate that at least one important variable –such as higher-order beliefs, perhaps– is missing.

Table 2: Explaining contributions in the one-shot game

	Neutral	Community	Wall Street	Environment
Belief	0.76***	0.82***	0.62***	0.86***
SVO	0.24***	0.17**	0.02	0.13*
R-Square	0.76	0.87	0.40	0.79
n	41	48	45	44

Note: OLS-regression with standardized beta coefficients.
*significant at 10%; **significant at 5%; ***significant at 1%.

3.3 Repeated Public Goods Game

Subjects' behavior in the iterated public goods game with stranger matching is examined next. In each iteration of the public goods game, subjects indicated how many points (from their endowment of 20 points) they want to contribute to the public account, and also reported how many points they believed their fellow group members would contribute, on average.

3.3.1 Beliefs

In the repeated public goods game, participants expected the other three members of their group to contribute an average of 10.5 points in the NEUTRAL frame, 11.9 in the COMMUNITY frame, 10.5 in the ENVIRONMENT frame and 9.0 in the WALL STREET frame. Although beliefs were statistically undifferentiated across the four treatments in the first round of the repeated game, beliefs diverged across treatments beginning

in Round 2, and the level of divergence grew for the subsequent rounds. Subjects' reported beliefs of others' contributions in the WALL STREET frame declined steadily, and remained persistently lower than in the other three treatments. In Round 10, subjects reported that they expected their other 3 group members to transfer a relatively high average of 10.4 points in the COMMUNITY frame as compared to only 6.5 points in the WALL STREET frame. Overall, final round beliefs in the COMMUNITY and ENVIRONMENT frames were significantly higher (Mann-Whitney U and Kolmogorov-Smirnov test $p < 0.05$) than in the NEUTRAL frame or the WALL STREET frame.

3.3.2 Contributions

Across all rounds, mean contributions were highest in the COMMUNITY (11.19) and lowest in the WALL STREET (7.68) frame (Figure 7, Mann-Whitney-U test for equivalence between these two treatments $p < 0.001$).⁵ Contributions in the first round were roughly equivalent in all treatments (Mann-Whitney U test $p = 0.26$, see Table 3). Interestingly, consistency between behavior in the one-shot situation and in the first round of the repeated interaction varies by treatment. That is, contributions in the first round of the repeated game were highly correlated with contributions in the one-shot decision in the NEUTRAL ($r = 0.843$), COMMUNITY ($r = 0.763$) and ENVIRONMENT ($r = 0.728$) frames, but only mildly correlated in the WALL STREET frame ($r = 0.264$). Overall, contribution levels diverged between treatments starting in Round 2. The decline in contributions observed from first to last round is steepest in the WALL STREET frame with a decline of 55 percent from 10.6 in Round 1 to 4.8 in Round 10. In the COMMUNITY frame we observe a decline of only 30 percent. A summary of the results is presented in Figure 7.

⁵Framing also influences the frequency of complete free-riding (subjects contributing 0 points to the public good), and full contributors (subjects who contribute all of their endowed points to the public good). In the COMMUNITY frame, the proportion of subjects who are full contributors is higher than in all other frames (on average 18.8), although it also declines over time. The share of free-riders is highest in the WALL STREET frame, and this share increases from the first to the final round starting at 12.5 and ending at 29.2. The latter observation is in line with Andreoni (1995), who observes twice as many free-riders in the negative frame condition compared to the positive frame.

