

Koessler, Ann-Kathrin; Page, Lionel; Dulleck, Uwe

**Working Paper**

## Public Statements of Good Conduct Promote Pro-Social Behavior

*Suggested Citation:* Koessler, Ann-Kathrin; Page, Lionel; Dulleck, Uwe (2018) : Public Statements of Good Conduct Promote Pro-Social Behavior, SSRN, Rochester, <https://doi.org/10.2139/ssrn.3184836>

This Version is available at:

<https://hdl.handle.net/10419/180669>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Public statements of good conduct promote pro-social behavior

Ann-Kathrin Koessler<sup>1</sup>, Lionel Page<sup>2</sup>, and Uwe Dulleck<sup>2</sup>

<sup>1</sup>Institute of Environmental Systems Research, University of Osnabrueck, Germany

<sup>2</sup>School of Economics and Finance, Queensland University of Technology, Australia

## Abstract

Voluntary and compulsory public statements of good conduct are frequently observed in the real world, such as the codes of good conduct for professionals or the requirements of academic journals to affirm that research was carried out ethically.

In this study, we investigate what effect public statements of good conduct have on contribution behavior in a public goods experiment. Using a ‘between-within subjects design’ we identify three channels by which non-enforceable statements of intent are associated with higher levels of contributions to the public good. First, in a selection effect, socially-oriented participants are more likely in the experiment to make a public statement. Second, in a commitment effect, participants who make a public statement are contributing more to the public good. Third, in a coordination effect, aggregate contributions are higher when ‘Statement-Makers’ observe that also other group members make the statement. The latter explains why compulsory statements of good conduct are in our experiment more effective over time.

*Keywords:* social dilemma, pro-social behavior, commitment, compliance, public good, policy making

*JEL Classification:* A13, C72, C91, H41

---

\*Koessler (corresponding author): akoessler@uos.de; +49-541-969-2424; Barbarastr. 12, 49076 Osnabrueck, Germany. Page and Dulleck: School of Economics and Finance, QUT, 2 George Street, 4000 Brisbane, Qld, Australia.

# 1 Introduction

‘Codes of conduct’ are widely used in practice to promote pro-social behavior. The ‘Banker’s oath’, the ‘Hippocratic oath’ or the requirement of many academic journals for authors to affirm that research was carried out ethically can serve as examples. These pledges can be understood as public statements of good conduct which are used to motivate individuals to remember their social orientation and consequently contribute to the public good. Given that these statements are a relative low cost policy instrument, we believe that it is important to understand how they affect behavior.

In many industries individual actors can gain private benefits by deviating from professional standards: a doctor or lawyer may charge for services not provided, a researcher may manipulate data, a bank employee may give advice that is not in the client’s best interest. These behaviors have often been analyzed and simply judged as criminal or unethical behavior in bilateral relationships. However, unethical conduct not only carries costs for the client or patient, but also implies potential negative externalities for the entire profession. The detection of academic misconduct, for instance, can lead to a loss of credibility in the entire scientific community (see Ioannidis (2005); Martinson et al. (2005); Fanelli (2009) for a general discussion on scientific fraud; Begley and Ellis (2012) for unreproducible ‘landmark’ studies in cancer research; List et al. (2001); Necker (2014) for a discussion in economic research.). Tighter regulations, and fewer investments or client interactions may result from the erosion of professional integrity. And even without obvious and publicized misconduct negative effects are present. In the case of scientific research, unethical methods, for example, limit the progress in the field.

On the basis of these market externalities we argue that professional conduct possesses the characteristics of a public good. In practice, one approach for addressing misconduct issues is to require oral and written statements from actors in the field in question, promising proper, pro-social conduct. Statements of this kind are either voluntary or compulsory. Whether a voluntary or compulsory statement is the better choice is surely context dependent, but we will show in this study that this choice also determines how the individuals get affected by the statement.

To better understand the link between public statements of good conduct and resulting changes in behavior we designed a public good experiment. Existing experiments focus on what effect statements of good conduct have in bilateral interactions (see for example Ellingsen and Johannesson, 2004, Charness and Dufwenberg, 2006; and Vanberg, 2008). The evidence for the

group context is limited. In our study, we focus on the wider social dimension of good conduct and examine whether non-enforceable compliance statements can reduce free-riding in a social dilemma situation. The experiment allows us to identify three ways in which public statements of good conduct can be associated with higher levels of contributions to a public good: a selection effect, a commitment effect, and a coordination effect.

The selection effect refers to the fact that the statement is more likely to be made by individuals who contribute to the public good anyway. We are able to disentangle this effect by starting our experiment with a baseline session before introducing the possibility make a pledge. We then identify the real effect of the public statement by comparing the levels of contributions before and after the statement introduction and between treatment groups. Changes in behavior after the pledge are due to two origins: first, a preference for keeping one’s word leads to a direct commitment effect and second, observing others making (also) the statement helps to coordinate on better group outcomes. This coordination effect is particularly dominant when the dynamic development of contributions is taken into account.

The remainder of the paper is organized as follows. In Section 2, we review the related literature. In Section 3, we present the experimental design. Behavioral predictions are listed in Section 4, followed by the experimental results in Section 5. Finally, we conclude with some discussion and potential insights for management practice in Section 6.

## 2 Related Literature

Academic interest on codes of good conduct as institutionalized coordination devices is only just emerging (de Bruin and Dolfsma, 2013). In economic experiments, Jacquemet and colleagues introduced a voluntary solemn oath to tell the truth during experiments. Following the oath, participants were more likely to reveal their true preferences (Jacquemet et al., 2013), less likely to lie (Jacquemet et al., 2018) and coordination improved through the more truthful communication (Jacquemet et al., 2017). Also Mazar et al. (2008) examined the behavioral impact of codes of good conduct. When participants were told that the study fell under the university’s honor code, significantly fewer students cheated in a self-reporting task. Shu et al. (2012) examined the effect of good conduct declarations in the field and found that signing an honesty statement at the beginning of a tax declaration rather than at the end promotes truthful reporting.

Contrary to these studies which investigate the effect of a vow on honesty,

we are interested in the impact of a public statement of intent on actions, i.e. pro-social behavior. Like in the real-world examples, we are examining public statements of intent given by an institution. The communicative act, in form of a public statement of good intent, is induced with the aim to motivate the agents to act pro-socially. The cognitive dissonance theory (Festinger, 1957) explains why individuals may change their behavior in response. According to this theory, individuals seek consistency in their behavior and experience emotional disutility when they act inconsistently with previous actions or statements. Avoiding the disharmony motivates the commitment towards the action stated. Hence, even in anonymous interactions, in which reputational concerns are non-existent, statements of good conduct can create commitment.

Complementary research in experimental economics shows that a substantial number of individuals avoid lying, even if they have to forgo a material gain by doing so (Gneezy, 2005; Kartik et al., 2007; Kartik, 2009; Hurkens and Kartik, 2009; Lundquist et al., 2009; Fischbacher and Föllmi-Heusi, 2013). And people who make a promise about future cooperative behavior are likely to keep their word.<sup>1</sup>

How strong a commitment is varies with the form of the statement. Research in social psychology stresses the importance of decision autonomy (Linder et al., 1967). Individuals who decide freely to express an intention about future behavior, feel committed to this action (Kiesler, 1971; Schlesinger, 2011). Elicited or pre-formulated promises, by contrast, have none or only a small effect on the consequent behavior (Charness and Dufwenberg, 2010; Belot et al., 2010). Public promises have a stronger commitment effect than private pledges (Joule and Beauvois, 1998); and being engaged in a pledge, i.e. the statement-making is in some form effortful, increases the binding function.

