# The effect of leadership on free-riding: results from a public-good experiment

Daniel M. Parsons
Department of Economics
Monash Business School
Clayton, VIC 3800, Australia
parsonsdaniel34@gmail.com

Nick Feltovich*
Department of Economics
Monash Business School
Clayton, VIC 3800, Australia
nicholas.feltovich@monash.edu

Philip J. Grossman
Department of Economics
Monash Business School
Clayton, VIC 3800, Australia
philip.grossman@monash.edu
and
Rasmuson Chair, University of Alaska Anchorage

**Abstract:** We examine the impact of two types of communication: (i) encouragement of honesty and (ii) encouragement of lying that benefits the group. Subjects choose contributions to a public good, with a portion of the contribution framed as determined by a self-reported die roll. While honesty is typically viewed as desirable, in our setting it is more equivocal, since it results in a sub-optimal group payoff. We find that when leaders encourage their followers to lie in a cooperative way, followers increase these "die roll" contributions. There is also a positive spillover into additional discretionary contributions to the public good. By contrast, the way leaders are chosen and their observed contribution history have little effect.

**Keywords:** leader; cheap talk; lying; honesty; group culture
**JEL codes:** D91, D23, C72, C92, H41, M14

* Corresponding author. This work was supported by the Australian Research Council (DP130101695). We thank three anonymous referees for helpful suggestions and comments.

## 1. Introduction

Successful cooperation is often difficult to achieve. Cooperation rates in one-shot prisoners' dilemmas and contribution rates in one-shot linear public good games are typically only around one-half. When either of these games is repeated with feedback between rounds, rates of cooperation or contribution typically start there and trend downward (Ledyard, 1995; Chaudhuri, 2011). In coordination games, maintaining coordination on the group-earnings-maximising outcome proves difficult in all but the smallest groups (Feltovich and Grossman, 2015).

Typically, a group leader is either a first mover in sequential-move games or a group member whose cheap-talk message is relayed to the rest of the group at the start of each period; the leader is usually randomly chosen, with no special information or power relative to other group members. Evidence from previous experiments suggests that both leading by example (see for example, Potters et al., 2005, 2007; Guth et al., 2007: Komai and Grossman, 2009; Komai et al., 2011), and cheap-talk messages (see, for example Levy, et al., 2011; Koukoumelis et al., 2012; Houser et al., 2014; Feltovich and Grossman, 2015) improve group cooperation and coordination.[1] Communication is usually found to increase coordination and cooperative behaviour, though the efficacy of communication depends partly on the form it takes (Ledyard, 1995, Sally, 1995, Bochet et al., 2006, Balliet, 2010, Chaudhuri, 2011).[2]

In this paper, we consider the impact of a leader who sends a cheap-talk message. Our design sets a high bar for the leader to have an impact on follower behaviour; our leaders send their messages only once, prior to the 26$^{th}$ round of the die rolls. This offers an extreme test of the effectiveness of leaders and their messages. We examine the impact of two specific types of communication: (i) encouragement of honesty and (ii) encouragement of lying that benefits the group. Importantly, we introduce a tension between these two kinds of communication, both of which promote pro-social behaviour. Our setting is a multi-round social dilemma with fixed groups of three, and without feedback between rounds. In each round, each subject can contribute between zero and four tokens to a linear public good. Half of these four tokens are framed as being allocated according to a self-reported die roll, with zero, one, and two each equally likely. In the paper (though not in the instructions to subjects), we call these "die-roll contributions"; the other two tokens are framed as completely discretionary ("top-up contributions"). As always, an individual subject earns the most by free-riding, including under-stating the die roll if necessary. An honest subject will contribute on average one token of the two allocated by the die roll, while a cooperative subject

[1] Koukoumelis et al. (2012), Angelova et al. (2019) with the unique design with permanent (i.e., members of the group through all interactions) and temporary (i.e., members who rotate out of the group after each period) leaders and followers provide counter evidence. They find that leaders who lead by example in a public good game reduce contributions.

[2] Besides leadership, other mechanisms have been found to increase cooperation, including endogenous group formation (Ehrhart and Keser, 1999; Cinyabuguma et al., 2005; Ahn et al., 2009), playing in continuous rather than discrete time (Friedman and Oprea, 2012; Oprea et al., 2014), and increasing group sizes gradually (Weber, 2006).

will contribute all four tokens, including both of those allocated by the die roll – and therefore over-stating the die roll if necessary.

So, while honesty is typically viewed – in experiments and elsewhere – as desirable behaviour, its desirability in our setting is more equivocal. Following a norm of honesty is better for the group than free-riding, but still results in a sub-optimal payoff for the group. The group would be best off if everyone always contributed the full four tokens irrespective of the die rolls, following the norm of cooperation at the expense of the norm of honesty. The conflict between norms is an important aspect of social behaviour that has not received much attention in the experimental economics literature.

Previous public-good experiments (Ledyard 1995) suggest that in our setting, both free-riding and cooperation are likely to be observed. Deception experiments suggest that some subjects will be reluctant to lie (Gneezy, 2005; Gibson et al., 2013), and that mutually beneficial lying will be especially prevalent in our setting when the resulting gains are shared with others (Wiltermuth, 2011; Erat and Gneezy, 2012).[3] Hence, in the absence of leaders, we would expect a wide range of behaviour, with some subjects contributing almost nothing, some contributing roughly half of the endowment, and others contributing more. This (conjectured) heterogeneity implies a role for leaders to guide the behaviour of their group.

In our experiment, fixed groups play two sets of 25 rounds.[4] The first set of 25 rounds is without feedback and with no leader. In our baseline treatment, the second set of 25 rounds is simply a continuation of the first set; we call this our ***No Leader*** (N) condition. In our other two treatments, one of the existing group members is chosen as the group leader, with the other group members remaining followers. The leader's role is quite limited, with only two facets. First, she can choose one of two pre-programmed messages (or if she prefers, no message), to send before the second set of 25 rounds starts. One message is a pro-honesty message; the other is a pro-cooperation message. Second, information about her performance in the first set of 25 rounds is displayed to all group members, also prior to the beginning of the second set of 25 rounds. The difference between these two treatments is in how leaders are chosen. In our ***Random Leader*** (R) condition, each group member is equally likely to become the leader, while in our ***Group-oriented Leader*** (G) condition, the leader is chosen based on contributions over the first 25 rounds. In all treatments, there is no further intervention in the second set of 25 rounds, and no feedback until the last round has ended.

[3] Studies that have examined lying in public good games have focused on players lying to each other regarding the amount they contributed to the public good (Irlenbusch and Ter Meer, 2013) or the experimenter lying to subjects and exaggerating the average contribution to the public good (Hoffman et al., 2013). To our knowledge there is no study in which subjects can lie to the experimenter to increase or decrease their contribution to the public good.

[4] Subjects are informed that there were two stages to the experiment and that instructions for each stage would be provided at the start of each stage.

Our design enables us to examine whether a leader can improve group outcomes by encouraging pro-social behaviour. If free-riding is initially rampant in the group, even a recommendation for honesty can benefit the group by raising die-roll contributions. A recommendation for cooperation can benefit the group even more if followed by erstwhile free-riders, and it can even benefit a group of primarily honest subjects if it leads them to over-state their die rolls. This trichotomy, where subjects can be selfish (always contributing zero from the die roll), honest (contributing one on average) or cooperative (always contributing two), is one of the unique features of our design.

Our design also enables us to examine if a leader who engages in non-selfish lying can, by lying and encouraging her followers to do likewise, influence her followers to overcome their incentive to free ride. If the honesty norm prevails when there is no leader, then we can detect both positive and negative treatment effects, in contrast to many experiments in which the baseline treatment implies no cooperation at all, so treatment effects can go in only one direction. Furthermore, and most importantly, we ask if such behaviour can help create a social norm with spillover effects encouraging subjects to more generally contribute to a group good. Can the encouragement to act cooperatively in one action (here, die-roll contributions) encourage more cooperative behaviour in other actions (top-up contributions)? We test whether a leader can, and will, bring about a socially beneficial outcome by encouraging privately disadvantageous, but socially advantageous, lies. A leader, by encouraging others to cooperatively lie, may give her followers the freedom to also lie, and help the group achieve the Pareto efficient action.

We can think of two clear counterparts to our experimental setting in real life. The more obvious of the two comprises settings where honesty, while socially beneficial, limits the success of a group undertaking. Many examples exist; a recent one is the series of scandals involving automobile makers' cheating on emissions tests, with dishonesty practiced by many individual employees within each of these companies. We hesitate to draw policy conclusions from our results to these kinds of settings for two reasons. First, cooperation amongst group members is harmful overall, and hence we do not wish to encourage it. Second, a casual look at corporate scandals over the last couple of decades suggests that there is little difficulty in overcoming any norm of honesty in those cases.

The second counterpart uses a less straightforward analogy. Some organisations put into place structures that limit the worst excesses by individuals, but (presumably unintendedly) also restrict positive behaviour that benefits the group. In our experiment, the honesty norm fills this role. We know of no example where honesty performs a similar role in the outside world, but other types of culture do. One example is micro-managing: restrictive travel policies, time-consuming performance reporting systems, and (in academic institutions) overly bureaucratic human ethics committees. These efforts are probably successful at reducing wasteful spending, screening out poor performers, or minimising lawsuits from unethical research, but at the expense of productive activity that would be beneficial for the institution. At a disaggregated level, micro-management is likely to deter the worst free-riders in an organisation, but also to weigh down those who are intrinsically motivated to champion the group. In this sense, it functions similarly to the honesty norm in our setting: raising cooperation by those who are disposed to low cooperation, but lowering cooperation by those disposed to high cooperation. Obviously, there are lots of differences between these organisational settings and our laboratory environment, including differences in moral and ethical status between honesty and micro-management. Thus, we consider our experiment as an analogy for comparable settings in the outside world, not a replica. As with all analogies, the correspondence to the real settings is not perfect, but we believe it is useful nonetheless.

Our results are fairly strong, and in some ways surprising. Behaviour in the first 25 rounds is illustrative of an unsuccessful group endeavour. Many subjects appear roughly honest in reporting their die rolls – though about a fifth seem to under-report, and a much smaller fraction over-reports – so that "die-roll contributions" average less than one (per subject-round). Average top-up contributions also average less than one, but with greater heterogeneity, as many subjects average close to zero or close to two. The honesty norm encourages those who had been reporting their die rolls honestly and on average contributing one as a top-up contribution to report a group beneficial die roll and to increase their top-up contributions. However, among the most cooperative subjects, the honesty norm has the effect of reducing their contributions.

Observed behaviour in the second set of 25 rounds suggests that neither the manner in which the leader is chosen, nor the leader's observed history, has an effect on the group's subsequent behaviour. The message, however, is very important. Groups receiving no message – either because there is no leader or because the leader chose not to send one – and groups receiving the pro-honesty message achieve similar results to those from the first 25 rounds. Groups receiving the pro-cooperation message, by contrast, substantially raise their contribution levels in the second 25 rounds. The effect is most pronounced in the die-roll contributions, which increase by about 24 percentage points.

We also observe a spillover effect. Leaders who encourage cooperation (at the expense of honesty) help to create a social norm encouraging subjects to more generally contribute to the public good. The increase in top-up contributions (about 7 percentage points after a pro-cooperation message, and no overall change otherwise), while smaller than the increase in die-roll contributions, is significant. Moreover, it is notable that there is any increase at all: subjects are not "robbing Peter to pay Paul" by contributing more based on the die rolls and compensating themselves by contributing less elsewhere.