Figure 7: Mean contributions in the repeated public goods game by treatment

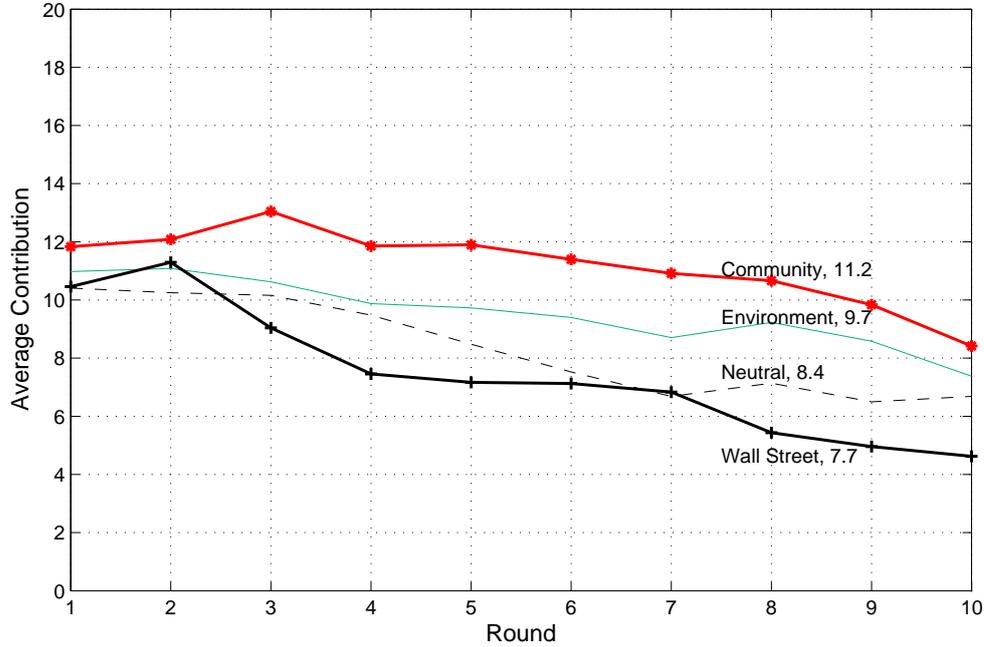


Table 3: Tests for differences in Round 1 contributions (p -values)

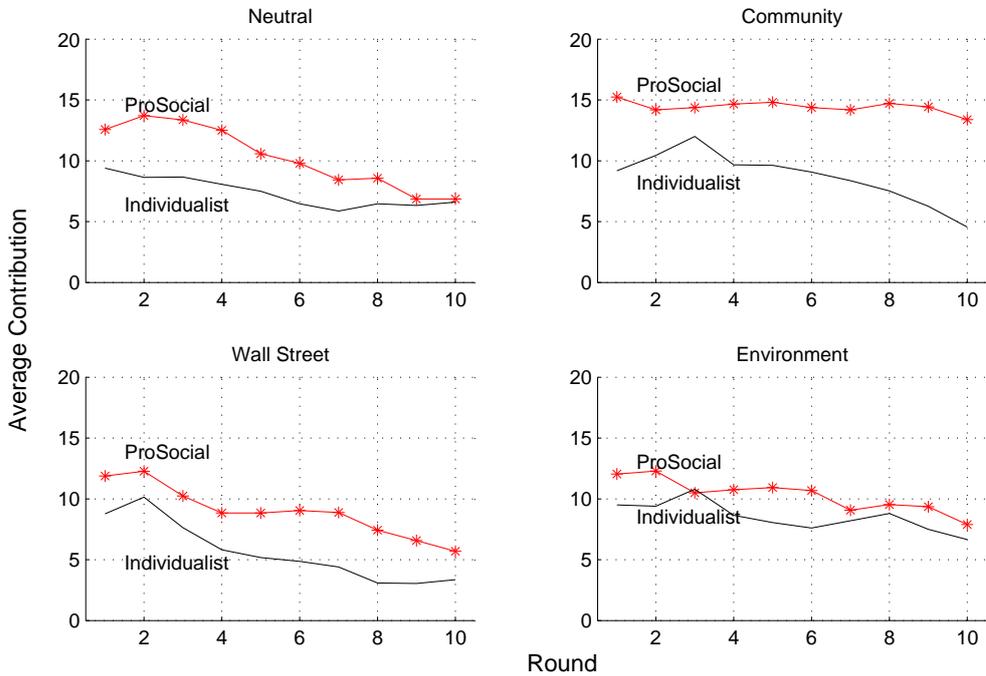
	KS Test	Ranksum Test
Community/Environment	0.33	0.53
Community/Neutral	0.07	0.26
Community/Wall Street	0.48	0.30
Environment/Neutral	1.00	0.64
Environment/Wall Street	0.95	0.67
Neutral/Wall Street	0.98	0.97

