Journal of Economic Psychology 31 (2010) 285–292

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

# Voting on a sharing norm in a dictator game

Christoph Vanberg *

*Department of Economics, University of Heidelberg, Bergheimer Str. 58, 69115 Heidelberg, Germany*

ARTICLE INFO

*Article history:*
Received 29 October 2007
Received in revised form 10 October 2008
Accepted 29 October 2009
Available online 11 November 2009

*JEL classification:*
C91
C92
D63
D64
D70

*PsycINFO classification:*
2300
2900
3000
2360
2960
3020

*Keywords:*
Dictator game
Communication
Voting
Agreements
Guilt aversion
Fairness
Obligations

ABSTRACT

A large body of experimental research has documented that fairness and cooperation in experimental games are often enhanced when subjects can (and do) make non-binding pre-play agreements. Many studies have demonstrated this phenomenon for verbal agreements reached via open discussion. In contrast, little evidence exists concerning the effects of agreements reached using formal procedures that do not allow for discussion. This paper begins to fill this gap by investigating the effectiveness of an agreement reached by majority voting. Specifically, I investigate how voting on a sharing norm affects subsequent behavior in a dictator game. The voting procedure takes place behind a 'veil of ignorance', and the result of the vote is referred to as a non-binding agreement. I find that this procedure does not induce higher offers than are observed under a no-vote baseline. In fact, dictators were significantly more likely to offer nothing under the voting treatment. This evidence suggests that formal decision procedures may be less effective than open deliberation in creating effective non-binding agreements.

© 2009 Elsevier B.V. All rights reserved.

## 1. Introduction

A large and growing literature in experimental economics shows that fairness and cooperation in experimental games are enhanced when subjects can engage in pre-play cheap talk communication (Bornstein, Rapoport, Kerpel, & Katz, 1989; Brosig, Weimann, & Ockenfels, 2003; Charness & Dufwenberg, 2006; Dawes, Mctavish, & Shaklee, 1977; Ellingsen & Johannesson, 2004; Kerr & Kaufman-Gilliland, 1991; Orbell, van de Kragt, & Dawes, 1988; Ostrom, Walker, & Gardner, 1992). Indeed, as emphasized by Brosig et al. (2003), "one of the few variables that is known to have a robust and strong positive effect on the level of cooperation is the opportunity to communicate" (see also Sally, 1995; Walker & Ostrom, 2007).

* Tel.: +49 6221 54 2947.
*E-mail address:* vanberg@uni-hd.de

0167-4870/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.joep.2009.10.006

As a typical example, Orbell et al. (1988) report on a public good experiment in which subjects were given the opportunity to engage in group discussion prior to interacting. Relative to a no communication benchmark, contributions were significantly enhanced, especially when promises had been exchanged. Ellingsen and Johannesson (2004) find that the exchange of promises enhances trust and trustworthiness in a two-player setting, suggesting that people have a preference for keeping their word. Charness and Dufwenberg (2006) collect similar evidence and suggest that the theory of *guilt aversion* (Dufwenberg & Gneezy, 2000; Battigalli & Dufwenberg, 2007) may account for players' tendency to uphold agreements.[1]

Notions such as commitment or guilt aversion may potentially explain why people feel motivated to abide by the terms of a contract *once is has been established*. However, a point worth emphasizing is that neither of these concepts helps to identify the *procedures* necessary for creating something that will be perceived as constituting a contract. Applying these theories to the experiments mentioned above requires us to accept the additional premise that a certain type of verbal behavior ("making a promise") causes expectations and/or obligations to be established. However, the theories themselves are silent as to what actually constitutes an instance of promise-making. Similarly, both guilt aversion and commitment may explain norm-abiding behavior. However, neither concept helps to identify the *process* by which a social norm is established. Therefore, even if we were to accept one or the other of these theories, we would still be left with the question: how are effective agreements actually created?