Examination of the disaggregated data suggests that the impact of the leader's cooperative message is mainly due to inducing honest followers to ditch the honesty norm in favour of higher contributions –

that is, to follow the suggestion in the message. For the non-contributing followers, the leaders' message has little impact. The differences we observe in behaviour are also reflected in money earnings: groups receiving a pro-cooperation message earn more on average than the other groups.

## 2. Literature review

Numerous studies offer evidence of lying behaviour in the lab and elsewhere. Self-serving lying is common but not universal; there is a substantial subset of individuals who do not lie even at considerable cost to themselves.[5] Not all lying is selfish lying: subjects may lie to avoid appearing greedy (Utikal and Fischbacher, 2013); subjects lie more when the benefits of doing so go to a charity (Lewis et al. 2012) and if gains from the lie are shared with others (Wiltermuth, 2011; Erat and Gneezy, 2012). Erat and Gneezy (2012) observe frequent lying when the lie is an altruistic "white" lie (i.e., lies that harm the liar but help others).

In our game, group members observe their leaders' previous play and can determine whether the leader acted selfishly or to benefit the group. Alempaki et al. (2016) find that paired players lie less if they were treated "fairly" in an earlier dictator game. d'Adda et al. (2017) report that self-interested group leaders influence the ethical conduct of their followers; followers of known self-interested, dishonest leaders exhibit even higher levels of dishonesty. Barr and Michailidou (2017) and Weisel and Shalvi (2015) observe more lying when subjects are trying to coordinate with their accomplice to maximize earnings. Diekmann et al. (2015) find that subjects lie more if they have seen a distribution of previous die roll reports, observing that others have likely lied appears to have freed them to lie more.

In our design, participants receive a message from their group leader as well as observing their leaders' tabulation of recorded rolls from the first 25 rounds. Thus, they have some indication of their leaders' tendency to record rolls favourable to herself or to the group. Past behaviour has been shown to influence current behaviour in a number of games.[6] Kahneman et al. (1986) and Eckel and Grossman (1996) find players will reward (punish) players who had been generous (selfish) in a prior dictator game. Duffy and Feltovich (2002) find that observing their partners' behaviour from the prior round increases coordination and cooperation in Prisoners' Dilemma, Stag Hunt, and Chicken games. Duffy and Feltovich (2006) report similar results when both prior behaviour and a cheap talk message is revealed, but only when behaviour and message are in agreement. For a public good game, Page et al. (2005) show that a player's past behaviour in the game increases the likelihood the player will be selected as a partner as well as

[5] See, e.g., Gneezy (2005), Gibson et al. (2013), Dai et al. (2016) and Abeler et al. (2014).
[6] The theory of reciprocity (Falk and Fischbacher 2006) argues that good or bad past behaviour will be rewarded or punished. See, for example, Bolton et al. (2005), Seinen and Schram (2006), and Charness et al. (2011).

increasing contribution rates in the current game. Similarly, Croson (2007) finds that a high contribution rate by a leader encourages stated higher contribution rates by potential followers using the strategy method.

## 3. Experimental design

Our underlying setting is a three-player, linear public-good game, under the voluntary-contributions mechanism (VCM). In each round, each subject is allocated four tokens, worth 0.08 Australian dollars (AUD) each, per round to distribute between an individual and a group account. At the time of the experiment, one AUD was worth roughly 0.75 US dollars. Tokens allocated to the individual account earn the subject their value. Tokens allocated to the group account are doubled and the proceeds are shared equally among the three group members. Individual earnings per round in dollars are given by

$$\pi_{it} = 0.08\left[(4 - x_{it}) + \frac{2}{3}\sum_{j=1}^{3} x_{jt}\right] \qquad (1)$$

where $x_{it}$ and $\pi_{it}$ are the contribution and earning of the i-th player in the t-th round, and the j's are the members of i's group (including i). As usual, the Nash equilibrium under standard preferences entails no contributions, while group earnings are maximised if all tokens are contributed.

Clearly, individualistic behaviour in this setting is simply free-riding (contributing nothing), while the group-optimal choice is to contribute all four tokens. Inducing a norm with an intermediate level of contributions is more complicated. One possibility would be simply to claim such a norm exists, perhaps by inserting appropriate wording into the experiment instructions (e.g., "imagine that in the past, members of your group have contributed X in each round"). This technique has been used with some success to induce focal agreements in bargaining experiments (e.g., Gächter and Riedl 2005, Karagözoğlu and Riedl 2015). However, in those cases it specifies one Nash equilibrium out of many, and thus has an element of stability. In our setting, unless the norm consisted of zero contributions (and was thus identical to individualistic behaviour), it would not be a Nash equilibrium, and we know of no evidence to suggest that attempts to induce a disequilibrium norm would be successful.

Instead of trying to induce a new norm, we therefore exploit an existing norm: that of honesty. As noted in the literature review, there is substantial experimental evidence showing that honesty is common: individuals often tell the truth even at a monetary cost. We take advantage of the honesty norm via the following procedure. While two of the four tokens in a subject's endowment are framed as completely at the subject's disposal, we frame the other two as being allocated according to the result of a die roll. In each round, each subject rolls a standard die, enters the result into the computer, and the contribution is determined as shown in Table 1. We call the contributions arising from this roll the subject's ***die-roll contributions***, and the contributions from the remaining two tokens her ***top-up contributions***. The choice of the top-up contribution is made immediately after reporting of the die roll (and hence determination of the die-roll contribution).

| Reported die roll | Number of tokens allocated to… | |
|---|---|---|
| | Group account | Individual account |
| 1, 2 | 0 | 2 |
| 3, 4 | 1 | 1 |
| 5, 6 | 2 | 0 |

Table 1: Die-roll contribution, depending on reported die roll

The information is Table 1 is made known to subjects in the experiment instructions. Also, die rolls are not observed by anyone other than the subject (neither the experimenter nor other subjects). So, there is both the opportunity and a monetary incentive to misreport the die roll. An individualistic subject can simply ignore the actual die roll and report a 1 or 2 in order to contribute zero. An honest subject, by contrast, will report the actual die roll, which entails an expected die-roll contribution of one token. This is more than an individualistic subject will contribute, but is sub-optimal from the standpoint of the group, which fares best if everyone's die-roll contribution is two (and of course the group-optimal top-up contribution is also two).

Thus, we induce a group norm in our experiment by taking advantage of the norm of honesty that prevails in the outside world. This norm has effects that depend on the prevalence of individualistic versus cooperative subjects in the group. If most group members are individualistic but conform to the honesty norm, the effect is positive: such people will contribute one token each to the group (die-roll contribution of one on average, top-up contribution of zero), versus zero if the honesty norm did not exist. However, if most group members are cooperative but conform to the honesty norm, its effect is negative: they will contribute three tokens each (die-roll contribution of one, top-up contribution of two) instead of all four. In this latter case, the honesty norm arguably limits the group's success.

To avoid confusion, we wish to emphasise that we view the parallel between honesty and inefficiency as limited to our laboratory environment, and applicable to real life only as an analogy, not literally. In our study, we take advantage of an existing norm of honesty by constructing a game in which complete honesty benefits the group more than some forms of lying (lying in order to free ride) but less than other forms of lying (lying in order to contribute the maximum). In that, the honesty norm functions similarly to other institutions that prevent some kinds of bad behaviour, but at the cost of limiting some kinds of good behaviour. As noted in the introduction, micro-management is a real-world analogue, probably familiar to anyone working in a large organisation (whether public, for-profit, or private non-profit). The organisation's central management seeks to reduce certain kinds of bad behaviour (e.g., free-riding that benefits the individual at the expense of the organisation), but in ways that also restrict good behaviour (e.g., intrinsically motivated cooperation) that would benefit the organisation. In that, it functions similarly to the way the honesty norm functions in our experimental setting. However, they differ in many other ways, and our exploitation of the honesty norm in this experiment should not be interpreted as advocating for an increase in dishonesty outside the lab.

Besides allowing us to distinguish between individualistic subjects bound by the honesty norm and cooperative subjects bound by the honesty norm, our use of top-up contributions (in addition to die-roll contributions) also limits experimenter demand effects. If all tokens are distributed by a die roll, subjects are restricted to either lying or being truthful. This might be perceived as a demand on the part of the experimenter that the subjects lie, resulting in an overestimation of the subjects' willingness to be deceitful.

### 3.1 Treatments

Our baseline ("No leader" or N) treatment has the VCM game played for 50 rounds, with fixed groups and no feedback between rounds. The large number of rounds allows us to distinguish with high likelihood between subjects who lie about the die rolls and those who merely get a run of high or low numbers. The lack of feedback between rounds prevents contamination of subjects' choices from observing what others have done.

Our "Random leader" (R) treatment differs by introducing a randomly chosen leader after the first 25 rounds. We refer to the remaining group members as "followers", though this term was not used in the instructions to subjects. The leader serves two primary purposes. First, she chooses a message to send to her followers. The message options are:

(1) No message ('n')

(2) Pro-honesty message ('h'): "Everyone should record their true roll."

(3) Pro-cooperation message ('c'): "Everyone should record a 5 or 6. If we all do this, we will all earn the most money."[7]

Second, the leader's total numbers of 1 or 2 rolls, 3 or 4 rolls, and 5 or 6 rolls from rounds 1-25 are revealed to the followers (see Appendix C for a sample screenshot); note that this is equivalent to providing information about the number of times the leader's die-roll contribution was 0, 1 or 2. After the leader's

[7] The message options were listed in the instructions, which were read aloud in an attempt to make them common knowledge. Followers largely treated a "No message" as neither a message to lie nor as an honesty message (see Figure 6).

message and history are broadcast to the followers, subjects play rounds 26-50 in the same groups as before. As previously noted, our design sets a high bar for the leader to have an impact on follower behaviour. This offers an extreme test of the effectiveness of leaders and "be honest" cheap-talk messages.[8]

Our "Group-oriented leader" (G) treatment is identical to the R treatment, except that the leader was the group member who recorded the most rolls of 5 or 6 (and thus the most die-roll contributions of 2) over the first 25 rounds. The selected leader is likely a liar, but a group-oriented liar. Comparison between the G treatment and the R treatment allows us to test whether the impact of a leader depends on the process by which the leader is chosen.

### 3.2 Research questions

Standard theory makes the same predictions in all of our treatments. The unique stage-game Nash equilibrium involves zero contributions by all group members to the public good. Since the number of rounds is finite and publicly announced, the unique Nash equilibrium for the entire session is simply this stage-game equilibrium played in every round. The existence of a leader, the process by which the leader is selected, the leader's message, and the content of the leader's history shown to followers are all irrelevant to the standard theory.

With this in mind, we do not state formal hypotheses, but instead list the research questions that underlie our experiment.

*1. Are contributions higher when a leader is selected based on past cooperation rather than randomly?*

In our G treatment, leaders are chosen based on their die-roll contributions, and are thus immediately verified as the group member most willing to contribute to the group (even perhaps at the expense of honesty). Past cooperation does not commit the leader to contribute in future rounds, but does serve as a costly signal that she is likely to do so. By contrast, leaders are randomly chosen in our R treatment. While followers can ascertain the leader's cooperativeness from the information they receive about past contributions, the leader has a two-thirds chance of not being the high contributor in the group, and even determining this may be cognitively taxing.