The behavioral impact of statements of good intent has been analyzed with a strong focus on one-on-one interactions. Promises are used in these settings to signal trustworthiness so that a consequent interaction can take place. We, on the contrary, study a promise addressed to a group and inves-

---

<sup>1</sup>Two theoretical explanations for promise-keeping are prominent in the literature. First, an aversion to lying exists, either because the person has a preference for keeping their word per se (Ellingsen and Johannesson, 2004; Vanberg, 2008; Sánchez-Pagés and Vorsatz, 2009; Erat and Gneezy, 2012; López-Pérez and Spiegelman, 2013) or because the promisor does not want to act against the social norm of not breaking a promise (Binmore, 2006; Bicchieri and Lev-On, 2007). Second, the effect may be indirect: the statement raises the expectations of others, the person making the statement anticipates the raised expectations and is motivated not to disappoint the expectations of others (Charness and Dufwenberg, 2006; Battigalli et al., 2013).

tigate whether the public declaration to act in a socially beneficial way can help to promote cooperation. This links our work to a vast amount of studies on the effect of communication in social dilemma situations (Ledyard, 1995; Chaudhuri, 2011, for reviews). In these situations, communication reduces strategic uncertainty and subjects treat one another’s messages as a serious attempt to coordinate (Ostrom et al., 1992). Consequently, promises and commitments to cooperate are prominent elements in the discussions which arise in open, pre-play communication; and the possibility to communicate is particularly beneficial when all group members promise to cooperate (Orbell et al., 1988, 1990; Bicchieri, 2002, e.g.).

In reality, howsoever, open discussions between all relevant actors may be time-consuming or just not be feasible (Messick et al., 1983), not to mention the challenge of finding consensus in an open discourse. To restrict communication in these settings may help to ease coordination. Koukoulis et al. (2012) provide in this context first evidence that the cooperation-enhancing-effect of communication also exists when not all group members can communicate with each other. In their study, the option to send a free message was only given to one randomly selected person. Contributions to the public good still increased substantially.

Also Dannenberg (2015) examined how group contributions to a public good change when one player can announce his or her intended contribution before group members have to make their allocation decision. In this study, the communicator was either chosen randomly or volunteered. In both cases communicators announced high contribution levels, but rarely realized them. The other group members anticipated this and rarely got affected by the announcement. Only in the volunteering treatment, a small increase in contributions could be found. The communication in Dannenberg’s study only contained the numerical announcement of the intended contribution level and was not in free text format. Compared to Koukoulis finding, this may explain the weaker effect of the message.

Our study is similar to both studies in the regard that also we offer only one message as a coordination device; a design feature which may ease efficiency and substantially lowers communication costs in reality. However, in contrast to the two previous studies, the message in our design is not endogenously chosen by one of the group members, but given and predetermined by an external institution. Moreover, our design elicits a form of self-commitment from each player. For previously listed reasons, we expect the effect of the statement thus to be stronger.

Our study adds to the literature by offering a highly controlled message as a coordination device in an abstract social dilemma situation. We identify

the channels through which a statement of intent may affect the contribution behavior. Existing promise or oath studies identified the effect of a statement so far in between-subject designs. Consequently, it is not clear whether an increase in pro-social behavior, observed in these studies, is due to a commitment effect or due to the selection effect that only socially oriented people make a statement about intended social behavior. In our within-subject design we can address this issue and can distinguish between a selection and a commitment effect. Additionally, we control in a between-subject treatment variation whether the commitment effect varies when the statement is voluntary or compulsory.

## 3 Experiment

### 3.1 Design

For this experiment, we employed a standard linear public good game with a voluntary contribution mechanism (Marwell and Ames, 1979). The experiment consisted of two stages and subjects played in each stage ten rounds of the standard public good game in partner matching. The first stage served as a baseline to measure heterogeneity and allows us to carry out a within-subject analysis.

At the beginning of each round, each subject was endowed with 20 Experimental Currency Units, which they could consume either privately or contribute to a public good. The payoff function was the following:

$$p = 20 - z_i + 0.4 * \sum_{j=1}^n z_j \quad (1)$$

$z_i$  denotes  $i$ 's contribution level, where  $0 \leq z_i \leq 20$  and  $0.4 * \sum_{j=1}^n z_j$  presents the income from the project. Contributions to the public good increased the collective output, but the marginal per capita return of a contribution was less than one. A profit-maximizing individual may thus have been tempted to keep the entire endowment and free-ride on the contributions of the other players. Free-riding, however, diminished the total welfare for all group members.

After the first ten rounds, experimental groups were re-matched<sup>2</sup> and the

---

<sup>2</sup>The rematching of players was done in a way that no group members interacted twice. A perfect stranger matching was guaranteed and this was common knowledge.

treatment variations were implemented. Participants were randomly allocated between three groups: Control, Voluntary and Compulsory. Table 1 summarizes the design.

Table 1: Experimental Design

	STAGE 1	STAGE 2	N
<b>CONTROL</b>	Standard PGG	Standard PGG	64
<b>VOLUNTARY</b>	Standard PGG	<i>Voluntary Statement</i> Standard PGG	64
<b>COMPULSORY</b>	Standard PGG	<i>Compulsory Statement</i> Standard PGG	64

In the two treatment groups, a statement was offered to communicate intended future contribution behavior. The promise was directed to the other group members and stated that the player will make significant contributions to the project, that is at least 75%<sup>3</sup> of the endowment. In the first treatment group (Voluntary), participants simultaneously decided whether or not to make this statement. In the second treatment group (Compulsory), players had to make the statement in order to proceed in the experiment. In both cases, it was made explicit that making the statement had no consequence on the set of possible future choices and did not limit the decisions later in the experiment. All participants who decided to pledge had to type in the following, ‘*I promise to contribute each round at least 15 ECUs to the project*’. According to research in social psychology (Linder et al., 1967; Kiesler, 1971), commitment is stronger when individuals engage in the act of promise making. For this reason, we decided that subjects had to key in the statement.<sup>4</sup> Before the first contribution decision in Stage 2, all players in

<sup>3</sup>We have chosen 75% instead of 100%, the Pareto optimal contribution level, to still allow for some latitude in the contribution choices.

<sup>4</sup>Subjects who decided not to pledge, had to key in a neutral text: ‘*I am a voluntary participant in this experiment, no coercion or interference has taken place.*’ This text



the two treatment groups learned which players in their group (also) made the statement. In the second stage, Statement-Makers were consequently labelled.<sup>5</sup>

To control how beliefs influenced the contribution choice, we asked subjects after their contribution decision to indicate their expectations about the contributions of the other players (first order beliefs) and the guess of others expectations with respect to their own contributions (second order beliefs).<sup>6</sup> After all subjects stated their beliefs, feedback was given about each group member's individual contribution.<sup>7</sup>

At the end of each stage, three rounds were randomly selected to determine the payments; one round for the contribution decision and two rounds for the accuracy of the beliefs. By this mechanism we minimized wealth effects and prevented hedging within a stage.

### 3.2 Experimental Procedure

The experiment was conducted at the Queensland University of Technology, we used the experimental software CORAL (Schaffner, 2013) and the on-line recruitment system ORSEE (Greiner, 2015). The experiment lasted for about an-hour-and-a-half. Sessions were equally distributed over the three treatment groups. Before participants could start the experiment, comprehension questions needed to be answered correctly. The average payment was 25.3 Australian dollars (app. 18 USD). The data comprises observations of 192 individuals and 3840 decisions in total.

---

was already introduced in the baseline stage and was of similar length as the contribution statement.

<sup>5</sup>Respective screenshots and instructions can be found in the appendix.

<sup>6</sup>The belief payment function was incentive compatible and based on the quadratic scoring rule. However, to make it easier for the students to understand, we described the possible outcomes verbally in the instructions.

<sup>7</sup>Croson and Marks (2001) find that feedback about single players' contribution compared to information about the total contribution does not change the average contributions. Also Fehr and Gächter (2000) do not find a difference in contributions for feedback on an average level or feedback that displays the entire contribution vector.