Existing studies have focused mostly on agreements reached through verbal (spoken or written) communication.[2] This procedure is justified if the researcher's goal is to prove the basic point that non-binding agreements *can* be effective. Open discussion seems intuitively to maximize the chance that subjects can establish an agreement with which they identify. Once the effectiveness of such agreements has been established, a question worth turning to is whether 'weaker' decision procedures can have similar effects. Such procedures are interesting because there are many real-life situations in which agreements in a group cannot feasibly be reached via discussion, and instead must rely on formal procedures.

Arguably the most common mechanism for collective agreement in real-world situations is majority voting. The current paper seeks to investigate the effectiveness of this procedure for creating a non-binding agreement or norm. Specifically, I investigate whether behavior in a dictator game is affected when, prior to interaction, subjects vote on a non-binding sharing norm. Previous research has shown that verbal communication in the form of written messages can have a significant positive effect on dictator offers (Mohlin & Johannesson (2008)). My goal here is to investigate whether a more formal procedure not involving verbal communication can have similar effects.[3]

The advantage of the dictator game setting is that it is a non-strategic situation and removes all coordination aspects of an agreement. The only reason for a subject to abide by an agreement in this context is that she feels obligated to do so. The design is based on the idea that participants may prefer to establish a norm of sharing when placed behind a "veil of ignorance". Therefore, voting takes place before subjects are told whether they will be the dictator or the recipient in the subsequent interaction.

My main hypothesis is that the opportunity to establish a norm using this mechanism causes dictator subjects to behave more 'fairly' in the subsequent game. Specifically, I test the hypothesis that the distribution of dictator offers is shifted to the right by the voting treatment, as compared to a no-vote benchmark. An alternative approach would have been to test whether the distribution of dictator offers is shifted in the direction of whatever agreement is reached. The problem with this is that the agreement is endogenous, and thus correlated with unobserved characteristics of the respective session's subject population. Therefore, it becomes impossible to causally interpret a correlation of behavior with the agreement itself. For this reason, I instead test whether offers are positively affected by the voting treatment as such.[4]

Surprisingly, the results of the experiment point in the *opposite* direction. Offers under the voting treatment were somewhat lower than under the baseline. A significantly higher proportion of dictators offered zero, as predicted by the standard model of narrowly self interested behavior. This evidence suggests that agreements arrived at through a formal mechanism such as the one employed may be less effective than those produced by open and informal deliberation.

The next section presents the experimental design. Section 3 describes the results, comparing behavior as well as measures of beliefs and feelings of obligation under the two treatments. Section 4 concludes and briefly outlines a possible explanation for the absence of a positive effect on dictator offers. The Appendix contains a translation of instructions.

[1] An interesting difference between these approaches lies in the causal 'proximity' ascribed to agreements in shaping motivations. According to Charness and Dufwenberg (2006), agreements may affect motivations and behavior *indirectly*. Specifically, it is not an agreement *per se* that motivates one player ($i$) to cooperate with another ($j$), but rather her ($i$'s) belief concerning what the other ($j$) expects to receive. Agreements can be effective because they can change these (second order) beliefs. In contrast, Ellingsen and Johannesson (2004) advocate a theory of genuine "commitment" to contracts, which posits that people are *directly* affected by agreements because they feel obligated to abide by them. This distinction will not be pursued in what follows. See Vanberg (2008) for evidence supporting the notion of genuine commitment.

[2] Subjects are usually given the opportunity to communicate verbally. Verbal communication can be implemented in a number of ways, including face-to-face discussions, anonymous chatting, one-way messages, and audio or video conferences. See Brosig et al. (2003) for a study comparing the associated effects in the context of a public goods experiment.

[3] More generally, behavior in dictator games has been shown to be sensitive to various contextual manipulations such as the degree of anonymity or the framing of the situation (see, e.g. Bohnet & Frey, 1999; Bolton, Katok, & Zwick, 1998; Branas-Garza, 2007; Hoffman, McCabe, & Simth, 1996). It thus constitutes a paradigmatic tool for measuring the effectiveness of contextual factors in promoting other-regarding or fairness-related motivations.