The implication for followers' behaviour is unclear. Having a leader chosen based on past cooperation could encourage followers to follow her lead by raising contributions (Duffy and Feltovich, 2002 and 2006,

[8] We acknowledge that having leaders send messages, and having their past choices observed by followers, in the R and G treatments introduces confounds, so that we will not be able to disentangle the effect of leaders per se from the effect of the information followers get. Our main treatment effects (differences among the R, G and N treatments) should therefore be interpreted as the effects of introducing a leader combined with this information.

provide evidence that observation of previous actions are better than cheap-talk messages at raising cooperation in social dilemmas). However, it could have no effect, or it could even lead followers to reduce their contributions (out of spite, or because any encouragement of lying – even pro-cooperation lying – destroys the norm of honesty and induces previously honest free-riders to begin under-reporting their die rolls).

*2. Are contributions higher when a leader sends a pro-cooperation message than other (or no) messages?*

As noted earlier, numerous studies have shown that pro-cooperation messages increase cooperation by followers. Here, the message serves a secondary purpose: not only encouraging cooperation, but also signalling that it is acceptable to break the norm of honesty to do so. We may expect a stronger-than-usual effect for this reason.

*3. Are contributions higher when the leader has visibly contributed more in the past?*

As with (1) above, a leader who contributed a high amount in absolute terms has sent a costly signal of commitment to the group's well-being which could lead followers to increase their contributions, but it is also possible that contributions decrease or remain roughly the same.

## 4. Experimental procedures

All sessions were conducted at Monash University's MonLEE lab. Subjects were recruited using ORSEE (Greiner, 2015) and the experiment was programmed in z-Tree (Fischbacher, 2007). Three sessions of each treatment were conducted with 171 subjects in total (see Table 2).

| Treatment | Sessions | Groups | Subjects |
|---|---|---|---|
| No leader (N) | 3 | 21 | 63 |
| Random leader (R) | 3 | 18 | 54 |
| Group-oriented leader (G) | 3 | 18 | 54 |

Table 2: Session information

At the beginning of each session, subjects were seated at partitioned desks in the same room. Instructions were distributed (see Appendix A), and displayed on the subjects' computer screens, along with a die and a cup. The cup aided in die-rolling and, along with the partitions between desks, made it difficult for subjects to see others' die rolls. All instructions were read aloud by the experimenter in an effort to make the information common knowledge. Any questions subjects were answered privately.

The experiment began with general and Part 1 instructions. The instructions were similar across treatments, with differences only as required by the treatments.[9] In the N treatment, subjects played 50 rounds consecutively. Subjects in a session began each stage of each round at the same time. A round began with subjects being prompted to roll their dice, though they were not forced to do so. Subjects entered the result of the die roll into the computer; after all had done this, each chose their top-up contribution. There was no end-of-round feedback, so play would then continue to the next round.

In the R and G treatments, subjects played Part 1 for 25 rounds, with leaders introduced between rounds 25 and 26. (In particular, subjects in Part 1 would not have any knowledge of leaders, how they are chosen, and what information about them would be made available.) Subjects saw Part 2 instructions on their screens and on hard copies that were distributed and read aloud. The leader was prompted to choose one message: the "honest" message, the "cooperative" message, or no message (see Section 3.1 for their wording). The leader's message was revealed to the followers, along with the leader's history of recorded die rolls from the first 25 rounds. Subjects then played rounds 26-50, which were identical in structure to rounds 1-25; in particular, the leader's message and history were no longer visible once the 26th round began.

After round 50, subjects saw their earnings for the entire session on their screen and were then directed to complete a survey comprising demographic and attitudinal questions involving honesty and other pro-social behaviour (see Appendix B).[10] After completing the survey, participants were paid privately and in cash. Payments were the sum of earnings from all rounds (rounded up to the next 50 cents) plus a $5 payment for completing the survey; these averaged $28.49, $29.86 and $31.08 in the N, R and G

[9] In all of the treatments, the instructions contained the same four examples of outcomes from a hypothetical round. This was to illustrate the game design and how individual and group decisions determined an individual's earnings. One example had a subject free-riding in both types of allocations and the other group members playing cooperatively. A second example had a subject playing completely cooperatively and the other group members free-riding. Two other examples were included, showing different variations within these extremes. A disclaimer reading: "These are examples only. They are not intended to instruct you how to distribute your tokens" was placed after each example.

[10] The attitudinal questions were answered on a 10-point Likert scale with higher scores associated with more pro-social views. In our analysis, we use the average (arithmetic mean) response as a measure of pro-social beliefs. We acknowledge several issues with this approach. First, the average of ordinal-scale responses can be difficult to interpret. Second, subjects self-report, and may lie for reasons such as to justify the decisions made in the main part of the experiment, or to project an honest image. Third, since some questions refer to truth-telling and others to more general good behaviour such as cooperation, they may actually associate with different behaviour in the main experiment (e.g., a hypothetical subject strongly agreeing with the honesty questions but not the cooperation questions may accurately report die rolls, while another hypothetical subject with the opposite pattern of responses may over-report 5s and 6s to contribute more to the public good). We note that this last possibility seems not to have happened. A factor analysis of responses to the 15 attitudinal questions yields an eigenvalue of over 5 for the first factor and under 1 for the second factor, suggesting that all of the questions elicited the same characteristic from subjects.

treatments, respectively.[11] All sessions lasted for approximately 50 minutes.

## 5. Results

Table D1 in Appendix D reports subject characteristics by treatment. Other than age, there are no significant differences across our three treatments.[12]

Before discussing our results in detail, we note that one complication in analysing data from an experiment like ours is that an important piece of information – the sequence of actual die rolls for each subject – is not available. Instead, we have only the reported die rolls. While some levels of die-roll contributions, such as those averaging close to 0 or close to 2 over many rounds, are unlikely to have arisen from chance and honest reporting, they are not impossible, so we can never completely rule out the possibility that they did occur. Hence, we acknowledge we are abusing terminology by referring to subjects as having lied – or indeed, having told the truth – though we follow the literature in doing so.

With this caveat in mind, we will find it expedient to refer to particular patterns of behaviour as (likely) lying or truth-telling. The most direct such measure is the reported die rolls themselves. A chi-square test shows that for a fair die and with honest reporting, there is less than a 5 percent chance that a given pair of numbers (e.g., 1 and 2) is observed more than half the time. Given the prominence of p-values of 0.05 in hypothesis testing in economics, this seems like a natural threshold of potentially dishonest reporting. We say "potentially" since with 171 subjects and three relevant pairs of numbers (1/2, 3/4, or 5/6), false positives can be a concern. However, even after adjusting for these 513 comparisons, the probability that at least one subject will report at least one of these pairs of numbers 80 percent of the time (or more) is less than 0.001 (given a fair die and honest reporting). Hence any pair of numbers reported at least 20 out of the 25 rounds can be regarded as evidence of near-certain dishonesty. Reports of between 13 and 19 occurrences of a pair of numbers are in a grey area: possible dishonesty but with a reasonable doubt.

### 5.1 First-half behaviour

Overall, subjects contribute an average of 1.78 tokens to the public good in rounds 1-25, of which 0.90 and 0.88 are die-roll and top-up contributions respectively.[13] Given the endowment of 4 tokens per round, this

[11] We acknowledge that paying the sum of all rounds' payoffs, rather than a single randomly chosen round, may lead to wealth effects. If these wealth effects differ according to the treatment or the message sent by the leader, this could partly explain the results we observe.

[12] Even after correcting for multiple comparisons (e.g., Benjamini and Hochberg, 1995), age still differs significantly across treatments. See note 8 for some evidence that differences in age do not drive our results.

[13] Our analysis focuses on 25-round averages, treating a subject's average contributions in rounds 1-25 and in rounds 26-50 as single observations, and ignoring time trends. However, even though we do not provide feedback so subjects between rounds, time trends may emerge due to continued introspection, or to changing beliefs about the play of other group members. In

corresponds to a 45-percent contribution rate overall, which is typical in the literature for linear public goods without feedback. There were no significant differences in first-half contributions across treatments (Kruskall-Wallis test, group-level data, $p > 0.20$ for both die-roll and top-up contributions), which is unsurprising since leaders were not introduced until these rounds had ended.[14]

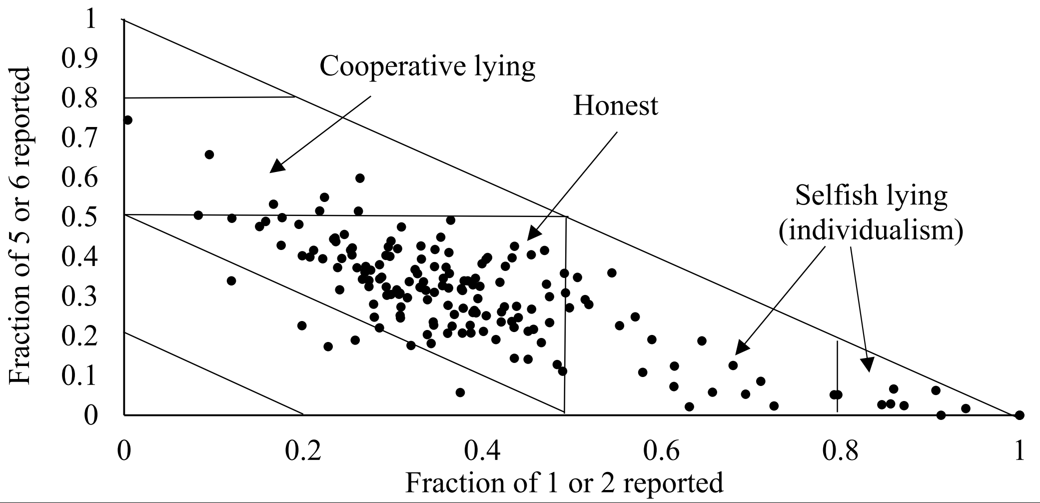



Figure 1: Distribution of reported die rolls,
all subjects, rounds 1-25 (first half)

Figure 1 illustrates some of the heterogeneity of subject behaviour in rounds 1-25, with a scatterplot of the fraction of reported die rolls of 5 or 6 (implying a contribution of 2), and 1 or 2 (implying a contribution of 0), with the residual being die rolls of 3 or 4 (implying a contribution of 1). The triangular region in the centre of this simplex corresponds to the region of "honest" behaviour as we defined it above, with "selfish lying" (individualism) towards the rightmost vertex and "cooperative lying" towards the uppermost vertex. Overall, 74.9 percent of subjects are in the centre "honest" region, and another 3.4 percent are in the region between this and the origin (which corresponds to possible dishonesty through over-reporting of 3 or 4 die rolls, though it is difficult to see a motive for lying in this manner). Most of the remaining subjects are in the "selfish lying" region (17.0 percent total, including 7.0 percent near-certain). No subjects are in the "near-certain cooperative lying" region, though 4.7 percent are in the "possible cooperative lying" region.

Appendix E, we provide an analysis of subject behaviour that allows for time trends. It can be seen there that even though there is some tendency for contributions to decline over time, our main results are not affected by time dependency.

[14] See Siegel and Castellan (1988) for descriptions of the non-parametric statistical tests used in this paper.