3.3.3 Relationship between Preferences, Contributions and Beliefs

We now turn our attention to the link between contributions and social preferences over the course of the repeated game. For the purpose of providing a visualization of how different social preference types behave across the repeated interaction, we categorize subjects in two idealized SVO types according to their choices in the SVO Slider Measure, namely prosocials (SVO angles between 22.45° and 57.15°) and individualists (SVO angles between -12.04° and 22.45°). On the basis of this categorization, 51.1% of

all subjects are individualistic, and 48.9% are prosocial. Figure 8 illustrates the contribution levels of the two SVO types over the 10 rounds per treatment. The plot shows that framing predominantly affected subjects with a prosocial SVO, whose cooperation rates are remarkably high and stable in the repeated COMMUNITY frame relative to the other three frames, while subjects with an individualistic SVO show the commonly observed decline of cooperation over the course of the iterations. That is, prosocial subjects in the COMMUNITY frame maintain the highest average contribution level of any social preference group in any of the frame conditions. Concretely, prosocial subjects in the COMMUNITY frame contribute an average of 72 percent of their endowment, most conspicuously more than WALL STREET frame prosocials, whose average contribution declines from 60 percent of their endowment in Round 1 to 29 percent in Round 10. In other words, the data suggest that the COMMUNITY frame had an effect on prosocial subjects, but not on individualistic subjects. Technically speaking, this means that the COMMUNITY frame effected an interaction between social preferences (SVO) and the number of round on contributions in the repeated public goods game. This finding is corroborated by the results of regression analyses indicating a significant interaction between SVO (here used as a continuous variable) and the round number on contributions in the repeated game in the COMMUNITY frame, but not in the other frames (see the results concerning Model 2 in Table 5).

Figure 8: Contributions by social preferences type by treatment



Next, we analyze to what extent framing affected the relation between both beliefs and social preferences as independent variables on the one side, and contributions in the repeated game as the dependent variable on the other side. Table 4 reports on the results of an OLS regression with contributions in the first round of the repeated game as dependent variable, while Table 5 shows the results of a random-effects GLS regression with contribution levels across the whole repeated interaction as the dependent variable. While social preferences do not seem to add predictive power beyond the predictive capacity of beliefs in the first round of the iterated game (see Table 4), they do carry significant predictive power across the repeated interaction as a whole (see Model 1 in Table 5).⁶ Figure 9 displays the results from a series of OLS regressions that employ preferences and beliefs as explanatory variables for contributions in each round separately to provide a clearer picture of the dynamics of the two independent variables' relative predictive capacities across the repeated play. We find that beliefs are generally the strongest predictor across all frames. However, in the COMMUNITY frame, the predictive power of social preferences increases with the number of rounds, such that social preferences are even more predictive than beliefs in the final two rounds. This is just another way in which the previously discussed interaction effect stemming from the sustained high cooperation levels among prosocials in the COMMUNITY frame expresses itself. In the tenth and final round, social preferences have a beta weight (i.e., standardized regression coefficient) of 0.5 in the COMMUNITY frame, which is higher than any of the beta coefficients assigned to social preferences in all of the other three frames irrespective of which round is considered. This makes sense according to the following explanation. If prosocial subjects in the COMMUNITY frame contribute high amounts throughout the repeated game (including the final round), while individualistic subjects' contributions decline to very low amounts, then the subjects' contributions in the later rounds correspond more closely to their social preferences than in the earlier rounds, resulting in a higher association between the two variables at later stages in the game. This is not an additional finding, but rather a manifestation of the previously indicated interaction effect.

⁶One thing worth keeping in mind here is that social preferences (SVO) are used as a static predictor –i.e., SVO was measured only once at the beginning of the experiment–, while beliefs are updated and assessed in every round, and thus used as a dynamic predictor. Hence, the predictive power of social preferences is likely underestimated because there is evidence that social preferences may be updated over the course of a repeated interaction as well (Murphy and Ackermann, 2014a), presumably because the behavior of others may be interpreted as a revelation of their types, and these revelations may then effect reciprocal reactions (Ackermann et al, 2014).