### 3.3 Behavioral Predictions and Identification Strategies

Under the assumption of purely self-serving and money maximizing behavior, contributions are expected to be 0 in all groups and stages. Statements, if they are made, are considered as meaningless by the participants and do not affect their choices of contribution. This also applies to the repeated setting of the game. A rational decision maker will always break the promise in the last round and try to free-ride on the contributions of the others. Applying backward induction, the statement is consequently non-credible in all rounds of the game.

Empirical evidence, however, shows human behavior deviates fundamentally from these predictions. Contributions in public good games are on average between 40 – 60% of the endowment in the first round of an experiment and deteriorate over the repetitions of the game (Ledyard, 1995; Fehr and Gächter, 2000; Chaudhuri, 2011). Moreover, non-institutionalized, multilateral communication enhances the contribution levels significantly (Sally, 1995; Bochet et al., 2006; Balliet, 2009). Koukoulis et al. (2012)’s study provides first evidence that also one-dimensional communication may be sufficient to motivate higher contribution levels.

Following these insights, public and institutionalized statements of intent should increase contributions to the public good in this experiment. This reasoning is also supported by the promise literature. Individuals are reluctant to lie; either because the person has a preference for keeping their word (Ellingsen and Johannesson, 2004; Vanberg, 2008) or because the promisor does not want to go against the social norm of not breaking a promise (Binmore, 2006; Bicchieri and Lev-On, 2007). Other authors argue that the effect is more indirect: the statement raises the expectations of others, the promisor anticipates this and is motivated not to disappoint the expectations of the interaction partners (Charness and Dufwenberg, 2006). Both theories, however, suggest that making a statement increases contributions to the public good in our experiment.

In the case of the voluntary statements, this increase can come from two potential explanations: First, a *selection effect* indicating that people who are intrinsically motivated to contribute more, tend to make the statement; second, a *commitment effect*, based on the theories above the pledge motivates an increase in contributions. Given our experimental design, the treatment group 1 (Voluntary) allows us to disentangle the commitment effect from the selection effect by comparing the Stage 1 to the Stage 2 behavior, in which the voluntary statement was offered. We can see whether participants who

voluntarily choose to make the statement in Stage 2, have on average, higher contributions in Stage 1 (*selection effect*). And we can determine if participants who voluntarily made the statement, increase their contributions in Stage 2 compared to Stage 1 (*commitment effect*).

**Hypothesis 1 (selection effect):** *Participants who voluntarily choose to make the statement in Stage 2, have on average higher contributions in Stage 1.*

**Hypothesis 2 (commitment effect):** *Participants who make the statement subsequently increase their contributions to the public good in Stage 2 compared to their contributions in Stage 1.*

If making a statement triggers commitment, due to lying aversion or cognitive dissonance, and the interaction partners are aware of this effect, they may also adapt in their contribution behavior. To be more specific, if subjects observe other group members are making the statement, they expect higher group contributions in the future. Following conditional cooperation (Fischbacher et al., 2001), this belief change motivates the subject to also contribute more to the public good. We refer to this as *coordination effect*, and see it as the second source for an increase in pro-social behavior after a pledge.

In the Compulsory treatment, this coordination effect should be stronger since all subjects (are forced to) make the statement. Particularly over time, we expect this coordination effect to unfold.

**Hypothesis 3 (coordination effect):** *In treatment group 2, where statements are compulsory, a positive increase in contributions can be sustained for longer than in treatment group 1, where statements are voluntary and the number of Statement-Makers is thus determined endogenously.*

In regard to the self-driven commitment, it is nevertheless to expect that a compulsory statement triggers a weaker commitment than a voluntary statement (Kiesler, 1971; Schlesinger, 2011). Decision autonomy is an integral prerequisite for commitment. The empirical support for this hypothesis comes from studies analyzing the effect of a promise in bilateral interactions (Linder et al., 1967; Charness and Dufwenberg, 2010; Belot et al., 2010). No evidence howsoever exists for groups or aggregate contributions. Hence, an open empirical question is to what extent the commitment effect of a voluntary statement is stronger than the commitment effect of a compulsory statement. By controlling for the amount of statements made in one group, we observe a coordination component varying in strength. Particularly interesting are hereby observations, where all group members voluntarily decide

to make the statement. The comparison with the Compulsory group can provide insights into how the obligation to make the statement impairs the commitment effect. We expect that contributions in the Voluntary groups with four Statement-Makers are higher than in the Compulsory treatment groups. In such groups, a strong commitment would be paired with a high coordination component.

## 4 Results

To investigate how public statements of good intent affect the contribution behavior, we first focus on voluntary statements of intent (treatment group 1). By controlling for a selection effect, we can identify whether an additional commitment effect existed. In a second step, this commitment effect is then compared with the behavioral change that was taking place when the statement was compulsory (treatment group 2). Lastly, we will analyze the dynamic development of the effects.

### 4.1 Voluntary Statement

Whilst the rational money maximizing assumption predicts that statements are irrelevant and therefore not worth being made, we find subjects made voluntary statements, and also got affected by them.

Overall, 48% of the subjects made the statement before entering Stage 2. We call these subjects Statement-Makers.<sup>8</sup> Subjects who had the option to make the statement, but decided against it, we refer to as Non-Takers. After the statement decision, these two groups contributed in substantially different ways. The Statement-Makers contributed significantly more than Non-Takers (Pairwise Wilcoxon rank sum test:<sup>9</sup>  $Z = -3.682$ ,  $p < 0.001$ ). Table 2 provides an overview of the average contribution levels for each group and stage.

**Result 1:** *Contributions were higher when subjects made (voluntarily) the public statement of good intent.*

[Table 2 here]

---

<sup>8</sup>A detailed analysis of the individual characteristics of the Statement-Makers is offered in the appendix.

<sup>9</sup>Mean contributions per person, averaged over all stage rounds, are chosen as level of analysis, if not indicated differently.

Table 2: Average contributions for each group and stage

Contributions	Stage 1 ( <i>Baseline</i> )	Stage 2	Number participants
CONTROL	8.70 (4.67)	7.65 (4.52)	64
VOLUNTARY	8.57 (3.35)	10.08 (5.77)	64
Non-Takers	7.56 (2.83)	7.52 (4.98)	33
Statement-Makers	9.65 (3.56)	12.82 (5.33)	31
COMPULSORY	10.61* (6.02)	13.13 (5.61)	64
Number of participants			192

Standard deviations are in parantheses.

*\*Note: Stage 1 contributions in the Compulsory group were on average higher than in the Voluntary and Control group. This was due to a strong variation across the experimental groups. In the regression analysis, we control for these initial differences and can show that they do not impair the findings presented in the following.*

Comparing the contributions in Stage 1 suggests that subjects who later made the statement were before the pledge already different to those who did not. Also in Stage 1, contributions from Statement-Makers were significantly higher than the contributions from Non-Takers ( $Z = -2.506$ ,  $p = 0.0122$ ).

**Result 2:** *Subjects who decided to voluntarily make the statement, acted already before the pledge in a more socially oriented way. We call this selection effect.*

At an aggregate level, mean contributions in Stage 2 were significantly higher in the Voluntary group than in the Control group ( $Z = -2.231$ ,  $p = 0.0026$ ), as well as compared to the contributions in Stage 1 ( $Z = -2.210$ ,  $p = 0.0271$ ). This difference was driven by the Statement-Makers, who significantly increased their contributions after the pledge (Voluntary Statement-Makers - Stage 1 vs. Stage 2:  $Z = -2.732$ ,  $p = 0.0063$ ).<sup>10</sup>

<sup>10</sup>The results also hold when the analysis is based only on the first round of Stage 2, i.e. when subjects did not yet receive any information about the other group members' contribution behavior and their reaction on the statement (Round 11 - Control vs. Voluntary group:  $Z = -2.150$ ,  $p = 0.0315$ ; Voluntary Statement-Makers - Round 1 vs. Round 11:

By looking at the differences in contribution levels between Stage 1 and Stage 2, we can ensure that the difference in contributions between Statement-Makers and Non-Takers was not purely due to selection. Figure 1 displays the corresponding average differences in contribution levels. It can be seen that beside the selection, also a change in behavior took place. Statement-Makers increased their contributions after the pledge, while Non-Takers retained their contribution levels (Difference in contribution levels between Stage 1 and Stage 2 of Voluntary Statement-Makers vs. Non-Takers:  $Z = -3.070$ ,  $p = 0.0021$ ).