Another way of measuring the effects of lying about the die rolls is through the contribution levels themselves. Assuming a fair die and honest reporting, the probability that a subject's average die-roll contribution over 25 rounds is below 0.6 tokens or above 1.4 tokens is about 1 percent, and there is a 0.1 percent chance of it being below 0.48 tokens or above 1.52 tokens. So, a potential rule of thumb for categorising 25-round individual average contributions would treat an average of 0.5 or less as "selfish lying", 1.5 or more as "cooperative lying", and between 0.5 and 1.5 as "honesty". According to this definition, the vast majority of subjects are honest (87.7 percent), with only 11.1 percent selfish liars and 1.2 percent cooperative liars.

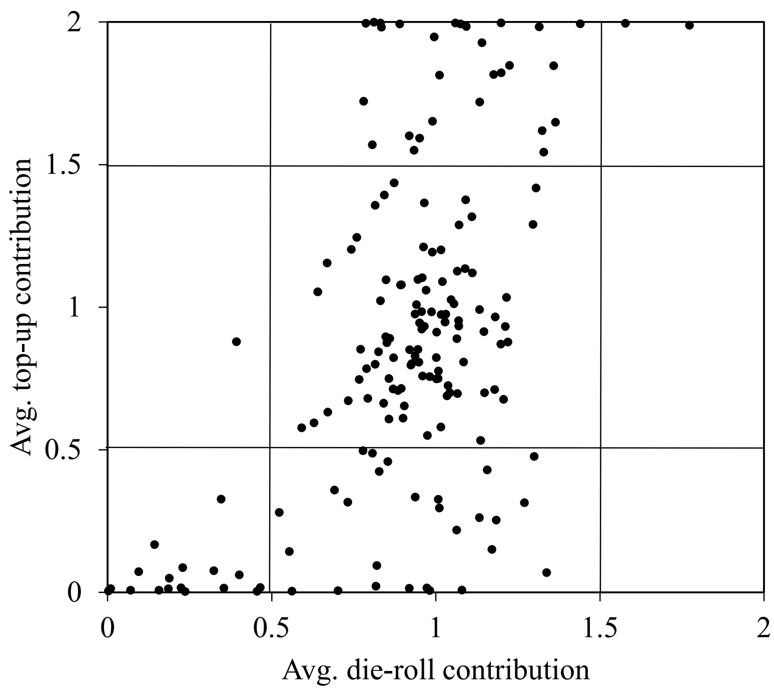



Figure 2: Scatterplot of die-roll and top-up contributions, all subjects, rounds 1-25

Figure 2 shows this distribution of average die-roll contributions in rounds 1-25, as the horizontal coordinate of a scatterplot with the corresponding top-up contributions as the vertical coordinate. The vertical lines at 0.5 and 1.5 distinguish the honest, the cooperative liars, and the selfish liars, as noted above. Additional information that reinforces this taxonomy is provided by the top-up contributions. The large numbers of plotted points in the bottom-centre and top-centre rectangles, corresponding to average die-roll contributions between 0.5 and 1.5 and average top-up contributions either less than 0.5 or more than 1.5, suggest that an aversion to lying constrains some individualistic subjects from free-riding by under-reporting their die rolls, but it also constrains some intrinsically cooperative subjects from further benefiting the group by over-reporting their die rolls. Thus, our use of the honesty norm to implement a norm with an intermediate level of contributions "works". By the same token, all but one of the points in the leftmost three rectangles lie in the bottom-left rectangle, indicating that low die-roll contributions typically result from individualism that outweighs any aversion to lying, rather than a run of low actual die rolls. Similarly, the fact that all of the points in the rightmost three rectangles lie in the top-right rectangle suggests that high die-roll contributions resulted from intrinsic group-oriented motivation that outweighs any aversion to lying – though admittedly this is based on a small number of observations.

### 5.2 Leaders' messages

Table 3 shows leaders' message frequencies. In both R (random-leader) and G (group-oriented-leader) treatments, cooperate messages predominate: they are sent almost two-thirds of the time in the R treatment and five-sixths of the time in the G treatment. (The difference in frequencies between these two treatments is not significant: chi-square test, $p \approx 0.26$.) Due to the rarity of messages other than cooperate, in our analysis below we will pool the blank and honest messages.

| Treatment | No message | "Blank" message | "Honest" message | "Cooperate" message |
|---|---|---|---|---|
| No leader (N) | 100% (21/21) | -- | -- | -- |
| Random leader (R) | -- | 16.7% (3/18) | 22.2% (4/18) | 61.1% (11/18) |
| Group-oriented leader (G) | -- | 0.0% (0/18) | 16.7% (3/18) | 83.3% (15/18) |

Table 3: Message choices by leaders

### 5.3 Second-half behaviour

Figure 3 shows some information about aggregate-level contributions across treatments: die-roll contributions in the left panel and top-up (extra) contributions in the right panel. Each panel shows a scatterplot with the average for the first half (rounds 1-25) as the horizontal coordinate and the second-half average (rounds 26-50) as the vertical coordinate. Five points are plotted in each panel, corresponding to the N (no-leader) treatment, the R (random-leader) and G (Group-oriented-leader) treatments conditional on a c (cooperate) message being sent, and the R and G treatments conditional on either other message (blank or honest) being sent.

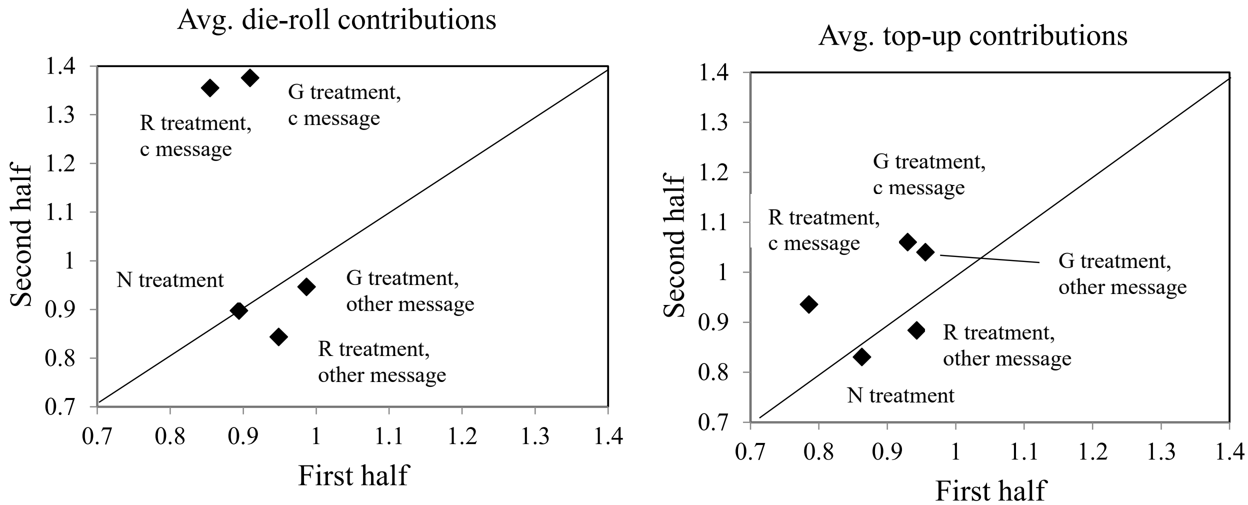



Figure 3: Aggregate contributions in first half and second half,
by treatment and leader's message

As the figure shows, when there is no group leader, or when the leader sends either the blank or the honest message, there is little change in average die-roll contributions from the first half of the session (before the message is sent) to the second half (after the message is sent). By contrast, when there is a leader and she sends the cooperate message, average die-roll contributions increase by about one-half token from the first half to the second half. These latter increases are highly significant (two-tailed Wilcoxon signed-ranks test, group-level data, $p < 0.001$ for R treatment, G treatment, or both pooled), while there is no significant difference in any of the cases without a cooperate message ($p > 0.20$ for N treatment, R treatment, G treatment, or all three pooled).

The effects on top-up contributions are smaller in scale, but largely reinforce the effects on die-roll contributions. There is no significant difference in top-up contributions when anything other than a cooperate message is sent ($p > 0.20$ for N, R, or G treatment, or all three pooled). After a cooperate message, top-up contributions increase, though the difference is significant at the 5-percent level only if the R and G treatments are pooled ($p \approx 0.016$, versus $p \approx 0.15$ and 0.087 for the R and G treatments individually).

Figures 4 and 5 show the individual-level analogues to Figure 3's left panel for followers and leaders respectively: scatterplots of first- and second-half contributions according to whether a cooperative message was sent (right panel) or not (left panel). When either no message (N treatment) or a message other than cooperate (R or G treatment) was sent, individual-level behaviour remained roughly the same from the first half to the second: subjects with low (high) die rolls in the first half tended also to have them in the second, while those with mid-level die rolls in the first half tended to have these in the second half as well (except for two subjects who reported low die rolls exclusively in the second half).

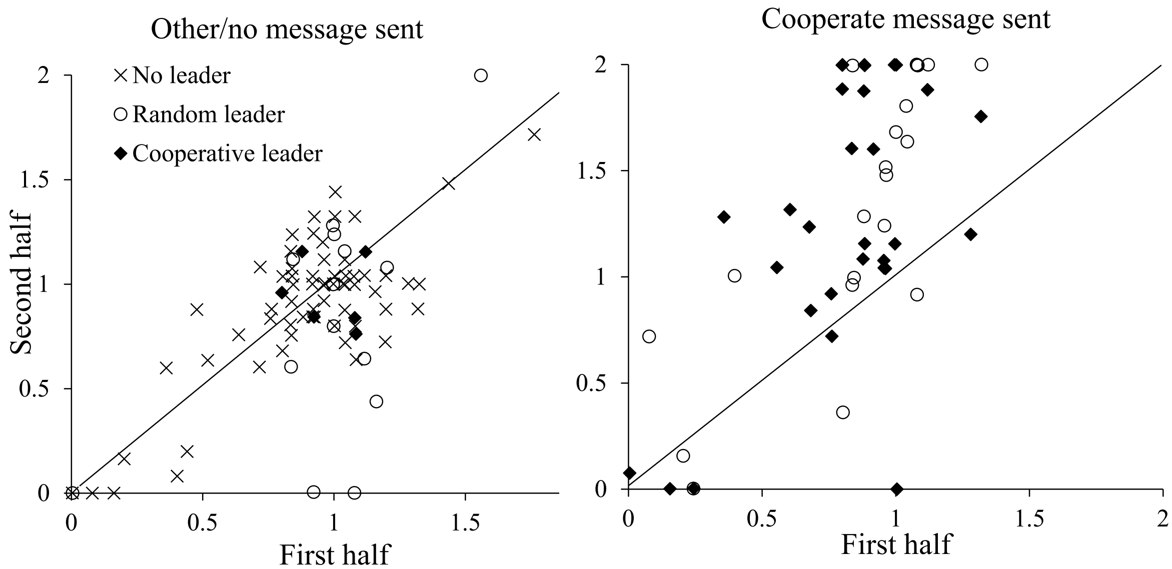



Figure 4: Individual followers' die-roll contributions in first and second half, by treatment/message

The effect of a cooperate message appears to be more heterogeneous than that of the other messages. The handful of subjects who were reporting low die rolls in the first half mostly continued to do so in the second half, though a few seem to switch to reporting honestly. The majority of subjects – who were reporting mid-level die rolls in the first half – split between continuing to do so and reporting high rolls. Only one subject reported low rolls in the second half after reporting honestly in the first half.