Table 4: Explaining contributions in the first round of the repeated game

	Neutral	Community	Wall Street	Environment
Belief	0.84***	0.87***	0.82***	0.79***
SVO	0.12	0.08	0.04	0.06
R^2	0.80	0.85	0.70	0.65
n	41	48	45	44

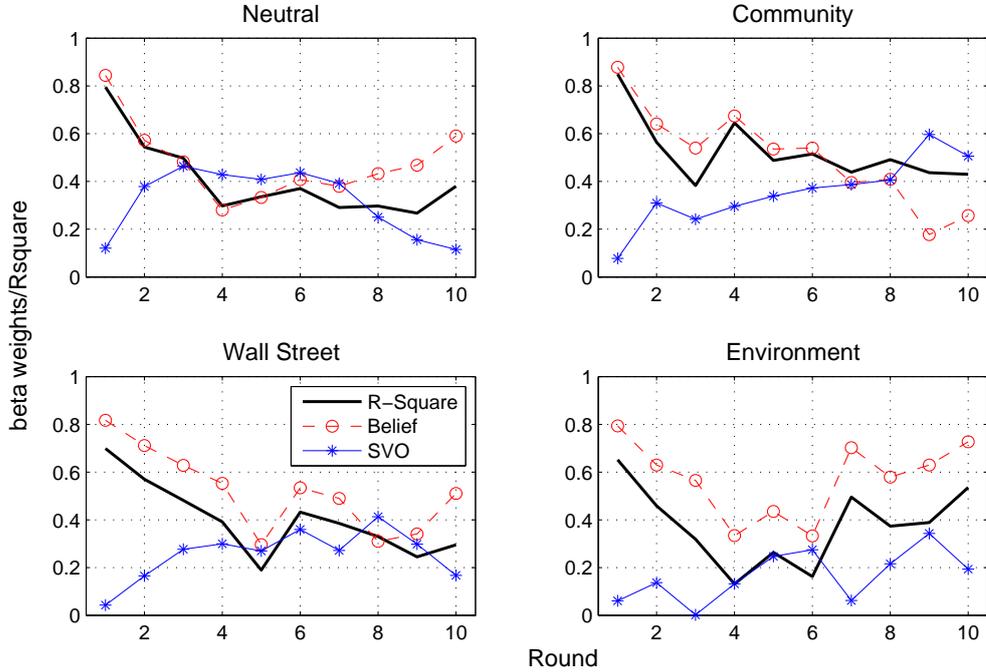
Note: OLS-regression with standardized beta coefficients. *significant at 10%; **significant at 5%; ***significant at 1%.

Table 5: Random-effects GLM regression on contribution levels in the repeated public goods game

	Model 1				Model 2			
	Neutral	Community	Wall Street	Environment	Neutral	Community	Wall Street	Environment
SVO	0.13*** (0.04)	0.17*** (0.04)	0.14*** (0.05)	0.07* (0.04)	0.17*** (0.05)	0.10** (0.04)	0.12* (0.07)	0.04 (0.04)
Belief	0.76*** (0.13)	0.71*** (0.10)	0.74*** (0.08)	0.67*** (0.09)	0.75*** (0.13)	0.72*** (0.09)	0.74*** (0.08)	0.68*** (0.10)
Round	-0.22*** (0.08)	-0.27*** (0.09)	-0.23*** (0.09)	-0.20*** (0.07)	-0.08 (0.11)	-0.52*** (0.14)	-0.30 (0.22)	-0.35*** (0.12)
SVO x Round					-0.01 (0.01)	0.01*** (0.004)	0.003 (0.008)	0.006 (0.005)
Constant	-0.03 (1.53)	0.75 (1.13)	-0.98 (1.61)	2.41 (1.53)	-0.70 (1.45)	2.15* (1.15)	-0.62 (2.00)	3.10 (1.47)
R^2	.43	.51	.43	.37	.43	.52	.43	.37
n	41	48	45	44	41	48	45	44

Notes: Robust standard errors clustered on subjects in parentheses. The symbols indicate: * significant at 10%; ** significant at 5%; *** significant at 1%.