**Result 3:** *Statement-Makers increased their contributions after they made the statement of intent. We call this commitment effect.*

We assume that the statement not only created intrinsic commitment but also changed the beliefs a player held about the behavior of the other players. Table 3 displays the expectations from other players, players' own second order beliefs and the respective contributions. To measure the pure effect of the statement in regard to others' expectations, we examine the beliefs and contributions only from the first round of the second stage. This was the first interaction after a new group was formed. Players were at this point not yet able to predict the other players' behavior based on the history.

[Table 3 here]

Expectations in the Voluntary treatment group were significantly higher than the expectations subjects formed in the Control group. Interestingly, this was not only the case for the expectations towards the Statement-Makers ( $Z = -4.114$ ,  $p < 0.0001$ ), but also for the expectations towards the Non-Takers ( $Z = -2.043$ ,  $p = 0.0410$ ). In general, the introduction of the statement made subjects more optimistic of what contributions to expect in the second stage. But within this optimistic belief, expectations towards Statement-Makers were significantly higher than towards Non-Takers ( $Z = -1.797$ ,  $p = 0.0723$ ).<sup>11</sup>

The second order beliefs, in contrast, were only higher for the Statement-Makers. They anticipated correctly that other players expected significantly higher contributions from them ( $Z = -5.572$ ,  $p < 0.001$ ). This is in line with previous research on expectation based guilt aversion (Charness and Dufwenberg, 2006, 2010). Statement-Makers knew that they raised the expectations of others (Second order beliefs) and felt guilty if they did not meet these expectations. However, it needs to be mentioned that with our design

---

$Z = -3.380$ ,  $p = 0.0007$ )

<sup>11</sup>When the beliefs from all Stage 2 rounds are taken into account, the divergence of these expectations becomes stronger.

Table 3: Beliefs in Stage 2

	Control	Voluntary	
		Non-Taker	Statement-Maker
<b>Expectation of others</b>	8.04 (3.82)	8.61 (4.06)	12.73 (4.17)
<b>Second order belief</b>	8.12 (4.41)	7.58 (4.48)	13.33 (4.79)
<i>Corr.coef.</i>	<i>0.92</i>	<i>0.93</i>	<i>0.93</i>
<b>Real contribution</b>	7.65 (4.53)	7.52 (4.98)	12.82 (5.33)

*Note: This table presents the average beliefs and contributions in the first round of Stage 2, with standard deviations in parantheses.*

*The first row displays the average expectation towards a player in the Control group (column 1), a player who decided against the statement (column 2), and a player who made the statement (column 3). The second row lists the average second order belief, i.e. the guess of a player about the average expectations of the other group members about his or her contribution behavior. The last row presents the average contributions realized by the respective players. The row in-between displays Pearson correlation coefficients, which measure the association between second order beliefs and realized contributions.*

we are not able to rule out that Statement-Makers reported higher second order beliefs simply to be consistent with the higher contributions made.

As last analysis in this section, we perform a multivariate analysis to control for group interdependencies and analyze how the take up of the statement by others affected the contribution behavior. Table 4 presents the results of Ordinary least square regression models estimating the average contributions for each stage and treatment group. Contributions of the Control group in Stage 1 serve as baseline.

[Table 4 here]

Model 1 estimates the contributions for each treatment group and stage, considering the average contributions of other players an individual experienced in Stage 1. In the Control group, contributions were lower in the second than in the first stage. Contrary, in the Voluntary treatment group contributions increased in Stage 2.

Model 2 differentiates between subjects, who made the statement and subjects who decided against it. Statement-Makers contributed significantly

Table 4: Average Contributions in Control and Voluntary

	Model 1	Model 2	Model 3
Voluntary	-0.127 (1.128)		
Voluntary $\times$ No		-1.138 (1.158)	-1.138 (1.165)
Voluntary $\times$ State		2.088** (0.816)	2.088** (0.821)
Stage 2	-1.050 (1.278)	-1.050 (1.283)	-4.542*** (1.392)
Stage 2 $\times$ Voluntary	2.562 (1.814)		
Stage 2 $\times$ Voluntary $\times$ No		1.005 (1.803)	2.640 (1.986)
Stage 2 $\times$ Voluntary $\times$ State		4.221** (1.930)	3.922* (2.238)
Stage 2 $\times$ Voluntary $\times$ No State $\times$ N_State			-1.056 (0.970)
Stage 2 $\times$ Voluntary $\times$ State $\times$ N_State			0.178 (0.833)
OthersContrib_S1			0.134*** (0.0339)
Constant	8.698*** (0.988)	8.698*** (0.992)	8.698*** (0.998)
Observations	256	256	256
Adjusted $R^2$	0.023	0.108	0.143

*Note: This table presents the results of OLS regressions, estimating average contribution levels for each stage in the Control and Voluntary treatment group. The variable N\_State lists how many of the other players in one's matching group made a statement. OthersContrib\_S1 is a control variable for the average contributions of other players an individual experienced in Stage 1. Standard errors, clustered on the matching group level, are in parentheses. \*\*\*, \*\*, \* indicates significance at the 1%, 5% and 10% levels, respectively.*



more in Stage 1 ( $p = 0.0129$ ) than Non-Takers, this is what we described before as the *selection effect*. These Statement-Makers additionally increased their contributions significantly in Stage 2 ( $p = 0.032$ ).

Model 3 considers as additional controls the subjects' experience of others' contributions in Stage 1 and the number of other players who (also) made a statement in the subject's matching group ( $N\_State$ ). When considering these additional influences, the commitment effect ( $Stage\_2 \times Voluntary \times N\_State$ ) was slightly weaker ( $p = 0.085$ ). The experience individuals made in Stage 1 had a strong positive effect on the contribution behavior in Stage 2. The presence of (other) Statement-Makers promoted contributions when the subject was also a Statement-Maker ( $Stage\_2 \times Voluntary \times State \times N\_State$ ). When the subject, on the contrary, decided against the statement, the presence of Statement-Makers weakened the average contributions. One possible interpretation for this is a diffusion of responsibility; when players decided not to take the statement and then observed that others pledged to contribute a substantial amount to the public good, then the Non-Takers felt no longer responsible for the maintenance of the public good.

To summarize, we observed that more socially oriented subjects selected themselves into the statement and second, that Statement-Makers increased their contributions after the pledge. This increase was anticipated by the other group member and the Statement-Makers were aware of the higher expectations. We also observed that the change in contribution behavior was not only driven by the pure commitment towards the statement, but was also influenced by the statement choices of the other players in the subject's matching group. The increase in contributions was consequently stronger when also other group members made the statement. We conclude that not only the self-driven commitment due to e.g. lying aversion determined the increase in contributions after a pledge but also the facilitated coordination with the other group members played a role. To investigate this influence we contrast in the following the results from the Voluntary treatment with the Compulsory setting.

## 4.2 Voluntary versus Compulsory Statements

In this section, we compare the behavioural changes of the Voluntary treatment group with a setting in which all group members had to make the statement. From the social psychological perspective, the pure, self-driven commitment associated with a compulsory statement should be negligible since the pledge is not self-chosen (Kiesler, 1971; Linder et al., 1967). In comparison with the Voluntary setting, the commitment effect in the Compulsory

setting should be smaller. Our data, however, contradicts this prediction.