[4] Note also that the median offer observed under the baseline was 3 EUR (out of 10), so that the norms agreed upon indeed called for larger dictator offers.

## 2. Experimental design

### 2.1. Baseline treatment

Subjects entered the laboratory and were randomly assigned to isolated terminals. Instructions (reproduced in the appendix) were handed out and read aloud by the experimenter (myself). The instructions were the same for all subjects and informed them that they would be randomly matched with another subject, that one of them would then be randomly chosen to be 'subject A', and that this subject would then have the opportunity to choose an integer-valued distribution of 10 EUR between the two subjects. They were informed that their payment at the end of the session would consist of a 5 EUR show-up fee plus the payoff from the experiment. Questions were answered privately at the subjects' seat (again, by myself).

Subjects A then saw a screen displaying a table of all possible integer distributions of 10 EUR. While they made their choices, subjects B saw the same table and were asked to report a hypothetical choice, had they been subject A.[5] After both had made their choices, subjects A were asked what they thought B expected to receive. Subjects B were asked what they expected to receive. The elicitation of beliefs was not incentivized.[6]

After choices and beliefs were submitted, all subjects were paid privately at their seats, with cash placed in envelopes. This procedure guaranteed that no subject would witness other subjects being paid. Subjects were informed of this procedure in the instructions.

### 2.2. Voting treatment

The voting treatment was exactly like the baseline treatment except for the following features. Instructions for the dictator game were marked 'Instructions for phase 2'. These were passed out first and read out loud, as before. In contrast to the baseline treatment, the instructions contained a sentence asking subjects to think about 'what subject A should do' in the situation described. Questions were answered before the instructions for phase 1 were handed out. These described the voting procedure. Subjects were told that they would have the opportunity to vote on a 'non-binding rule' prescribing how subjects A 'should' behave in phase 2. Subjects did not know at this time which role they would ultimately be assigned to. Thus, voting on the sharing rule took place behind a 'veil of ignorance'.

The voting procedure worked as follows. A table displaying all integer allocations of the available surplus was displayed on the subjects' screens. Next to the table, a green arrow initially pointed at a randomly chosen allocation. Each subject could then vote to move the arrow up or down, or to keep it where it was. The instructions asked subjects to truthfully state their preferences concerning the agreement, i.e. they said 'If you think that A should choose a distribution further down in the table, please vote *down*', etc. and 'If you agree with the distribution indicated, please vote *stay*'. If a majority of subjects voted to move the arrow in the same direction, it was moved and another vote was taken. After each round of voting, subjects were informed about the aggregate number of 'up', 'down' and 'stay' votes as well as the implied movement of the arrow.[7] This procedure was repeated until the arrow was not moved for two consecutive rounds.[8]

After voting had ended, subjects were informed verbally that phase 1 was now over and phase 2 would begin. The rest of the experiment worked exactly as in the baseline, except that the table of distributions displayed on the screen was accompanied by a red arrow pointing at the 'norm', and all on-screen instructions contained the sentence 'The red arrow indicates the rule agreed upon in the previous phase'.

### 2.3. Details

The experiment was conducted in the experimental laboratory of the Max Planck Institute of Economics in Jena, Germany. Subjects were students of the Friedrich-Schiller University of Jena. Subjects were recruited using the online recruitment system ORSEE (Greiner, 2004). The experiment was programmed and conducted with the software *z*-Tree (Fischbacher, 2007).

We ran a total of ten sessions, four baseline sessions and six voting sessions. All sessions except two involved 16 participants, i.e. 8 pairs. Due to subjects failing to show up, one voting session involved only 12 participants (6 pairs), and one baseline session involved only 14 participants (7 pairs). Thus, the data collected include 31 offers made under the baseline and 46 offers made under the voting treatment.

[5] This procedure was chosen mainly in order to give subjects B 'something to do' during the submission of choices, thereby guaranteeing that subjects A could not be localized by the sound of mouse clicks.