Figure 9: Contributions explained by preferences and beliefs per round



3.4 Donation to Public Good

After subjects were informed about their earnings from the experiment, they were offered the option to donate a portion of what they earned to a real world carbon offset project. The average percent of earnings donated was significantly higher in the ENVIRONMENT frame with 9.19% as compared to the COMMUNITY frame with 3.67% (Kolmogorov-Smirnov and Mann-Whitney tests $p < 0.01$). Surprisingly, donations in the NEUTRAL frame (6.45%) were also significantly higher than donations in the COMMUNITY frame (Kolmogorov-Smirnov and Mann-Whitney tests $p = 0.08$ and $p = 0.03$), while donation amounts in the WALL STREET frame (7.91%) were in between the range of—and not statistically different from—the donation amounts in the other three frames. With respect to accounting for donation decisions, the data indicate that SVO is the most reliable predictor of donations to the naturally-occurring public good among the variables we have assessed. The bivariate correlations between average contributions in the repeated public goods game and donations to the naturally-occurring public good are $r = .2$ (COMMUNITY), $r = .03$ (ENVIRONMENT), $r = .27^*$ (NEUTRAL), and $r = .36^{**}$ (WALL STREET), while social preferences (SVO) correlate with donations significantly across all frames ($.30^{**} \leq r \leq .40^{***}$). The results provided in Table 6 show that average contributions in the repeated public goods game are not significant predictors of

donations when social preferences are statistically controlled. That is, social preferences (SVO) mediate the –when existent– generally weak relationship between contributions in a public goods game and donations to a naturally-occurring public good, except for the WALL STREET condition in which the two predictors seem to be redundant.

Table 6: Results of OLS regression with donations as dependent variable

	Neutral	Community	Wall Street	Environment
Social Preferences	0.31*	0.45***	0.23	0.31**
Average Contribution over 10 Rounds	0.11	-0.08	0.25	-0.03
R^2	0.14	0.17	0.17	0.09
n	41	48	45	44

Note: OLS-regression with standardized beta coefficients. *significant at 10%; **significant at 5%; ***significant at 1%.

4 Discussion and Conclusions

The main purpose of this study was to clarify whether previously found frame-dependent differences in cooperation levels can be attributed to an effect of framing on preferences, an effect of framing on beliefs, or both. In other words, we wanted to test the extent to which social preferences, beliefs, or both mediate the association between the exogenous label on a social dilemma, and the corresponding cooperation levels that are subsequently observed. However, we could only partially replicate the main effect of framing on contribution levels in a social dilemma. In contrast to previous studies (e.g. Liberman et al, 2004) we find no evidence for an effect of framing on contribution levels in a one-shot social dilemma, but we do find a main effect of framing on contribution levels in a repeated interaction. Furthermore, while we find no evidence for a general effect of framing on preferences or beliefs in both a one-shot and a repeated interaction, our results indicate that framing may affect the extent to which contribution levels are linearly dependent on preferences and beliefs, and moreover that framing may have an impact on the relative importance of preferences and beliefs in informing contribution decisions. Hence, the pattern of results we obtain does not allow for a simple and straightforward answer to the question how exactly framing affects behavior in a social dilemma. Nevertheless, several important conclusions follow from our results.

First, while we do not find a framing effect on social preferences, beliefs, or contribution levels in the one-shot situation, we observe a remarkably lower association between the two independent variables –social preferences and beliefs– and contribution levels in the WALL STREET condition compared to the other three conditions. This means that in the WALL STREET condition a linear model taking into account social preferences and beliefs does not yield a good fit to the data. One explanation for this could be that due to the WALL STREET framing, a further predictor may become important, such as higher-order beliefs. Previous evidence suggests that higher-order beliefs may indeed be

affected by framing (see Dufwenberg et al, 2011), but the question to what extent this is the case under a WALL STREET frame is not clearly answered yet and our data do not allow us to test this potential explanation explicitly. However, the fact that social preferences are not predictive of contribution levels in the WALL STREET frame at all, but do show at least some predictive power in the other three frames, is in support of the conclusion that missing belief components may be responsible for the lower model fit. A further potential indication of this interpretation stems from observations regarding the first round of the repeated interaction, where higher-order beliefs are very likely to play a significant role. More precisely, in the first round of the iterated public goods game, social preferences are not a significant predictor of contribution levels across all frames, while the predictive power of beliefs is very high and of comparable magnitude across all frames, including the WALL STREET frame.