When we compare the change in contributions on an individual level<sup>12</sup>, we find contributions increased more when statements were made voluntarily than when subjects were forced to make the statement. However, this difference is not statistically different to the average increase we observed for compulsory Statement-Makers ( $Z = 0.099$ ,  $p = 0.9210$ ). Figure 1 displays the average change in contributions for each round between Stage 1 and Stage 2, for Statement-Makers and Non-Takers in each group. The similarity remains when we exclude potential learning effects and base our analysis only on the first round after the statement (Increase from Round 10 to Round 11: Voluntary Statement-Makers - 4.903 ECUs vs. Compulsory Statement-Makers - 4.734 ECUs:  $Z = -0.120$   $p = 0.904$ ). This finding stands in contrast to the results found in bilateral interactions, where elicited promises were significantly less effective than self-chosen commitments (Charness and Dufwenberg, 2010; Belot et al., 2010).

**Result 4:** *Contributions increased significantly after subjects made the compulsory statement. This increase was not statistically different to the increase in contributions of voluntary Statement-Makers.*

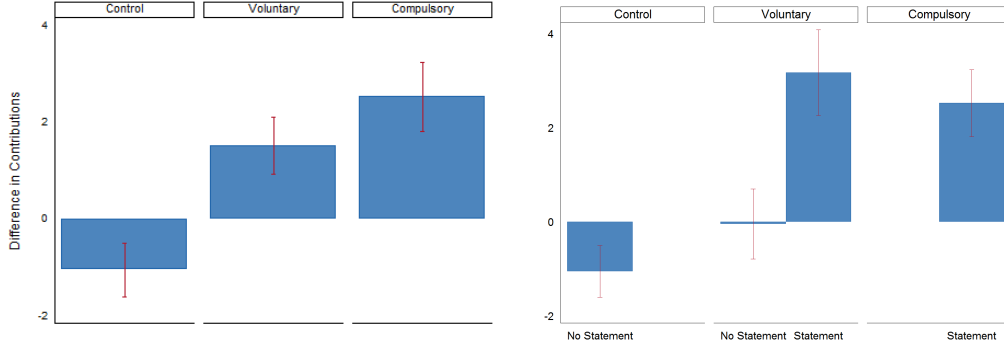
Furthermore, the difference in the average increases between the Voluntary and the Compulsory treatment group was close to statistical significance when we pooled the data of voluntary Statement-Makers and Non-Takers ( $Z = -1.466$ ,  $p = 0.1428$ ). Hence, we confidently reject the hypothesis that compulsory commitments in a public good setting were less effective than voluntary commitments.

[Figure 1 here]

---

<sup>12</sup>Since baseline contributions in the Compulsory group were higher than in the Voluntary and Control group, we base the comparison between treatments on the difference between Stage 1 and Stage 2 contributions.

Figure 1: Commitment: Difference between Stage 1 and Stage 2 contributions



Note: This graph displays the average difference in contributions between Stage 1 and Stage 2, with error bars: As in other studies, contributions in the Control group (left bar) were lower when the game was played for the second time. At the aggregate level, contributions increased more in the Compulsory treatment group than in the Voluntary treatment group. We call this *coordination effect*. In the right panel, contributions of Statement-Makers and Non-Takers in the Voluntary treatment group (middle bar) are displayed separately. Contributions of voluntary Statement-Makers significantly increased, while the contributions of Non-Takers remained constant. Also for the compulsory Statement-Makers (right bar) contributions increased significantly. This increase we call *commitment effect*.

Also on side of the beliefs, expectations towards Statement-Makers were higher in the Compulsory group than in the Voluntary group (First order beliefs: Voluntary Statement-Makers 12.73 ECUS (sd: 5.31), Compulsory Statement-Makers 13.75 ECUS (sd: 5.83):  $Z = -4.489$ ,  $p < 0.001$ )<sup>13</sup> and the respective Statement-Makers anticipated these higher expectations correctly (Second order beliefs: Voluntary Statement-Makers 13.33 ECUS (sd: 6.09), Compulsory Statement-Makers 14.05 ECUS (sd: 5.85):  $Z = -2.21$ ,  $p = 0.03$ ). We attribute this finding also to the *coordination effect*. The assurance that everyone signed up for the same ‘rule’, shifted expectations and thus promoted higher contributions. Ultimately, this increase balanced the weaker intrinsically motivated commitment following the imposition of the pledge.

To investigate a possible difference between the voluntary and compulsory statements in the self-driven commitment effect, we look at the data of two experimental groups, in which all four group members voluntarily chose to make a statement. Contributions, as well as the differences in contributions

<sup>13</sup>When only the beliefs of the first round in Stage 2 are chosen as unit of analysis, this difference is even stronger.

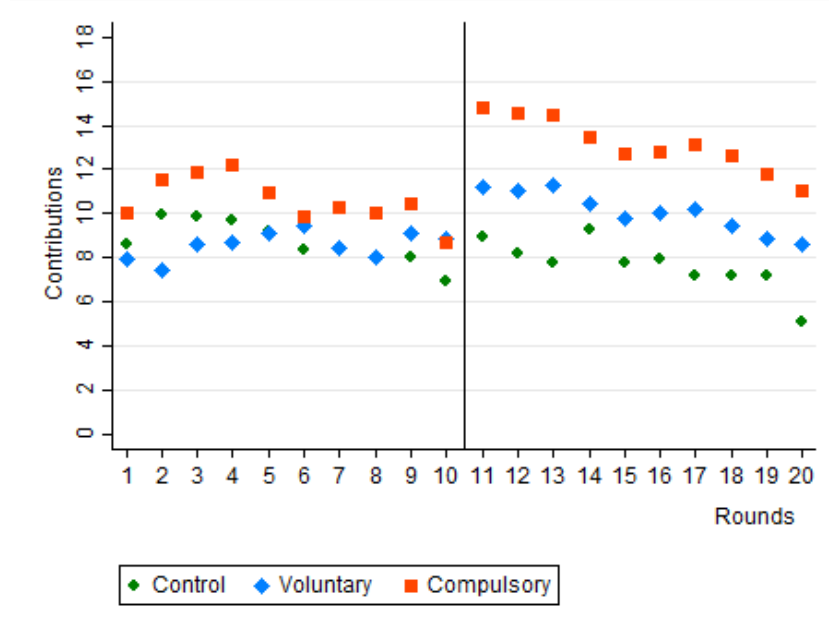
between Stage 1 and 2 were significantly higher in these two groups than in all other groups (average increase in the Voluntary group 3.59 ECUs (sd: 2.33) compared to 2.52 ECUs (sd: 6.53) in the Compulsory group, a graph with the respective contributions can also be found in the appendix). While the high contribution levels can be explained by the selection effect, the stronger increase points to the positive combination of coordination and self-driven commitment. This observation, howsoever, is based on two experimental groups and can thus be taken only as indicative finding.