Figure 5 shows that the relationship between leaders' messages and their own subsequent behaviour is similar to that for followers' behaviour. In both random- and cooperative-leader treatments, any message other than the cooperate one has no systematic relationship with their subsequent die-roll contributions (two-tailed Wilcoxon test, $p > 0.20$ for either treatment separately or both pooled), while the cooperate message is associated with an increase in leaders' die-roll contributions ($p \approx 0.028$ and 0.003 for the R and G treatments respectively, $p < 0.001$ for both pooled). The results are similar for total contributions (not shown in the figure): a significant increase following a cooperate message ($p \approx 0.04$ and 0.02 for the R and G treatments respectively, $p \approx 0.002$ for both pooled), but no systematic effect after any other message ($p > 0.20$ for either treatment separately and for both pooled).

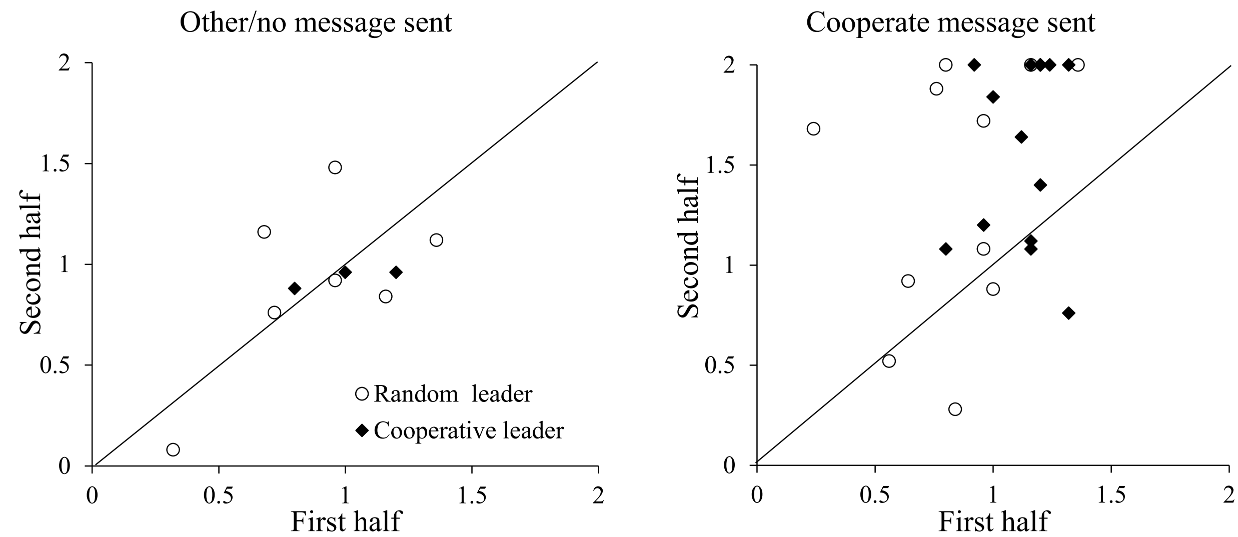



Figure 5: Individual leaders' die-roll contributions in first and second half, by treatment/message

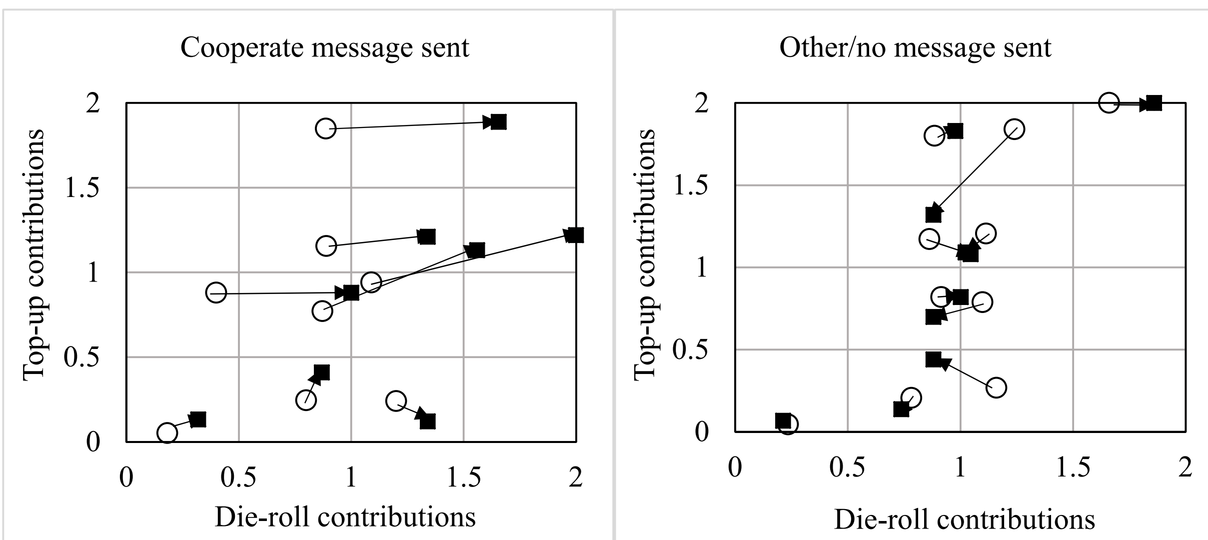



Figure 6: Followers' first- and second-half die-roll and top-up contributions, grouped by coordinates

Figure 6 provides additional information about followers' changes in behaviour from the first to the second half. The figure shows ordered pairs of die-roll and top-up contributions, for the first half as open circles, and for the second half as dark squares. To avoid cluttering the figure, instead of showing each individual separately, we aggregate them according to their first-half behaviour. Specifically, we divide the 2x2 coordinate square into 16 equal-sized regions (note the similarity of these regions to those in Figure 2), and pool all of the subjects whose first-half combination of die-roll and top-up contributions lay in the same region. (Note that some regions are empty.) The circle in a given region represents average first-half behaviour of the subjects in that region, and the square connected by an arrow to the circle represents the same subjects' second-half behaviour. Thus, the arrow itself shows the magnitude and direction of change from first to second half. We pool the R and G treatments with cooperate message in the left panel, and all three treatments with any other (or no) message in the right panel.

As we have seen previously, when anything but a cooperate message is sent (right panel), there seems to be no systematic effect on followers' die-roll or top-up contributions. We can also see that this (lack of) effect does not vary systematically according to first-half behaviour, as the arrows in the right panel appear fairly random in their magnitudes and directions.

On the other hand, when a cooperate message is sent (left panel), we observe increase in some groups' second-half contributions, which were obscured in Figure 3's aggregate results. Interestingly, the effect on die-roll contributions seems to depend mainly on followers' first-half ***top-up*** contributions, which as noted earlier, reflect subjects' propensity to cooperate when the honesty norm does not apply. Followers with low top-up contributions in the first half – irrespective of their die-roll contributions – are relatively unaffected by the cooperate message, with those receiving a cooperate message after top-up contributions of less than one-half token per round increasing their die-roll contributions by only 0.106 tokens on average. By contrast, followers with first-half top-up contributions larger than this who receive a cooperate message increase their die-roll contributions by 0.663 tokens in the second half.

This heterogeneity based on first-half top-up contributions extends to changes in the top-up contributions themselves. After receiving a cooperate message, followers with more (less) than half a token per round in first-half top-up contributions increase their top-up contributions in the second half by 0.197 (0.096) tokens. Thus, it seems as if the positive effect of cooperate messages comes despite having little or no impact on individualistic followers – whether "pure free riders" who have low die-roll and low top-up contributions, or "acculturated free riders" who have low top-up contributions but moderate die-roll contributions due to reporting their rolls honestly (as the honesty norm prescribes). Rather, the effect of these messages is driven by their impact on group-oriented followers. This includes "acculturated cooperators" who make moderate to high top-up contributions but average die-roll contributions due to honest reporting, and even nearly pure cooperators who over-report their die rolls but not completely so.

Tables 4 and 5 show results from Tobit regressions. The dependent variable is either die-roll or top-up contributions, averaged over the 25 rounds of the second half. The sample is either all subjects and treatments (for Table 4) or followers in the R and G treatments (for Table 5). In Table 4, the main explanatory variables are indicators either for the treatment (R or G) or for combinations of treatment and message (cooperate or other), so that the no-leader treatment is the baseline. The results presented above suggest that the message has at least as much of an effect on contributions as the treatment, but the effect of the treatment itself can be interpreted as either a standard treatment effect or that of "intent to treat" for a cooperate message.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable (Tobit): | Die-roll contrib. ($2^{nd}$-half avg.) | | Top-up contrib. ($2^{nd}$-half avg.) | |
| R treatment | 0.224** | | 0.039 | |
| | (0.096) | | (0.100) | |
| R treatment + c message | | 0.469*** | | 0.114 |
| | | (0.061) | | (0.105) |
| R treatment + other message | | −0.191* | | −0.071 |
| | | (0.097) | | (0.160) |
| G treatment | 0.344*** | | 0.075 | |
| | (0.079) | | (0.082) | |
| G treatment + c message | | 0.407*** | | 0.084 |
| | | (0.077) | | (0.091) |
| G treatment + other message | | −0.049 | | 0.034 |
| | | (0.071) | | (0.076) |
| Leader | 0.095 | 0.091 | 0.130 | 0.126 |
| | (0.103) | (0.097) | (0.090) | (0.089) |
| Die-roll contribution | 0.701*** | 0.772*** | 0.078 | 0.104 |
| ($1^{st}$-half avg.) | (0.127) | (0.097) | (0.148) | (0.149) |
| Top-up contribution | 0.285*** | 0.270*** | 0.888*** | 0.883*** |
| (1st-half avg.) | (0.073) | (0.058) | (0.065) | (0.064) |
| Demographics? | Yes | Yes | Yes | Yes |
| Observations | 171 | 171 | 171 | 171 |
| \|ln(L)\| | 129.33 | 107.16 | 131.29 | 130.04 |

Table 4: Factors affecting contributions (all subjects) – average marginal effects and standard errors (clustered by group)

Additional right-hand-side variables are an indicator for the leader, the subject's average die-roll and top-up contributions in the first half, a constant term, and the demographic and attitudinal variables collected in the post-experimental survey.[15] The models are estimated using Stata (version 15) and incorporate clustering by group (clustering by session yielded similar results).

The main results in Table 4 are consistent with our earlier descriptive statistics. When we do not control for the message, die-roll contributions are significantly higher in both the R and G treatments compared to the N treatment, while top-up contributions are also higher but not significantly so. When we do control for the message, we see that cooperate messages account for roughly the entire treatment effects. Die-roll contributions are significantly higher after the leader sends a cooperate message in either the R or G treatment compared to the N treatment, while either the honesty message or no message has either no significant effect on die-roll contributions or actually decreases them slightly, and conditional on the message, there are no significant differences between the R and G treatments. The lack of any significant positive effects of our "other message" variables here (and similarly for our "Group-oriented leader" dummy in Table 5 below), suggest that restart effects – which were possible in our R and G treatments but not our N treatment – do not play a large role in explaining behaviour.