[6] In experimental studies which focus on the interaction of beliefs and behavior, it is common to incentivize the elicitation of beliefs. In the present study, my main focus is on the effects of the voting treatment on subjects' subsequent behavior. I chose not to incentivize beliefs because I felt that the extra instructions and additional source of earnings might have distracted subjects' attention from the central task. In particular, a 'correct' elicitation of beliefs using the quadratic scoring rule is rather complicated when choices are not binary. See Sonnemans and Offerman (2001) for a discussion of this issue. They collect evidence indicating that the quality of elicited beliefs is not significantly improved by incentives.

[7] Subjects were never informed about the vote cast by any particular individual. That is, voting was anonymous.

[8] The latter rule ensured that subjects would not agree on something by accident. It also allowed subjects to 'ratify' an agreement that had been reached.

**Table 1**
Voting on the norm.

| Session | Norm | Stay votes |
|---|---|---|
| V1 | 5 EUR | 10 (of 12) |
| V2 | 5 EUR | 13 (of 16) |
| V3 | 4 EUR | 6 (of 16) |
| V4 | 4 EUR | 7 (of 16) |
| V5 | 5 EUR | 10 (of 16) |
| V6 | 5 EUR | 14 (of 16) |

**Table 2**
Dictator offers (out of 10 EUR).

| Offer | Baseline | Voting | Norm of 4 | Norm of 5 |
|---|---|---|---|---|
| 0 EUR | 5 (16%) | 17 (37%) | 4 (25%) | 13 (43%) |
| 1 EUR | 4 (13%) | 3 (7%) | 3 (18%) | 0 (0%) |
| 2 EUR | 4 (13%) | 2 (4%) | 1 (6%) | 1 (3%) |
| 3 EUR | 6 (19%) | 6 (13%) | 1 (6%) | 5 (17%) |
| 4 EUR | 3 (10%) | 9 (20%) | 4 (25%) | 5 (17%) |
| 5 EUR | 9 (29%) | 9 (20%) | 3 (19%) | 6 (20%) |
| | ($N$ = 31) | ($N$ = 46) | ($N$ = 16) | ($N$ = 30) |

## 3. Results

### 3.1. Voting on the norm

Table 1 summarizes the voting stages in each of the voting sessions. In 4 of the 6 sessions (V1, V2, V5, and V6) subjects agreed upon a norm of sharing 50% of the pie. In sessions V3 and V4, the collective agreement was to share 40%. Table 1 also shows the number of subjects voting to 'stay' in the final round of voting. According to this measure, the norm of 4 was associated with a lower level of agreement in the groups. (Recall that this round is essentially about confirming the decision that the group has arrived at.)

### 3.2. Distribution of offers

Table 2 displays the distribution of offers made under the baseline and voting treatments, as well as detailed results for each of the sharing norms. A look at these data immediately suggests that the voting treatment did not cause dictators to share a larger fraction of the pie. Indeed, the most striking feature of these data is that a very large number of subjects (43%) decided to give nothing after a norm of giving 5 had been agreed on (as compared to 16% under the baseline).

These impressions are supported by nonparametric tests. In order to test the hypothesis that the offers under the voting treatment are the same as under the baseline condition, I compare the distribution of offers using rank-sum (RS) and Kolmogorov–Smirnov (KS) tests. Both fail to reject the hypothesis that all offers are drawn from the same underlying distribution (RS: $p = 0.25$, KS: $p = 0.31$).[9] We can therefore reject the hypothesis that voting had a positive effect on dictator offers.