**Result 5:** *Compulsory statements of intent were effective in motivating higher contributions to the common. The assurance that everyone had to make the statement eased coordination. We call this coordination effect.*

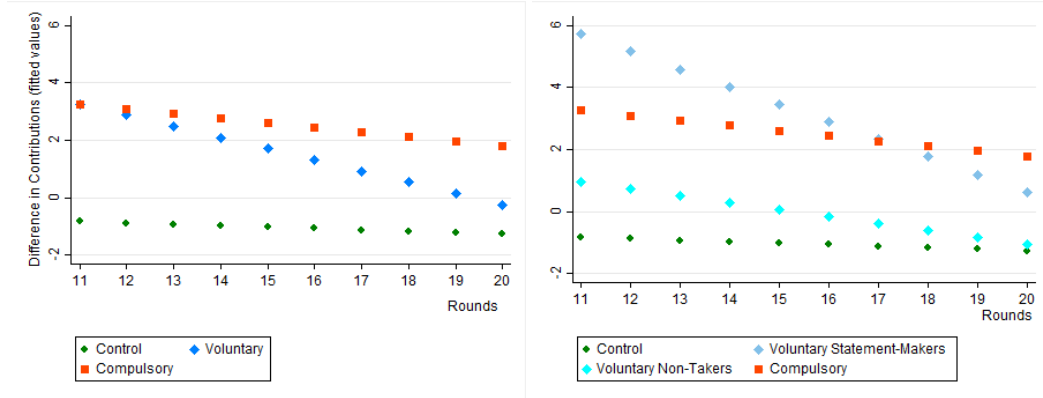
Further support for the effectiveness of compulsory statements and for the existence of a coordination effect is found in the development of contributions over time. Figure 2 shows first the dynamic development of contributions in each treatment group, then the fitted values of the differences in contributions between Stage 1 and 2. The effect of a voluntary statement was directly after the pledge as strong as the effect of the compulsory statement, but contributions deteriorated faster when the statement was voluntary ( $p = 0.0753$ ). The pattern was caused by participants who made the statement voluntarily and reduced their initial high contributions over time. This can be explained with conditional cooperation. After the first round, subjects learned how the other group members contributed and adjusted their contributions accordingly; over time the good intentions of contributing 15 ECUs or more vanished. In the Compulsory group, contrarily, everyone committed to the statement and contributions stayed higher for longer.

[Figure 2 here]

Figure 2: Dynamic Development



(a) Note: This graph shows the average contributions of the Control, Voluntary and Compulsory treatment group in all rounds of Stage 1 and 2.



(b) Note: This graph displays the dynamic development of contributions between Stage 1 and 2. We used fitted values, based on the average change of individual contributions for each round: The left column shows that the increase in contributions is deteriorating faster in the Voluntary group than in the Compulsory group, another aspect of the coordination effect. In the right column, it can be seen that the decline of average contributions in the Voluntary treatment group is driven by the voluntary Statement-Makers.

[Table 5 here]

Table 5: Dynamic Change in Contributions, per Treatment in Stage 2

	Model 4	Model 5	Model 6
Round	-0.0485 (0.110)	-0.0485 (0.110)	-0.0485 (0.110)
Voluntary	7.838** (2.991)		
Voluntary $\times$ No		3.700 (3.132)	3.516 (3.101)
Voluntary $\times$ State		8.542** (3.322)	8.775** (3.320)
Compulsory	5.366** (2.475)	5.366** (2.477)	6.445** (2.498)
Voluntary $\times$ Round	-0.340* (0.187)		
Voluntary $\times$ No $\times$ Round		-0.174 (0.190)	-0.174 (0.190)
Voluntary $\times$ State $\times$ Round		-0.517** (0.231)	-0.517** (0.231)
Compulsory $\times$ Round	-0.116 (0.156)	-0.116 (0.156)	-0.116 (0.156)
AvgContrib_S1			-0.189*** (0.0231)
Constant	-0.298 (1.993)	-0.298 (1.994)	4.625** (1.980)
Observations	1920	1920	1920
Adjusted $R^2$	0.039	0.052	0.131

*Note: This table presents the results of OLS regressions on the change in contribution levels for each round of Stage 1 and Stage 2. Standard errors, clustered on the matching group level, are in parentheses.*

*\*\*, \*\*, \* indicates significance at the 1%, 5% and 10% levels, respectively.*

The OLS regressions in Table 5 support these findings statistically. All estimations are based on the change of individual contributions for each round between Stage 1 and Stage 2. The left side of Figure 2 is based on Model 4. The coefficients *Voluntary* and *Compulsory*, which measure the average difference in contributions between Stage 1 and 2, are not statistically different from each other ( $F = 0.86$ ,  $p = 0.359$ ), but statistically different to the change in contributions in the Control group ( $p < 0.05$ ). In this model, the development over time is not statistically different between the treatment groups ( $F = 0.144$ ,  $p = 0.2362$ ).

Model 5 distinguishes between voluntary Statement-Makers and subjects who decided not to make a statement. The lower right side of Figure 2 is based on this estimation. The coefficients *VoluntaryState* and *Compulsory*, measuring the mean change in contributions between Stage 1 and 2 for voluntary and compulsory Statement-Makers, are statistically different ( $F = 6.61$ ,  $p = 0.0133$ ). This result indicates that voluntary statements possibly induced stronger, self-driven commitment at the beginning of Stage 2. The consequent increase in contributions, however, vanished faster for the voluntary than for the compulsory Statement-Makers ( $H_0: \text{Voluntary} \times \text{State} \times \text{Round} = \text{Compulsory} \times \text{Round}: F = 3.02$ ,  $p = 0.0885$ ). The results remain robust when individuals' average contribution levels in Stage 1 are taken into account (*AvgContrib\_S1*). Groups, who made the compulsory statement were more steered against the deterioration of contributions over time. This means that compulsory statements of intent may not only help to reach higher level of contributions on an aggregate level in public good situations, they may also help to sustain this increase in contributions for longer.

**Result 6:** *The positive effect of the statement lasted longer in the Compulsory treatment group than in the Voluntary treatment group.*

Finally, subjects' compliance with the statement shall be discussed. In the first round of Stage 2 both groups of Statement-Makers, voluntary and compulsory, were with their contributions close to stated level (Voluntary Statement-Makers: 14.387 ECUS (sd: 5.018) and Compulsory Statement-Makers: 14.797 ECUs (sd: 5.265)). The compliance rates of 87% for the Voluntary Statement-Makers and 83% for the compulsory Statement-Makers were not statistically different ( $Z = 0.534$ ,  $p = 0.593$ ). The compliance rates, however, deteriorated over the course of Stage 2 and reached on average 68% for the voluntary and 73% for the compulsory Statement-Makers. Overall, compulsory Statement-Makers were thus weakly significantly more often compliant than voluntary Statement-Makers ( $Z = -1.671$ ,  $p = 0.095$ ). The reason for this lies in the faster depletion of the motivation of voluntary

Statement-Makers to fulfill the statement when they were playing with Non-Takers. The two groups, in which all group members voluntarily chose to make the statement, met in all rounds, except in the last two, the required contribution level (Average contribution: 15.462 ECUs (sd: 5.619)). This finding points again to the strong influence of the coordination effect.

**Result 7:** *Compliance with the statement of intent was more frequent in the Compulsory treatment group than in the Voluntary treatment group.*

## 5 Conclusions

Our results suggest that public statements of good conduct, used in oaths and professional codes of conduct, can help to promote pro-social behavior in public good settings. This stands in contrast to the rational choice assumption that such statements would be cheap talk and behavioral changes must be merely attributed to a *selection effect*.

With a within-subject design, we controlled for this endogenous selection and demonstrate that the public statements motivate an additional increase in pro-social behavior, the so-called *commitment effect*. Once a statement is made, it has a positive effect on the level of contributions. Furthermore, we found that this commitment effect has two components, the behavioral change due to the self-driven commitment and the effect of improved coordination. On basis of this *coordination effect*, contributions and compliance with the public statement were in our study higher when all group members made a statement of good conduct. In contrast to bilateral interactions, this finding suggests that compulsory statements of good conduct are not less effective in promoting public good contributions, but may even work better under certain conditions, for example, when the general willingness to make a voluntary pledge is low.

To what extent are these results specific to the artificial situation in an experimental laboratory and what conclusions can be drawn for the real world? Contributions to the public good were contrived in our design and misconduct was clearly defined. In real world settings, this is clearly not the case. Sometimes deviations from good conduct can be easily identified, but in most cases the distinction between violating the code of conduct and acting in a simply less socially acceptable way is blurry. Hence, the abstraction in our laboratory experiment naturally impairs the external validity of our findings. The purity of incentives and clarity of the decision setting allowed us to identify three effects that potentially motivate behavioral changes when public statements of good conduct are in use. But there is a great potential and



need for future research to probe the workings of our selection, commitment and coordination effect in the field.