Cooperate messages also have a positive but insignificant effect on top-up contributions (model 4 in Table 4), so that their effect on ***total*** contributions is positive and significant (marginal effects in the R and G treatments of +0.193*** and +0.147** respectively, with associated standard errors of 0.062 and 0.058). As with die-roll contributions, other messages have no significant effect on top-up contributions, nor does the way the leader is chosen (R versus G treatment).

The Tobits in Table 5 use similar methodology to models (2) and (4) in Table 4. In addition to those differences mentioned earlier, we include the leader's average part-1 die-roll contribution (which followers were informed of at the beginning of part 2) as an explanatory variable. We also interact this variable, as well as the follower's die-roll and top-up contributions, with the cooperate-message dummy, to allow the effect of the message to vary according to the leader's observed history and the follower's own past behaviour. None of these interaction terms turns out to be significantly different from zero, so we leave them out of the table to save space (though we leave them in the regressions themselves).

The results in Table 5 largely reinforce those in Table 4, and like those earlier results, are fairly robust to whether demographic and attitudinal variables are included. As before, we see that the leader's message is very important, while how the leader is chosen has little effect. Indeed, a cooperate message not only increases followers' die-roll contributions, but also their top-up contributions, despite the message not

[15] Our results are robust to leaving out these demographic variables, suggesting that the treatment effects we observe are not driven by the significant differences in age across treatments noted at the beginning of Section 5. Moreover, age itself is not significant in any of the regressions.

referring to these at all. We also see that "leading by example" is not particularly important: while the leader's observable past (part 1) behaviour is positively related to followers' (part 2) die-roll and top-up contributions, neither of these effects is significant.

| | (5) | (6) | (7) | (8) |
|---|---|---|---|---|
| Dependent variable (Tobit): | Die-roll contrib. ($2^{nd}$-half avg.) | | Top-up contrib. ($2^{nd}$-half avg.) | |
| Group-oriented leader | 0.093 | 0.015 | −0.015 | 0.016 |
| | (0.142) | (0.096) | (0.103) | (0.084) |
| c message | | 0.707*** | | 0.303*** |
| | | (0.098) | | (0.109) |
| Leader die-roll contribution | 0.175 | 0.030 | 0.393 | 0.246 |
| ($1^{st}$-half avg.) | (0.339) | (0.210) | (0.273) | (0.188) |
| Die-roll contribution | 0.821*** | 1.123*** | 0.197 | 0.397 |
| ($1^{st}$-half avg.) | (0.209) | (0.203) | (0.228) | (0.262) |
| Top-up contribution | 0.347*** | 0.242*** | 0.836*** | 0.778*** |
| ($1^{st}$-half avg.) | (0.125) | (0.112) | (0.097) | (0.101) |
| Demographics? | Yes | Yes | Yes | Yes |
| Observations | 72 | 72 | 72 | 72 |
| \|ln(L)\| | 69.32 | 55.18 | 62.70 | 57.06 |

Table 5: Factors affecting contributions (followers in leader treatments) – average marginal effects and standard errors (clustered by group)

### 6. Conclusion

We investigate the effect of leadership – how leaders are chosen, what they communicate, and their past behaviour – on two of the ways groups can underperform: individualism (free-riding at the expense of the group) and a strict adherence to a norm. We use a novel public-good setting in which a portion of the endowment is entirely discretionary ("top-up contributions") while the rest of the endowment, while also ultimately discretionary, is framed as being determined by the (reported) result of a die roll ("die-roll contributions"). To the extent that subjects feel bound to report die rolls honestly, the norm of honesty prevents the worst excesses of individualism (an honest free rider contributes more than the minimum), but at the cost of restricting cooperative behaviour that would have helped the group (an otherwise group-oriented person will contribute less if honest).

We find that behaviour is not affected by the presence of leaders per se, by the way leaders are chosen (randomly or based on past cooperation), or by observed information about leaders' past decisions. By contrast, we find that the leader's message can have a strong effect. A message that renounces both individualism and the norm of honesty leads to increases in public-good contributions, and these increases average over half a token per subject-round – an increase of more than 25 percentage points relative to the per-round endowment of 4 tokens. The effect is seen in money earnings as well: subjects receiving these pro-cooperation messages earn about 20 percent more in the second half of the session (after the message is sent) than in the first half, while everyone else earns the same or slightly less in the second half than in the first half.

Examination of the disaggregated data suggests that most of the effect of these messages is due to their impact on those group members who had already shown some propensity towards group-oriented behaviour (as evidenced by substantial top-up contributions in the first half), but appear to have been constrained by the norm of honesty. After receiving the message, these subjects went from reporting mostly honestly to sizably over-reporting the die rolls – increasing their die-roll contributions – while maintaining or even increasing their top-up contributions (see Figure 6). By contrast, subjects who mostly free rode prior to receiving the leader's message typically continued to do so.[16]

Perhaps just as important as what we did see in the experiment is what we did ***not*** see. As usual in public-good experiments, we did not see complete opportunism in the sense of widespread free-riding (though some degree of free-riding was certainly observed). But we also did not see substantial evidence of opportunistic leaders sending cooperative messages and then free-riding. Nor did we see opportunistic followers observing cooperative messages or cooperative past leader behaviour and then increasing their free-riding.

Other potential patterns of behaviour that were not observed suggest that the honesty norm worked well as a group culture. Despite the identical roles that die-roll and top-up contributions played in the individual (and group) payoff functions, subjects did not treat them symmetrically. For example, Figure 2 shows many subjects with first-half die-roll contributions averaging close to half the available tokens and top-up contributions averaging nearly zero or nearly all of the available tokens, while the reverse pattern is almost never seen. Again, it seems that a substantial fraction of subjects acted as though they were constrained by the honesty norm from either free-riding or group-optimal contributing. Relatedly, we do not see changes in die-roll contributions from the first half to the second half countervailed by opposite-direction changes in top-up contributions – that is, the increase in die-roll contributions resulting from a

[16] It is not unusual for treatment effects to vary across individuals based on behavioural characteristics. As a recent example, Jacquemet et al. (2019) find that the effect of asking subjects to swear a truth-telling oath reduces lying by "partial liars", but not by "chronic liars" or, of course, truth tellers.

cooperate message does not crowd out top-up contributions – as would result from subjects' having a preferred total contribution in mind, and using top-up contributions to offset the variation in die rolls.

We would like to emphasise yet again that our use of the honesty norm to implement an inefficient outcome is not meant as an attack on honesty more broadly. We are well aware that in a wide variety of settings, institutions "work" because people behave more honestly than narrow self-interest would dictate (as one example, people cheat on self-assessed income tax returns far less than the probability of and punishment for getting caught would imply – reducing the cost of enforcement to the taxpayer). Indeed, our successful use of honesty here is a testament to its prevalence.

Our findings have immediate implications for decision makers in organisations, particularly those designing governance structures. The trend towards increased managerialism in organisations is based chiefly on the assumption that if free-riding (or more generally, rent-seeking behaviour by individuals) is not prevented, it becomes pervasive. Our results suggest that although systems designed to limit free-riding may succeed in doing so, they do so at the cost of stifling the best performers. Our results raise the possibility that managers may do better with a more Pollyannaish approach: forgo imposing rules that can inhibit creativity, promote the message that when the group succeeds the individual succeeds, and trust that their employees' desire to be part of a successful group will win out.[17]

[17] We hope that these results will convince decision makers to think carefully about how their groups currently operate – and whether increasing micro-management will improve or worsen their performance – before introducing them, though we have little confidence that those decision makers would read a paper like this. A further complication is that behaviour that is optimal for the group may not be optimal for the larger society (e.g., price fixing by a group of firms). Because of this possibility, we stop short of making policy recommendations based on our findings.

## References


Abeler, J., A. Becker and A. Falk. 2014. Representative evidence on lying costs. Journal of Public Economics 113, 96-104.

Ahn, T.K., R.M. Isaac and T.C. Salmon. 2009. Coming and going: experiments on endogenous group sizes for excludable public goods. Journal of Public Economics 93, 336-351.

Alempaki, D., G. Doğan and S. Saccardo. 2016. Deception and Reciprocity.

Balliet, D. 2010. Communication and cooperation in social dilemmas: a meta-analytic review. Journal of Conflict Resolution 54, 39-57.

Barr, A. and G. Michailidou. 2017. Complicity without connection or communication. Journal of Economic Behavior and Organization 142, 1-10.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B Methodological) 57, 289-300.

Bochet, O., T. Page and L. Putterman. 2006. Communication and punishment in voluntary contribution experiments. Journal of Economic Behavior and Organization 60, 11-26.

Bolton, G.E., E. Katok and A. Ockenfels. 2005. Cooperation among strangers with limited information about reputation. Journal of Public Economics 89, 1457-1468.

Charness, G., N. Du and C.L. Yang. 2011. Trust and trustworthiness reputations in an investment game. Games and Economic Behavior 72, 361-375.

Chaudhuri, A. 2011. Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. Experimental Economics 14, 47-83.

Cinyabuguma, M., T. Page and L. Putterman. 2005. Cooperation under the threat of expulsion in a public goods experiment. Journal of Public Economics 89, 1421-1435.

Croson, R.T.A. 2007. Theories of commitment, altruism and reciprocity: evidence from linear public goods games. Economic Inquiry 45, 199-216.

d'Adda, G., Darai, D., Pavanini, N. and Weber, R. A. 2017. Do leaders affect ethical conduct? Journal of the European Economic Association 15, 1177-1213.

Dai, Z., Galeotti, F. and Villeval, M. C. 2017. Cheating in the lab predicts fraud in the field: an experiment in public transportation. Management Science 64, 1081-1100.

Diekmann, A., W. Przepiorka and H. Rauhut. 2015. Lifting the veil of ignorance: an experiment on the contagiousness of norm violations. Rationality and Society 27, 309-333.

Duffy, J. and N. Feltovich. 2002. Do actions speak louder than words? An experimental comparison of observation and cheap talk. Games and Economic Behavior 39, 1-27.

Duffy, J. and N. Feltovich. 2006. Words, deeds, and lies: strategic behaviour in games with multiple signals. Review of Economic Studies 73, 669-688.

Eckel, C. C. and P.J. Grossman. 1996. The relative price of fairness: gender differences in a punishment game. Journal of Economic Behavior and Organization 30, 143-158.

Ehrhart, K.-M and C. Keser. 1999. Mobility and cooperation: on the run. Working Paper 99s-24. CIRANO, Montreal.

Erat, S. and U. Gneezy. 2012. White Lies. Management Science 58, 723-733.

Falk, A. and U. Fischbacher. 2006. A theory of reciprocity. Games and Economic Behavior 54, 293-315.

Feltovich, N. and P.J. Grossman. 2015. How does the effect of pre-play suggestions vary with group size? Experimental evidence from a threshold public-good game. European Economic Review 39, 263-280.

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

Friedman, D. and R. Oprea. 2012. A continuous dilemma. American Economic Review 102, 337-363.

Gächter, S. and A. Riedl 2005. Moral property rights in bargaining with infeasible claims. Management Science 51, 249-263.

Gibson, R., C. Tanner and A. Wagner. 2013. Preferences for truthfulness: heterogeneity among and within individuals. American Economic Review 103, 532-548.

Gneezy, U. 2005. Deception: the role of consequences. American Economic Review 95, 384-394.

Greiner, B. 2015. Subject pool recruitment procedures: organizing experiments with ORSEE. Journal of the Economic Science Association 1, 114-125.

Guth, W., V.M. Levati, M. Sutter, and E. van der Heijden. 2007. Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics,* 91, 1023-1042.