Indeed, the numbers reported in Table 2 suggest that the number of subjects keeping the entire amount for themselves was significantly larger after voting on a norm (see Table 3). This conjecture can be tested using Fisher's exact (FE) test or a Chi-squared test (CHI2). Using these tests, it indeed appears that a significantly larger number of dictators offered nothing under the treatment (both tests: $p = 0.04$).[10]

To summarize, the evidence suggests that voting on a sharing norm did not have a positive effect on dictator offers. Surprisingly, the opposite appears to be the case. Specifically, it appears that a significantly larger number of subjects chose to offer zero under the treatment, as would be expected based on 'standard' theory.[11]

[9] These tests assume that the observations are statistically independent. Some clarification may be required to justify their use in our context. The 31 baseline offers constitute independent observations because they are one shot decisions. The 46 offers made under the voting treatment are not independent, in the sense that subjects jointly experienced the voting stage prior to submitting their choices. However, the subsequent choices are still one shot, and therefore the only source of statistical dependence is the treatment condition itself. We are therefore justified in treating these observations as independent if the treatment condition was similar in each session. This assumption may not be justified when we pool data from voting sessions which involved agreement on different norms. In order to verify the robustness of our inference, we can separately compare the baseline distribution to those obtained under a norm of 4 and 5, respectively. These tests confirm the result reported in the text. The Baseline vs. Norm of 5 comparison involves 31 vs. 30 independent observations and yields: RS $p = 0.25$, KS $p = 0.15$; The Baseline vs. Norm of 4 comparison involves 31 vs. 16 independent observations and yields: RS $p = 0.52$, KS $p = 0.95$.

[10] Separately comparing the baseline condition to sessions in which norms of 4 and 5 were established, we find that the number of zero offers was significantly larger in groups that agreed on a norm of 5 (FE and CHI2 $p = 0.02$). Under a norm of 4, the difference is not significant (FE $p = 0.36$, CHI2 $p = 0.46$). The latter may reflect the fact that there are relatively few observations.

[11] More detailed tests at the individual level do not reveal a systematic relationship between voting and subsequent dictator behavior. That is, it is not the case that subjects who agreed to a norm behaved differently from those who did not.

**Table 3**
Number of dictators offering zero.

| | Baseline | Voting | Norm of 4 | Norm of 5 |
|---|---|---|---|---|
| $s = 0$ | 5/31 (16%) | 17/46 (37%) | 4/16 (25%) | 13/30 (43%) |
| $s > 0$ | 26/31 (84%) | 29/46 (63%) | 12/16 (75%) | 17/30 (57%) |

### 3.3. Beliefs and obligations

The introduction discussed two theories that have been proposed to account for the effectiveness of non-binding verbal agreements. The theory of guilt aversion emphasizes that such promises may affect subjects' second order beliefs (i.e. beliefs about what other subjects expect to receive). The theory of commitment emphasizes that people may have a preference for 'keeping their word', i.e. for living up to their obligations.

As discussed above, neither of these theories offers predictions as to which processes are capable of inducing similar changes in beliefs and/or obligations. None the less, they do offer perspectives on *why* a given procedure does or does not work. In order to understand why the voting procedure investigated here did not induce higher dictator offers, we can investigate whether voting affected subjects' beliefs as measured, or their feelings of obligation as reported in a post-experimental questionnaire.

Fig. 1 displays the distributions of second order beliefs by treatment. These distributions are virtually identical. Thus, the ineffectiveness of the voting treatment in this experiment is consistent with the theory of guilt aversion.

Fig. 2 shows dictator subjects' reported agreement with the statement that 'I felt an obligation to share the money with B'. There is no statistically significant difference in the distribution of answers given under the baseline and voting treatments