In general, our research suggests that public statements of good conduct, even though they are non-binding, can help to promote pro-social behavior in social dilemma situations. In comparison to legally binding rules, such public commitments may be politically a easier and cheaper solution to implement, which is still effective in promoting contributions to public goods.

## Acknowledgements

The authors would like to thank Uri Gneezy, Nicholas Jaquemot, Richard Jefferson, Sarah Necker, Pedro Rey-Biel, Benno Torgler, Caroline van Bers and Israel Waichman for their valuable comments on earlier versions of this paper. Ann-Kathrin Koessler thanks the QUT Business school for the funding of the experiments and the Alexander von Humboldt-Foundation for the support.

## References

- Balliet, D. (2009). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution*, 54(1):39–57.
- Barger, S. D. (2002). The Marlowe-Crowne affair: Short forms, psychometric structure, and social desirability. *Journal of Personality Assessment*, 79(2):286–305.
- Basinger, K. S. and Gibbs, J. C. (1987). Validation of the social moral reflection objective measure - A short form. *Psychological Reports*, 61(1):139–146.
- Battigalli, P., Charness, G., and Dufwenberg, M. (2013). Deception: The role of guilt. *Journal of Economic Behavior & Organization*, 93:227–232.
- Begley, C. G. and Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, 483(7391):531–533.
- Belot, M., Bhaskar, V., and van de Ven, J. (2010). Promises and cooperation: Evidence from a TV game show. *Journal of Economic Behavior & Organization*, 73(3):396–405.
- Bicchieri, C. (2002). Covenants without swords: Group identity, norms, and communication in social dilemmas. *Rationality and Society*, 14(2):192–228.
- Bicchieri, C. and Lev-On, A. (2007). Computer-mediated communication and cooperation in social dilemmas: An experimental analysis. *Politics, Philosophy & Economics*, 6(2):139–168.
- Binmore, K. (2006). Why do people cooperate? *Politics, Philosophy & Economics*, 5(1):81–96.

- Bochet, O., Page, T., and Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60(1):11–26.
- Charness, G. and Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6):1579–1601.
- Charness, G. and Dufwenberg, M. (2010). Bare promises: An experiment. *Economics Letters*, 107(2):281–283.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14(1):47–83.
- Croson, R. and Marks, M. (2001). The effect of recommended contributions in the voluntary provision of public goods. *Economic Inquiry*, 39(2):238.
- Dannenberg, A. (2015). Leading by example versus leading by words in voluntary contribution experiments. *Social Choice and Welfare*, 44(1):71–85.
- de Bruin, B. and Dolfsma, W. (2013). Oaths and codes in economics and business - Introducing the Special Issue. *Review of Social Economy*, 71(2):135–139.
- Ellingsen, T. and Johannesson, M. (2004). Promises, threats and fairness. *The Economic Journal*, 114(495):397–420.
- Erat, S. and Gneezy, U. (2012). White lies. *Management Science*, 58(4):723–733.
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PloS one*, 4(5).
- Fehr, E. and Gächter, S. (2000). Cooperation and punishment in public goods experiments. *The American Economic Review*, 90(4):980–994.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.
- Fischbacher, U. and Föllmi-Heusi, F. (2013). Lies in disguise: an experimental study on cheating. *Journal of the European Economic Association*, 11(3):525–547.

- Fischbacher, U., Gächter, S., and Fehr, E. (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics Letters*, 71(3):397–404.
- Fischer, D. and Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne social desirability scale. *Educational and Psychological Measurement*, 53(2):417–424.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4):25–42.
- Gibbs, J. C., Basinger, K. S., Fuller, D., Fuller, R. L., et al. (2013). *Moral maturity: Measuring the development of sociomoral reflection*. Routledge.
- Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, 95(1):384–394.
- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125.
- Hurkens, S. and Kartik, N. (2009). Would i lie to you? on social preferences and lying aversion. *Experimental Economics*, 12(2):180–192.
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, 2(8):e124.
- Jacquemet, N., Joule, R.-V., Luchini, S., and Shogren, J. F. (2013). Preference elicitation under oath. *Journal of Environmental Economics and Management*, 65(1):110–132.
- Jacquemet, N., Luchini, S., Rosaz, J., and Shogren, J. F. (2018). Truth telling under oath. *Management Science*, 0.
- Jacquemet, N., Luchini, S., Shogren, J. F., and Zylbersztejn, A. (2017). Coordination with communication under oath. *Experimental Economics*, pages 1–23.
- Joule, R. and Beauvois, J. (1998). *The free will compliance: How to get people to freely do what they have to do*.
- Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395.

- Kartik, N., Ottaviani, M., and Squintani, F. (2007). Credulity, lies, and costly talk. *Journal of Economic Theory*, 134(1):93–116.
- Kiesler, C. A. (1971). *The psychology of commitment: Experiments linking behavior to belief*. Academic Press New York.
- Koukoulis, A., Levati, M. V., and Weisser, J. (2012). Leading by words: A voluntary contribution experiment with one-way communication. *Journal of Economic Behavior & Organization*, 81(2):379–390.
- Ledyard, J. O. (1995). *Public Goods: A Survey of Experimental Research*. Princeton: Princeton University Press.
- Linder, D. E., Cooper, J., and Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, 6(3):245.
- List, J. A., Bailey, C. D., Euzent, P. J., and Martin, T. L. (2001). Academic economists behaving badly? A survey on three areas of unethical behavior. *Economic Inquiry*, 39(1):162–170.
- López-Pérez, R. and Spiegelman, E. (2013). Why do people tell the truth? experimental evidence for pure lie aversion. *Experimental Economics*, 16(3):233–247.
- Lundquist, T., Ellingsen, T., Gribbe, E., and Johannesson, M. (2009). The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1):81–92.
- Martinson, B. C., Anderson, M. S., and De Vries, R. (2005). Scientists behaving badly. *Nature*, 435(7043):737–738.
- Marwell, G. and Ames, R. E. (1979). Experiments on the provision of public goods. i. resources, interest, group size, and the free-rider problem. *American Journal of Sociology*, pages 1335–1360.
- Mazar, N., Amir, O., and Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6):633–644.
- Messick, D. M., Wilke, H., Brewer, M. B., Kramer, R. M., Zemke, P. E., and Lui, L. (1983). Individual adaptations and structural change as solutions to social dilemmas. *Journal of Personality and Social Psychology*, 44(2):294–309.

- Necker, S. (2014). Scientific misbehavior in economics. *Research Policy*, 43(10):1747–1759.
- Orbell, J., Dawes, R., and Van de Kragt, A. (1990). The limits of multilateral promising. *Ethics*, 100(3):616–627.
- Orbell, J. M., Van de Kragt, A. J., and Dawes, R. M. (1988). Explaining discussion-induced cooperation. *Journal of Personality and Social Psychology*, 54(5):811.
- Ostrom, E., Walker, J., and Gardner, R. (1992). Covenants with and without a sword: Self-governance is possible. *American Political Science Review*, 86(02):404–417.
- Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, 7(1):58–92.
- Sánchez-Pagés, S. and Vorsatz, M. (2009). Enjoy the silence: an experiment on truth-telling. *Experimental Economics*, 12(2):220–241.
- Schaffner, M. (2013). Programming for experimental economics: Introducing CORAL - a lightweight framework for experimental economic experiments. QuBE Working Papers 016, QUT Business School.
- Schlesinger, H. J. (2011). *Promises, oaths, and vows: On the psychology of promising*. Taylor & Francis.
- Shu, L. L., Mazar, N., Gino, F., Ariely, D., and Bazerman, M. H. (2012). Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proceedings of the National Academy of Sciences*, 109(38):15197–15200.
- Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica*, 76(6):1467–1480.