Hoffmann, M., T. Lauer and B. Rockenbach. 2013. The royal lie. Journal of Economic Behavior and Organization 93, 305-313.

Houser, D., Levy, D. M., K. Padgitt, S.J. Peart, and E. Xiao. 2014. Raising the price of talk: An experimental analysis of transparent leadership. *Journal of Economic Behavior and Organization,* 105, 208-218.

Irlenbusch, B. and J. Ter Meer. 2013. Fooling the nice guys: explaining receiver credulity in a public good game with lying and punishment. Journal of Economic Behavior and Organization 93, 321-327.

Jack, B. and M. Recalde. 2015. Leadership and the voluntary provision of public goods: field evidence from Bolivia. Journal of Public Economics 122, 80-93.

Jacquemet, N., S. Luchini, A. Malézieux and J.F. Shogren (2019), Who'll stop lying under oath? Empirical evidence from Tax Evasion Games. Working paper, University of Exeter.

Kahneman, D., J. L. Knetsch and R. H. Thaler. 1986. Fairness and the assumptions of economics. Journal of Business 59, S285-S300.

Karagözoğlu, E. and A. Riedl 2015. Performance information, production uncertainty and subjective entitlements in bargaining. Management Science 61, 2611-2626.

Komai, M. and P.J. Grossman. 2009. Leadership and group size: An experiment. *Economics Letters,* 105, 20-22.

Komai, M., P.J. Grossman, and T. Deters. 2011. Leadership and information in a single-shot collective action game: An experimental study. *Managerial and Decision Economics,* 32, 119-134.

Koukoumelis, A., M. Levati and J. Weisser. 2012. Leading by words: a voluntary contribution experiment with one-way communication. Journal of Economic Behavior and Organization 81, 379-390.

Ledyard, J.O. 1995. Public goods: some experimental results. In J. Kagel and A. Roth (Eds.), Handbook of experimental economics. Princeton: Princeton University Press (Chap. 2).

Levy, D.M., K. Padgitt, S.J. Peart, D. Houser, and E. Xiao. 2011. Leadership, cheap talk and really cheap talk. *Journal of Economic Behavior and Organization*, 77, 40-52.

Lewis, A., A. Bardis, C. Flint, C. Mason, N. Smith, C. Tickle and J. Zinser. 2012. Drawing the line somewhere: an experimental study of moral compromise. Journal of Economic Psychology 33, 718-725.

Oprea, R., G. Charness and D. Friedman. 2014. Continuous time and communication in a public good experiment. Journal of Economic Behavior and Organization 108, 212-223.

Page, T., L. Putterman and B. Unel. 2005. Voluntary association in public goods experiments: reciprocity, mimicry and efficiency. Economic Journal 115, 1032-1053.

Potters J., M. Sefton, and L. Vesterlund. 2005. After you: Endogenous sequencing in voluntary contribution games. *Journal of Public Economics,* 89, 1399-1419.

Potters J., M. Sefton, and L. Vesterlund. 2007. Leading by example and signaling in voluntary contribution games: An experimental study. *Economic Theory* 33, 169-182.

Sally, D. 1995. Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. Rationality and Society 7, 58-92.

Seinen, I. and A. Schram. 2006. Social status and group norms: indirect reciprocity in a repeated helping experiment. European Economic Review 50, 581-602.

Siegel, S.and N.J. Castellan, Jr. 1988. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

Utikal, V. and U. Fischbacher. 2013. Disadvantageous lies in individual decisions. Journal of Economic Behavior and Organization 85, 108-111.

Weber, R.A. 2006. Managing growth to achieve efficient coordination in large groups. American Economic

Review 96, 114-126.

Weisel, O. and S. Shalvi. 2015. The collaborative roots of corruption. Proceedings of the National Academy of Sciences 112, 10651-10656.

Wiltermuth, S. 2011. Cheating more when spoils are split. Organizational Behaviour and Human Decision Processes 115, 157-168.

## Appendix A: Instructions from the experiment

Below are the instructions from the experiment, as they appeared to subjects with two exceptions: (1) to save space, horizontal lines are used to separate different screens; (2) wording inside square brackets is added here for the reader's information, but was not seen by subjects. Subjects received the general and part 1 instructions at the beginning of the experiment and part 2 instructions before beginning part 2.

---

**General Instructions [N treatment]**

Welcome and thank you for participating.

We ask that you do not talk with the other participants during the experiment. If you have questions at any time, please feel free to raise your hand and someone will come to you to answer them.

You are about to participate in a decision-making experiment. Please read the instructions carefully, as how much money you earn may depend on how well you understand them.

The experiment consists of two parts. Part 1 is a decision-making task. You will receive detailed instructions before beginning Part 1. Part 2 is a survey and you will receive a flat payment of $5 for completing the survey.

Your total earnings will be comprised of the $5 payment for completing the survey and your earnings from part 1. You will be paid privately and in cash at the end of the experiment.

You will be assigned a unique identification number. This ID number is used to maintain the anonymity of your decisions from other participants. You will also use this number to collect your earnings at the end of the experiment.

Please turn off your mobile phone.
Please do not write on the instruction sheets.
At the end of the session, please hand back your instruction sheet.

---

**General Instructions [R/G treatment]**

Welcome and thank you for participating.

We ask that you do not talk with the other participants during the experiment. If you have questions at any time, please feel free to raise your hand and someone will come to you to answer them.

You are about to participate in a decision-making experiment. Please read the instructions carefully, as how much money you earn may depend on how well you understand them.

The experiment consists of two parts. Part 1 is a decision-making task. You will receive detailed instructions before beginning Part 1. Part 2 is a decision-making task. You will receive detailed instructions before beginning Part 2. Part 3 is a survey and you will receive a flat payment of $5 for completing the survey.

Your total earnings will be comprised of the $5 payment for completing the survey and your earnings from part 1. You will be paid privately and in cash at the end of the experiment.

You will be assigned a unique identification number. This ID number is used to maintain the anonymity of your decisions from other participants. You will also use this number to collect your earnings at the end of the experiment.

Please turn off your mobile phone.
Please do not write on the instruction sheets.

At the end of the session, please hand back your instruction sheet.

---

**Part 1 Instructions [all treatments]**

You have been randomly divided into groups of three. Your group will **not** change during the remainder of the session.

Each of you will have 4 tokens (worth $0.08 each) to distribute.

Tokens will be distributed between a group account and your individual account.

For the group account, every token that is allocated by you and the other two members of your group to the group account will be doubled. You and your fellow group members will share equally in the group account.

For your individual account, every token that is allocated by you to your individual account will remain yours.

Your earnings for each period will equal your earnings from the tokens in your individual account plus your equal share of the earnings from the group account.

In each period you will roll the die that was given to you. You will record your roll. The roll you record will dictate how 2 of your 4 tokens will be distributed between the group's account and your individual account. The table below indicates what the number you record means for the allocation of your tokens.

| | Tokens allocated to… | |
|---|---|---|
| Roll Recorded | Group Account | Individual Account |
| 1,2 | 0 | 2 |

| | | |
|---|---|---|
| 3,4 | 1 | 1 |
| 5,6 | 2 | 0 |

For your remaining 2 tokens, you will decide how you wish to allocate them between your individual account and the group account.

---

**Example 1 [all treatments]**

You record a die roll of 5. Two of your tokens will automatically be allocated to the group account and 0 will automatically be allocated to the individual account. You allocate the remaining 2 tokens to your group account. Your fellow group members both recorded a 1, they allocate 0 token to the group account and 2 tokens to the individual account. They each allocate their remaining 2 tokens to their individual accounts. There would be a total of 4 tokens in the group account. Doubling this gives 8, so each member of the group would receive 2.66 tokens from the group account. You would have a total of 2.66 tokens (2.66 from the group account and 0 from the individual account). Your earnings for the period would be $0.21 (= 2.66*$0.08).

*These are examples only. They are not intended to instruct you how to distribute your tokens.*

---

**Example 2 [all treatments]**

You record a die roll of 1. None of your tokens will automatically be allocated to the group account and 2 will automatically be allocated to the individual account. You allocate the remaining 2 tokens to your individual account. Your fellow group members both recorded a 5, they allocate 2 token to the group account and 0 tokens to the individual account. They each allocate their remaining 2 tokens to the group account. There would be a total of 8 tokens in the group account. Doubling this gives 16, so each member of the group would receive 5.33

tokens from the group account. You would have a total of 9.33 tokens (5.33 from the group account and 4 from the individual account). Your earnings for the period would be $0.75 (= 9.33*$0.08).

*These are examples only. They are not intended to instruct you how to distribute your tokens.*

---

**Example 3 [all treatments]**

You record a die roll of 1. None of your tokens will automatically be allocated to the group account and 2 will automatically be allocated to the individual account. You allocate the remaining 2 tokens to the group account. Your fellow group members both recorded a 4 they allocate 1 token to the group account and 1 token to the individual account. They each allocate their 1 of their remaining tokens to the group account and 1 to their individual accounts. There would be a total of 6 tokens in the group account. Doubling this gives 12, so each member of the group would receive 4 tokens from the group account. You would have a total of 6 tokens (4 from the group account and 2 from the individual account). Your earnings for the period would be $0.48 (= 6*$0.08).

*These are examples only. They are not intended to instruct you how to distribute your tokens.*

---

**Example 4 [all treatments]**

You record a die roll of 5. Two of your tokens will automatically be allocated to the group account and 0 will automatically be allocated to the individual account. You allocate the remaining 2 tokens to the group account. Your fellow group members both recorded a 4 they allocate 1 token to the group account and 1 token to the individual account. They each allocate their 1 of their remaining tokens to the group account and 1 to their individual accounts. There would be a total of 8 tokens in the group account. Doubling this gives 16, so each member of the group would receive 5.33 tokens from the group account. You would have a total of 5.33

tokens (5.33 from the group account and 0 from the individual account). Your earnings for the period would be $0.43 (= 5.33*$0.08).

*These are examples only. They are not intended to instruct you how to distribute your tokens.*

*This task will be played for 50 periods [25 periods in R/G treatments]*

*If no questions we will begin Part 1*

---

**Part 2 Instructions [R treatment]**

For Part 2, you will continue to play the same Part 1 game with two changes.

First, a leader for your group will be selected at random. A summary of what the leader recorded for his/her first 25 period die rolls will be provided to the other members of his/her group.

Second, the leader will have the option to send a message to his/her fellow group members about how to play the game for the next 25 periods.

The leader will choose from among the following three options:

1. No message
2. Everyone should record their true roll
3. Everyone should record a 5 or 6. If we all do this we will all earn the most money

---

**Part 2 Instructions [G treatment]**

For Part 2, you will continue to play the same Part 1 game with two changes.

First, a leader for your group will be selected. The player selected to be the leader is the player in your group who recorded the most 5,6 rolls during the first 25 periods. A summary of what the leader recorded for his/her first 25 period die rolls will be provided to the other members of his/her group.

Second, the leader will have the option to send a message to his/her fellow group members about how to play the game for the next 25 periods.

The leader will choose from among the following three options:

1. No message
2. Everyone should record their true roll
3. Everyone should record a 5 or 6. If we all do this we will all earn the most money

---

**[Survey] Instructions [all treatments]**

You are about to begin Part 2 [3 for R/G treatments] of the experiment, a 25 question survey. For completing the survey you will be paid $5. If at any time you have questions, raise your hand and we will come to you to answer them.