Second, while we do not find a framing effect on social preferences, beliefs, or contribution levels in the first round of the repeated public goods game, we observe a framing effect that unfolds and reveals itself over the course of the iterated game. That is, we observe a framing effect that manifests itself in the later rounds of the repeated interaction resulting in higher contributions in the COMMUNITY as compared to the WALL STREET frame across rounds. The natural question that emerges is of course: How can this occur given the observation that framing did not appear to affect preferences, beliefs, or contributions in the first round? The most plausible answer is based on the observation of an interaction effect of framing and social preferences on contribution levels, such that prosocial subjects in the COMMUNITY frame manage to sustain a high level of cooperation, while prosocials in the other experimental conditions show the same decline of cooperation as individualists do across all treatments.

In conclusion, our results may help understand why findings from different studies on framing in experimental games are mixed and sometimes contradictory. The current data indicate that framing may not necessarily and primarily affect social preferences or first-order beliefs per se, but that the effects of framing may be based on more subtle mechanisms. For instance, based on our results we suspect that framing may affect higher-order beliefs, or trigger the contemplation of levels of reasoning in the first place. However, the magnitude and direction of a framing effect on higher-order beliefs may not be straightforward either, but rather depend on other context-dependent cues, too. The findings from Dufwenberg et al (2011) support both these two conjectures because they show that higher-order beliefs can be affected by framing, but not necessarily in expected ways. Furthermore, framing may also affect factors that are difficult to measure and formalize. For instance, one explanation for the observation that prosocial subjects in the COMMUNITY frame sustain high cooperation levels notwithstanding that neither their preferences, nor their beliefs or first-round contributions had been affected by the name of the game may be that the COMMUNITY frame somehow activated (or increased) their prosocial identity (for similar arguments along these lines, see for instance Bowles, 1998; Tett and Guterman, 2000). That is, the COMMUNITY frame may not have made subjects more prosocial, but may have increased the already prosocial subjects' tendencies to act

consistently according to their intrinsic prosocial preferences, irrespective of what others do. This explanation would be consistent with the pattern of results we observe and could also account for the fact that the predictive capacity of social preferences increases in the rounds of the repeated public goods game solely under the COMMUNITY frame. If the contributions of individualists –who may cooperate at the beginning of the iterated game mainly out of strategic considerations– decline over the course of the game while the contributions of prosocials stay high until the end, then both types would act more in line with their baseline social preferences at the end of the game as compared to the beginning of the game; and this is exactly what we observe.

Hence, research on framing effects in experimental games may have neglected –or at least may not yet have sufficiently emphasized– the subtleties that observations of framing effects depend on. Context, for instance, may turn out to be extremely important. It may be crucial whether experimental subjects know each other personally as was the case in the study by Liberman et al (2004), whether frames are mixed (take vs. give *and* community vs neutral), whether labels have particular connotations in different cultures or groups (see Dufwenberg et al, 2011), and many other things. All these factors may moderate the effect of framing on behavior in divergent and complicated ways, such that not only beliefs –of first or higher order– or preferences may (or may not) be affected by framing, but also the consistency and reliability with which these components inform behavior. Importantly, this insight also holds a warning against generalizing the effects of framing. A positive frame, such as a “community” label or “give-some” situation, may promote cooperation in some contexts, but may also have no effect or even detrimental effects in some other contexts.