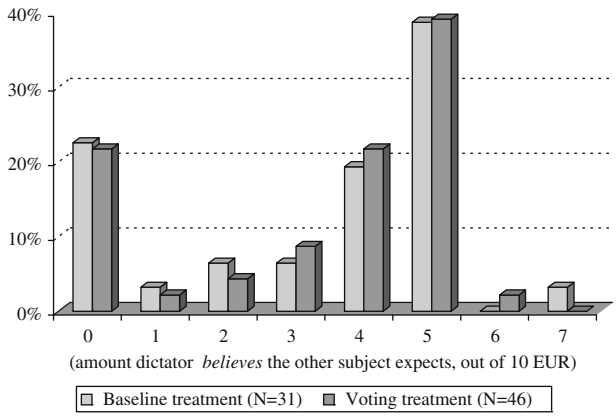


**Fig. 1.** Dictator subjects' second order beliefs.

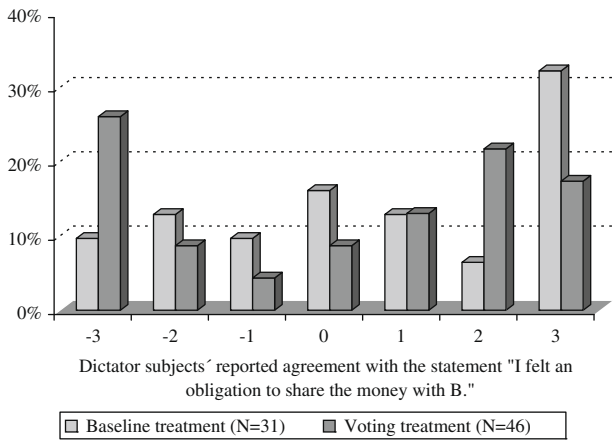


**Fig. 2.** Dictator subjects' feelings of obligation.

(RS $p = 0.3$, KS $p = 0.62$). If anything, it appears that dictators were *less* likely to agree with this statement after the voting procedure.

## 4. Conclusion

The experiment reported on in this paper sought to test whether majority voting can enhance fairness concerns and induce higher offers in a dictator game. Previous studies have documented that (written) verbal communication can enhance fairness in this setting. The hypothesis to be tested was that similar effects would be observed when agreement is reached using a formal procedure without verbal communication, namely majority voting. Surprisingly, the data suggests the opposite conclusion. In particular, the overall distribution of offers was not positively affected by the voting treatment. Indeed, the number of subjects keeping the entire amount was significantly higher after voting than under the no-vote baseline condition.

A possible – yet purely speculative – explanation for this result is that dictators experience a fixed payoff from abiding by a norm. When an excessively demanding norm is established by the group, it may make this payoff too 'expensive' for those who would otherwise have perceived (and abided by) a less demanding norm. As a consequence, dictators who are otherwise inclined to share moderate amounts (less than %40) may have felt overburdened and therefore kept the whole amount for themselves.[12]

Another explanation is that subjects may come to the laboratory holding certain *personal* norms, and that the establishment of an explicit *social* norm disrupts these convictions. There are at least two reasons why such an effect may occur. First, being asked to formally establish a norm may destroy a subject's confidence in the appropriateness of a personal norm.[13] Second, norm compliance may be motivated by automatic and spontaneous emotional reactions. The explicit establishment of a norm may induce subjects to consciously reflect upon their incentives, crowding out unreflected emotional responses.[14]

The result of this experiment suggests that agreements reached using formal decision procedures may be less effective than those which arise out of open discussion. The experiment also raises questions concerning the usefulness of establishing explicit social norms in areas where pro-social behavior may be motivated by heterogeneous personal norms. In particular, explicit norms may crowd out intrinsic motivations to abide by personal norms. This interpretation is speculative. Further experimental studies should aim at exploring these effects.

## Acknowledgements

I thank Luis Miller for the discussions which lead to the idea for this project. The comments of an anonymous referee helped to greatly improve the paper. Thanks to Frederic Bertels for assistance with programming and organizing experimental sessions. This research was funded by the Max Planck Society.

## Appendix A

### A.1. Translation of instructions

#### A.1.1. General instructions (all treatments)

The following instructions were printed on a single page placed at each seat before subjects entered the laboratory.

*Instructions*

Welcome. Please carefully read the following instructions.