## 6 Appendices

### 6.1 Subject pool

#### 6.1.1 Demographics

Table 6: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Econ (%)	0.56	0.50	0	1	192
Undergrad (%)	0.85	0.23	0	1	192
Female (%)	0.46	0.50	0	1	192
Age	22.60	6.35	17	63	192
Income	171.53	16.05	50	1000	144
Finsat (%)	0.32	0.468	0	1	175
lessRelig	4.43	1.55	1	6	175
CRT	1.21	1.12	0	3	192
SRM	5.83	1.51	4	11	191
SDS	4.43	2.35	0	11	190

Table 7: Demographics over Treatment groups

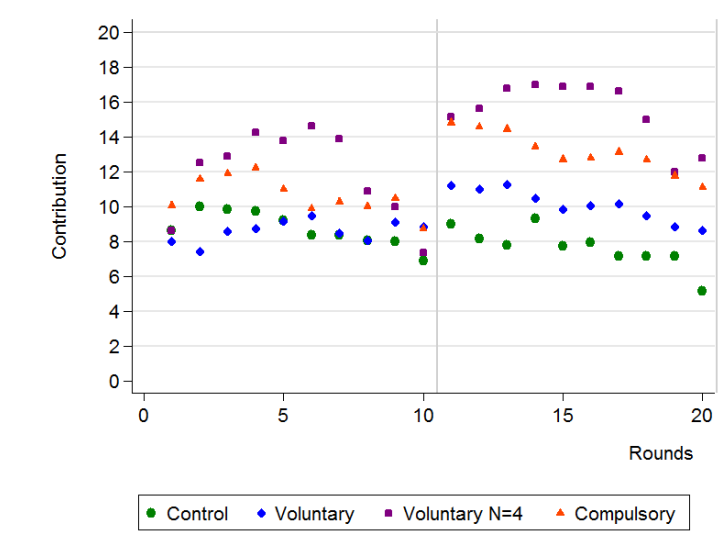
	Treatment groups			Total
	Control	Voluntary	Compulsory	
Econ-Fin-Bus (%)	57.8	54.7	56.3	56.3
Undergrad (%)	92.2	84.4	78.1	84.9
Female (%)	45.3	45.3	48.4	46.4
Age	22.94	23.64	22.23	22.93
Income	174.44	169.79	170.59	170.53
Finsat	0.29	0.34	0.33	0.32
lessRelig	4.40	4.63	4.27	4.43
CRT	0.95	1.34	1.33	1.21
SRM	5.67	5.86	5.97	5.83
SDS	4.27	4.46	4.57	4.43
N	64	63	63	190*

\* Two subjects did not answer the psychological questionnaires.

## 6.2 Contributions of groups with $n=N$ Statement-Makers

[Figure 3 here]

Figure 3: Contributions in Stage 1 and 2, Statements



*Note: This figure displays average contributions for each round in Stage 1 and Stage 2. The development of the two groups in the Voluntary treatment, in which all group-members decided to make the statement voluntarily, is plotted separately.*

## 6.3 Demographic Characteristics of Statement-Makers

In the following, we shed light on the characteristics of subjects who voluntarily made the statement in the Voluntary treatment group.

For this purpose, we use demographic data and information from psychological measures we elicited in a post-experimental questionnaire. We asked subjects for their sex, age, degree, course and religiosity.<sup>14</sup> We also asked five questions from the Socio-moral Reflection Measure Questionnaire (Basinger and Gibbs, 1987; Gibbs et al., 2013), which address socio-moral

<sup>14</sup>To have an indication how religious participants were, we asked "Apart from weddings, funerals and christenings, how often do you attend religious services these days?" The variable was coded with "More than once a week" (1), "Once a week" (2), "Once a month" (3), "Once a year" (4), "Less often than once a year" (5), "Never" (6). The observed average of 4.43 suggests that participants on average went to church once per year or less; apart from weddings, funerals and christenings.



values like truth-telling. The questions we selected asked for participant's attitude towards promises and lying.<sup>15</sup>

We also elicited a short version of the Crowne and Marlow Social Desirability Scale (SDS). This scale is often used in Psychology and Clinical Research to measure the need for social approval.<sup>16</sup> A person with a high SDS score is more likely to perform certain behavior due to a desire to be socially accepted or approved.

Ultimately, as an estimator for strategic reasoning, we integrated the cognitive reflection test (CRT) (Frederick, 2005). This test is designed to assess an individual's ability to suppress an intuitive response, which is incorrect, and engage in further reflections that lead to the correct response. Answers were incentive compatible so that participants were paid 1 AUD for each correct answer. The CRT measure ranges from 0 to 3, indicating a person with a high CRT score is able to resist intuitively compelling responses.

Table 8 shows the likelihood that a participant takes the statement voluntarily in Stage 2 based on the demographic characteristics. Model 1 takes into account a participant's study major and degree, gender and age as well as the experience with economic experiments. The variables Female, Econ and Postgraduate are dummy variables which take the value one when the participant was respectively female, studied Economics or enrolled in a post-graduate course. Model 2 predicts the likelihood of making the voluntary statement based on the extent to which a participant was satisfied with his or her financial situation and the degree of religiosity. Model 3 uses the psychological measurements we elicited in the experiment as explanatory variables. Model 4 combines all previous three models and Model 5 controls additionally for the experience a participant has made in the previous stage (average contribution of the other group members in Stage 1) and the own contribution behavior in Stage 1. The results show that only gender has a significant and robust impact on the decision to voluntarily make a statement about intended social behavior. When a participant was female, she was 40% less likely to make the voluntary statement ( $p = 0.016$  in Model 5).

---

<sup>15</sup>We asked the following questions: 1) *How important is it for people to keep promises, if they can, to friends?* 2) *How important is it for people to keep promises, if they can, even to someone they hardly know?* 3) *How important is it for parents to keep promises, if they can, to their children?* 4) *How important is it for people to tell the truth?*

The variable is coded in reverse order: very important (1), important (2), not important (3). Thus a high score in SRM indicates that the person stated that he or she perceives promise-keeping as less important.

<sup>16</sup>The original version includes 50 items, we used a shorter version from Fischer and Fick (2003) which is proofed to be also valid and internally consistent (Barger, 2002).

Table 8: Demographic characteristics of Statement-Makers

VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5
Econ	-0.0611 (0.138)			-0.170 (0.171)	-0.0588 (0.189)
Postgrad	0.187 (0.209)			0.383 (0.273)	0.535* (0.310)
Experiment	0.153 (0.146)			0.219 (0.181)	0.259 (0.209)
Female	-0.276** (0.130)			-0.418*** (0.156)	-0.395** (0.163)
Age	-0.00695 (0.0107)			-0.00876 (0.0162)	-0.00993 (0.0182)
Finsat		-0.0876 (0.143)		-0.153 (0.169)	-0.150 (0.185)
lessRelig		0.0537 (0.0446)		0.0573 (0.0555)	0.0632 (0.0579)
CRT			-0.0241 (0.0560)	-0.0571 (0.0732)	-0.0497 (0.0756)
SRM			0.0607 (0.0388)	0.0393 (0.0470)	0.0555 (0.0503)
SDS			0.0540* (0.0310)	0.0407 (0.0375)	0.00774 (0.0416)
Contrib_S1					0.0662** (0.0305)
OthersContrib_S1					-0.00856 (0.0140)
Observations	64	56	63	55	55

*Note: This table presents the marginal effects (calculated at the means of all variables) from a probit regression on the likelihood that a participant takes voluntarily the statement. Standard errors are in parentheses. \*\*\*, \*\*, \* indicates significance at the 1%, 5% and 10% levels, respectively.*