Nevertheless, with caution, we interpret our data to suggest that a COMMUNITY frame *can* serve as an effective and low cost promoter of cooperation, at least in a repeated setting and for a substantial portion of the population. Effectively, social preferences can transform a social dilemma into a coordination problem, and *under certain conditions* frames can help promote this transformation and hence facilitate coordination. Two other reliable promoters of cooperation in these decision environments are (1) allowing subjects to punish each other (Fehr and Gächter (2002)), which carries high costs especially in the short term, and (2) allowing subjects to develop a reputation as a cooperator (Milinski and Rockenbach (2012); Chaudhuri (2011); Rockenbach and Milinski (2006)), clear implementation of which is challenging or even unrealistic in many large-scale real world situations. Engaging individuals based on their endogenous social preferences, and using framing to emphasize these preferences and their associated positive beliefs when legitimate, may provide policy makers and leaders a largely untapped avenue by which large-scale cooperation can be effectively and inexpensively encouraged.

An additional motivation for this work was to better understand to what extent a social preference measure, and behavior in a laboratory based public goods game, are predictive of cooperation toward a naturally-occurring public good; moreover to what extent these laboratory measures are externally valid (i.e., what can they predict in the wild?). We find that donations to a naturally-occurring public good were highest in the ENVI-

RONMENT and NEUTRAL frames, and were better predicted by social preferences (SVO scores) than by public goods game decisions or by answers to survey questions. These results indicate that the association between behavior in experimental social dilemmas and real-world cooperation, in the cases in which the former is predictive of the latter, is fully mediated by social preferences (SVO). This is evidence for SVO as a robust construct that generalizes between the laboratory and the wild.

We are not the first to explore whether results observed in laboratory games can be extended to predict decisions outside of the laboratory (e.g., Levitt and List (2007); Benz and Meier (2008); Antonovics et al (2009); Carpenter and Seki (2011)). In a closely-related study, Laury and Taylor (2008) test whether donations to a real-world public good (e.g., the purchase of trees to be planted in an urban area) were predicted by behavior in a lab-based public goods game. Our study departs from theirs by our inclusion of a separate validated social preference measure, which allows for disentangling the effects of preferences and beliefs. While their results indicate that subjects who made high average contributions to the laboratory public good were more likely to donate to the naturally-occurring public good, they did not collect information about subjects' preferences apart from their strategic public goods contributions. Our separate measure of social preference allowed us to discern among different explanatory variables in accounting for donation chances. Indeed, social preferences do emerge as the most useful predictor of donation decisions, fully mediating the relationship between public goods game decisions and donations to a real-world cause, suggesting that a simple social preferences model is a more direct way to account for real world cooperation than choice behavior in a laboratory public goods game.

References

- Ackermann KA, Fleiss J, Murphy RO (2014) Reciprocity as an individual difference. Forthcoming in *Journal of Conflict Resolution*
- Andreoni J (1995) Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. *The Quarterly Journal of Economics* 110(1):1–12
- Antonovics K, Arcidiacono P, Walsh R (2009) The Effects of Gender Interactions in the Lab and in the Field. *The Review of Economics and Statistics* 91(1):152–162
- Benz M, Meier S (2008) Do people behave in experiments as in the field? — evidence from donations. *Experimental Economics* 11(3):268–281
- Bowles S (1998) Endogenous Preferences : The Cultural Consequences of Markets and other Economic Institutions. *Journal of Economic Literature* XXXVI(March):75–111
- Brandts J, Schwiorek C (2009) Frames and Economic Behavior: An Experimental Study
- Brewer MB, Kramer RM (1986) Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology* 50(3):543–549
- Carpenter J, Seki E (2011) Do Social Preferences Increase Productivity? Field Experimental Evidence From Fishermen in Toyama Bay. *Economic Inquiry* 49(2):612–630
- Chaudhuri A (2011) Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics* 14(1):47–83
- Crosetto P, Weisel O, Winter F (2012) A Flexible z-Tree Implementation of the Social Value Orientation (SVO) Slider Measure (Murphy et al. 2011) - Manual
- Cubitt RP, Drouvelis M, Gächter S (2011) Framing and free riding: emotional responses and punishment in social dilemma games. *Experimental Economics* 14(2):254–272
- van Dijk E, Wilke H (2000) Decision-Induced Focusing in Social Dilemmas: Give-Some, Keep-Some, Take-Some, and Leave-Some Dilemmas. *Journal of Personality and Social Psychology* 78(1):92–104
- Dreber A, Ellingsen T, Johannesson M, Rand DG (2012) Do people care about social context? Framing effects in dictator games. *Experimental Economics* 16(3):349–371
- Dufwenberg M, Gächter S, Hennig-Schmidt H (2011) The framing of games and the psychology of play. *Games and Economic Behavior* 73(2):459–478
- Ellingsen T, Johannesson M, Mollerstrom J, Munkhammar S (2012) Social framing effects: Preferences or beliefs? *Games and Economic Behavior* 76:117–130
- Fehr E, Gächter S (2002) Altruistic punishment in humans. *Nature* 415(6868):137–40