*General rules*

[12] Concretely, suppose that a decision maker is asked to divide a pie of size $X$. Suppose that she compares her action (the share $s_i$ that she gives away) to what she perceives as the *norm* (the share $n$ that she 'should' give away), and that she receives a payoff $\alpha_i > 0$ from abiding by the norm. Assume that this 'psychological' reward is simply added to her material payoff. Then, her utility from offering a share $s_i$ of the pie is equal to

$$U_i(s_i, n) = \begin{cases} (1 - s_i) \cdot X & \text{if} \quad s_i < n, \\ (1 - s_i) \cdot X + \alpha_i & \text{if} \quad s_i \geqslant n. \end{cases}$$

Then, subject $i$'s optimal choice is given by

$$s_i^*(n) = \begin{cases} 0 & \text{if} \quad \alpha_i < n \cdot X, \\ n & \text{if} \quad \alpha_i \geqslant n \cdot X. \end{cases}$$

It follows that this subject will follow a norm of sharing $n$ only as long as $n \leqslant \overline{n}_i = \frac{\alpha_i}{X}$. The establishment of any norm $n > \overline{n}_i$ will cause such a subject to offer zero.

[13] Miller (2006) suggests a similar explanation for the confounding effects of introducing normative principles in the context of an asymmetric coordination game.

[14] Krupka and Weber (2005) report on an experiment in which subjects behaved more generously after thinking about or viewing the behavior of others. In that experiment, subjects were probably consciously processing the treatment stimulus. However, acts of generosity were not explicitly *labeled* as norm compliance.

The experiment will last for approximately 30 min. During this time, we ask you to abide by the following rules:

- Do not speak to the other participants.
- Turn off and stow away your cellular phone.
- Stow away any reading or writing materials. Starting now, your table should contain only these instructions.
- In case you should have questions at any time, please raise your hand and wait until an assistant comes to your table.

*Payment at the end of the experiment*

Regardless of the outcome of the experiment, each participant will receive a minimum payment of 5 EUR. Your total earnings may depend on your own decisions and those of other participants.

At the end of the experiment, please remain quietly at your seat. Payment is conducted at your seat. The following procedure ensures that no other participant will learn what you have earned:

- An assistant will bring you an envelope and a receipt.
- Please verify *immediately* that the content of the envelope corresponds to the amount indicated on the receipt.
- Sign the receipt and return it to the assistant.
- Quietly leave the room.

Participants who do not abide by these rules will be excluded from the experiment and payment.

In a few moments, you will receive additional instructions regarding the specifics of the experiment.

*A.1.2. Baseline treatment*

The following instructions were printed on a single sheet and handed out after subjects were given time to read the general instructions. After giving subjects time to read, these instructions were read out loud by the experimenter. After this, subjects were again given time to read the instructions and ask questions. Questions were answered privately at the subjects' tables.

*Information regarding the experiment*

The 16 participants present will be randomly divided into 8 pairs. From each pair, one participant will be (once again, randomly) determined to be participant 'A', and the other to be participant 'B'. Participant A will then decide how 10 EUR will be divided between himself and participant B. Only integer numbers will be allowed. That is, A will choose a distribution from the following table:

Possible distributions of the 10 EUR.

| A receives | B receives |
|---|---|
| 10 EUR | 0 EUR |
| 9 EUR | 1 EUR |
| 8 EUR | 2 EUR |
| 7 EUR | 3 EUR |
| 6 EUR | 4 EUR |
| 5 EUR | 5 EUR |
| 4 EUR | 6 EUR |
| 3 EUR | 7 EUR |
| 2 EUR | 8 EUR |
| 1 EUR | 9 EUR |
| 0 EUR | 10 EUR |

The amounts assigned by A (plus the 5 EUR mentioned above) will be payed out to the participants at the end of the experiment.

*A.1.3. Voting treatment*

In the voting treatment, subjects first received the same instructions as above, with the following modifications:

- A sentence was added at the top of the page that read, "The experiment consists of two phases. For reasons that will become clear in a moment, we first describe the second phase".
- Beneath this sentence, an additional heading in bold read, "Second phase":
- At the bottom of the page, a sentence was added that read, "Before instructions for the second phase are passed out, we ask you to think about the following question: In your opinion, how should a participant in role A behave?"

As in the baseline treatment, these instructions were read out loud and subjects were encouraged to ask questions before proceeding. After questions had been answered, the following instructions for phase 1 were passed out. Again, subjects were given time to read, then instructions were read aloud, after which again subjects could read and ask questions.

*First phase*

In the first phase, the participants present will have the opportunity to agree, by way of voting, on a non-binding rule according to which those chosen as participant 'A' in phase 2 should behave. At the time of the vote, no participant will know which role (A or B) he himself will occupy.

The vote will proceed according to the following rules:

All participants will see the table printed on the preceding page on their screen. An arrow next to the table will initially point at a randomly chosen allocation. If it is your opinion that A should choose an allocation further down in the table, please choose 'down'. If it is your opinion that A should choose an allocation further up in the table, please choose 'up'. If you agree with the allocation indicated, please choose 'agree'.

If a majority of the participants present (at least 9 people) vote for a movement in thea same direction, the arrow will be shifted accordingly. If no such majority exists, the arrow will remain in place.

This procedure will be repeated until the arrow is not shifted in two consecutive rounds.

After each round of voting, you will learn the number of participants who have voted for each of the options.

The allocation rule agreed upon in this fashion will continue to be indicated by an arrow in phase 2. This rule is non-binding. That is, those persons chosen as participant A will still be able to choose any allocation in the table.

## References


Battigalli, P., & Dufwenberg, M. (2007). Guilt in games. *American Economic Review Papers and Proceedings, 97*, 170–176.

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

Bolton, G. E., Katok, E., & Zwick, R. (1998). Dictator game giving: Rules of fairness versus acts of kindness. *International Journal of Game Theory, 27*, 269–299.

Bornstein, G., Rapoport, A., Kerpel, L., & Katz, T. (1989). Within- and between-group communication in intergroup competition for public goods. *Journal of Experimental Social Psychology, 25*(5), 422–436.

Branas-Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology, 28*(4), 477–486.

Brosig, J., Weimann, J., & Ockenfels, A. (2003). The effect of communication media on cooperation. *German Economic Review, 4*(2), 217–241.

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica, 74*(6), 1579–1601.

Dawes, R., Mctavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology, 35*, 1–11.

Dufwenberg, M., & Gneezy, U. (2000). Measuring beliefs in an experimental lost wallet game. *Games and Economic Behavior, 30*, 163–182.

Ellingsen, T., & Johannesson, M. (2004). Promises, threats and fairness. *Economic Journal, 114*, 397–420.

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

Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen 2003* (pp. 79–93). Göttingen: GWDG Bericht 63. Ges. für Wiss. Datenverarbeitung.

Hoffman, E., McCabe, K., & Simth, V. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review, 86*, 653–690.

Kerr, N., & Kaufman-Gilliland, C. (1991). Communication, commitment, and cooperation in social dilemmas. *Journal of Personality and Social Psychology, 66*(3), 513–529.

Krupka, E., & Weber, R. (2005). The focusing and observational effects of norms. *Working paper*. Carnegie Mellon University.

Miller, L. (2006). The double nature of conventions – An experimental analysis. Papers on Strategic Interaction No. 7-2006. Max Planck Institute of Economics.

Mohlin, E., & Johannesson, M. (2008). Communication: Content or relationship? *Journal of Economic Behavior and Organization, 65*, 409–419.

Orbell, J., van de Kragt, A., & Dawes, R. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology, 54*(5), 811–819.

Ostrom, E., Walker, J., & Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *The American Political Science Review, 86*(2), 404–417.

Sally, D. (1995). Conversation and cooperation in social dilemmas. A meta-analysis of experiments from 1958 to 1992. *Rationality and Society, 7*, 58–92.

Sonnemans, J., & Offerman, T. (2001). Is the quadratic scoring rule really incentive compatible? *Working paper*. University of Amesterdam.

Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica, 76*(6), 1467–1480.

Walker, J., & Ostrom, E. (2007). Trust and reciprocity as foundations of cooperation: Individuals, institutions, and context. *Working paper*.