- Fischbacher U (2007) z-Tree: Zurich toolbox for ready-made economic experiments
- Fischbacher U, Gächter S, Fehr E (2001) Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71:397–404
- Frederick S (2005) Cognitive Reflection and Decision Making. *Journal of Economic Perspectives* 19(4):25–42
- Gächter S, Renner E (2010) The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics* 13:364–377
- Kocher MG, Cherry T, Kroll S, Netzer RJ, Sutter M (2008) Conditional cooperation on three continents. *Economics Letters* 101(3):175–178
- Laury SK, Taylor LO (2008) Altruism spillovers: Are behaviors in context-free experiments predictive of altruism toward a naturally occurring public good? *Journal of Economic Behavior & Organization* 65(1):9–29
- Levitt SD, List JA (2007) What do Laboratory Experiments Measuring Social Preferences Tell Us About the Real World? *Journal of Economic Perspectives* 21(2):153–174
- Liberman V, Samuels SM, Ross L (2004) The name of the game: predictive power of reputations versus situational labels in determining prisoner’s dilemma game moves. *Personality & Social Psychology Bulletin* 30(9):1175–85
- List JA (2007) On the Interpretation of Giving in Dictator Games. *Journal of Political Economy* 115(3):482–493
- McCusker C, Carnevale P (1995) Framing in Resource Dilemmas: Loss Aversion and the Moderating Effects of Sanctions. *Organizational Behavior and Human Decision Processes* 61(2):190–201
- Milinski M, Rockenbach B (2012) On the interaction of the stick and the carrot in social dilemmas. *Journal of Theoretical Biology* 299:139–43
- Murphy RO, Ackermann KA (2014a) Explaining behavior in public goods games: How preferences and beliefs affect contribution levels. Manuscript in Preparation
- Murphy RO, Ackermann KA (2014b) Social value orientation theoretical and measurement issues in the study of social preferences. *Personality and Social Psychology Review* 18:13–41
- Murphy RO, Ackermann KA, Handgraaf M (2011) Measuring Social Value Orientation. *Judgment and Decision Making* 6(8):771–781
- Park ES (2000) Warm-glow versus cold-prickle: a further experimental study of framing effects on free-riding. *Journal of Economic Behavior & Organization* 43(4):405–421
- Rabin M (1998) Psychology and Economics. *Journal of Economic Literature* 36(1):11–46

- Rockenbach B, Milinski M (2006) The efficient interaction of indirect reciprocity and costly punishment. *Nature* 444(7120):718–23
- Selten R (1967) Die Strategiemethode zur Erforschung des Eingeschränkten Rationalen Verhaltens im Rahmen eines Oligopolexperimentes. In: H Sauermann (ed) *Beiträge zur Experimentellen Wirtschaftsforschung*, J.C.B. Mohr, Tübingen, pp 136–168
- Sonnemans J, Schram A, Offerman T (1998) Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization* 34(1):143–161
- Tett RP, Guterman HA (2000) Situation trait relevance, trait expression, and cross-situational consistency: Testing a principle of trait activation. *Journal of Research in Personality* 34(4):397–423
- Tversky A, Kahneman D (1981) The Framing of Decisions and the Psychology of Choice. *Science* 211(4481):453–458
- Weber JM, Kopelman S, Messick DM (2004) A conceptual review of decision making in social dilemmas: applying a logic of appropriateness. *Personality and Social Psychology Review* 8(3):281–307