## Appendix B: Survey questions

### Demographic questions

Age (in years, as on last birthday): [positive integer]

What is your sex? [Female/Male]

Are you employed? [No/Yes, Part time/Yes, Full time]

Residency status in Australia? [Australian citizen/Australian permanent resident/New Zealand citizen/ New Zealand permanent resident/Student visa/ Aboriginal or Torres Island/Other]

How long have you been in Australia (in years)? [positive integer]

Class? [First year/Second year/Third Year/Fourth Year/Honours/Masters/PhD]

Major: [Economics/Other Business/Psychology/Sciences/Other]

How many Economics classes have you taken at the university level? [None/One/Two/Three/Four/Five/Six/More than Six]

What major political party do you identify with? [Liberal-National/Labor/Australian Green/Neither]

How often do you attend religious services? [Never/Seldom/Often/Always]

### Attitudinal questions

Each question was answered on a 10-point Likert scale from not true to very true.

You should always obey laws, even if you are unlikely to get caught.

You should never try to get even.

You should always declare everything at customs.

You should never drive faster than the speed limit.

You should never take things that don’t belong to you.

You should never copy material and turn it in as your own work.

You should never do less than your share of work in a group project.

If you receive too much change from a salesperson, you should tell him or her.

You should never take sick leave from school unless you are actually sick.

If you damage a library book or a store’s merchandise, you should report it.

When you hear people talking privately, you should avoid listening.

You should never drop litter on the street.

You should never cheat on an exam.

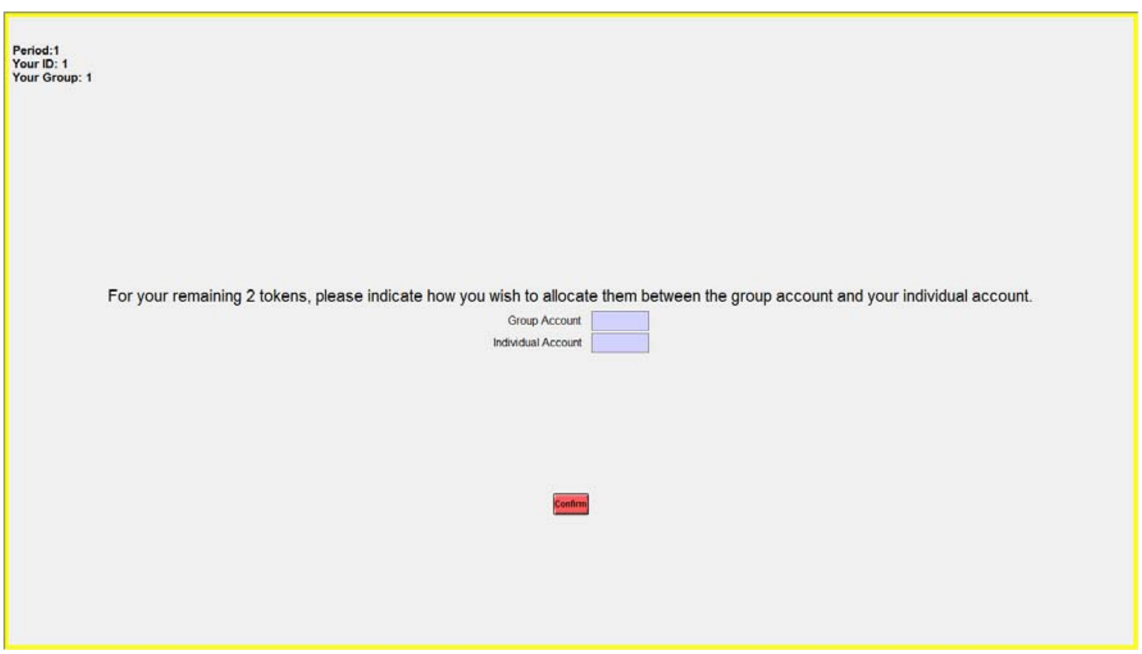
You should never help anyone cheat on an exam.

You should never lie.

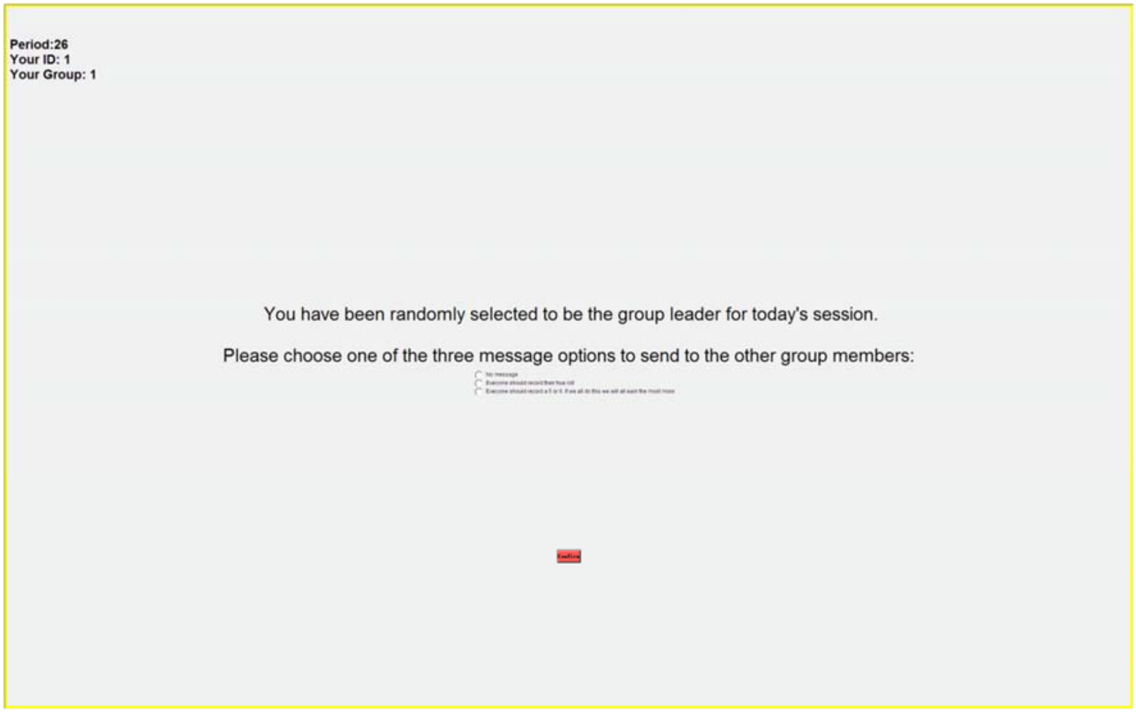
## Appendix C: Screenshots

Die-roll contribution screen (all treatments):


Top-up contribution screen (all treatments):

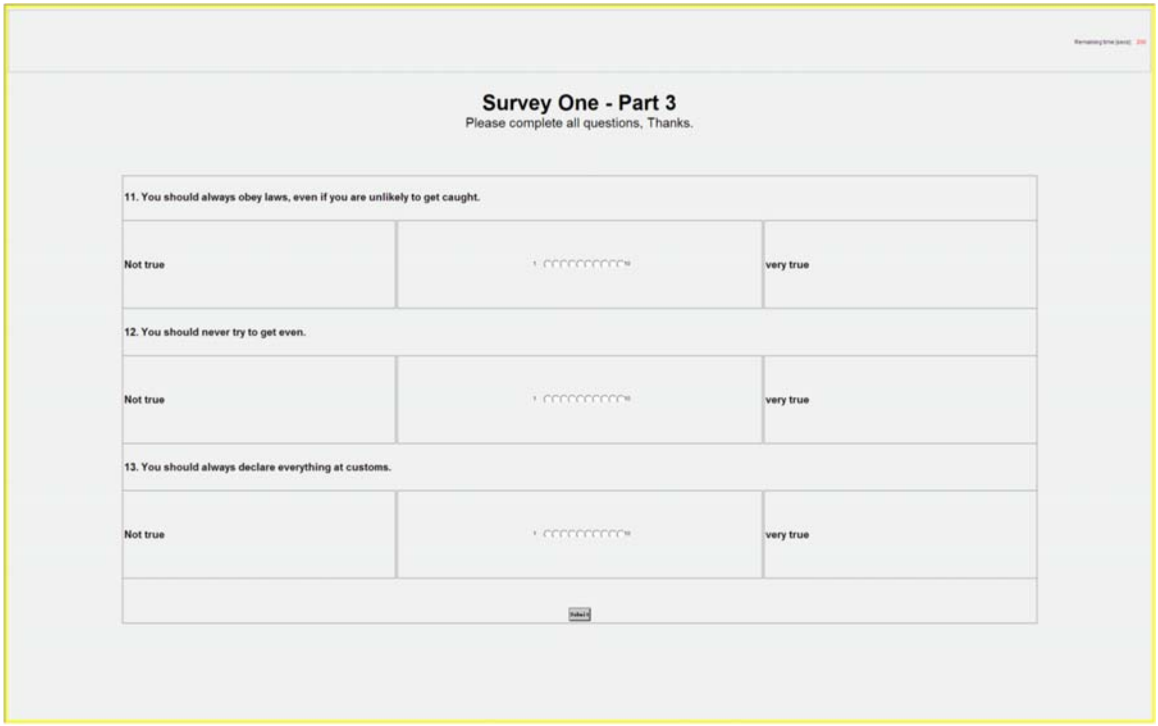


Leader message choice screen (R treatment):




Follower information screen (R/G leader treatments):

Questionnaire screen:



## Appendix D: Subject characteristics by treatment

| Characteristic | Treatment N | Treatment R | Treatment G | Test statistic p-value |
|---|---|---|---|---|
| Age | 23.8<br>(0.5) | 22.4<br>(0.6) | 21.1<br>(0.3) | *0.0003*[a] |
| Male | 0.56 | 0.59 | 0.67 | 0.47[b] |
| Employed | 0.24 | 0.31 | 0.39 | 0.21[b] |
| Economics major | 0.06 | 0.07 | 0.07 | 0.97[b] |
| Economics classes | 1.3<br>(0.3) | 1.6<br>(0.3) | 1.3<br>(0.3) | 0.67[a] |
| *Attend religious services* | | | | |
| Never | 27 | 16 | 23 | |
| Seldom | 25 | 23 | 21 | 0.21[a] |
| Often | 5 | 9 | 5 | |
| Always | 6 | 6 | 5 | |
| Pro-social (average response from attitudinal questions) | 8.0<br>(0.2) | 8.0<br>(0.2) | 7.9<br>(0.2) | 0.53[a] |
| *Political party* | | | | |
| Liberal/National (right) | 9 | 9 | 10 | |
| Labor (left-centre) | 6 | 8 | 6 | 0.85[a] |
| Greens (left) | 6 | 2 | 4 | |
| None of these | 42 | 35 | 34 | |
| No. of observations | 63 | 54 | 54 | |

a: Kruskal-Wallis test, b: chi-square contingency table test

Table D1: Subject characteristics by treatment (standard deviations in parentheses)

## Appendix E: Additional analysis

### E1 Dynamic analysis

Even though subjects were not given end-of-round feedback, it is still possible that their behaviour changed over time within each half of the experiment. Such changes could be due to continued introspection, or to reciprocal behaviour (e.g., "conditional cooperation) combined with beliefs that other subjects' contributions changed over time.

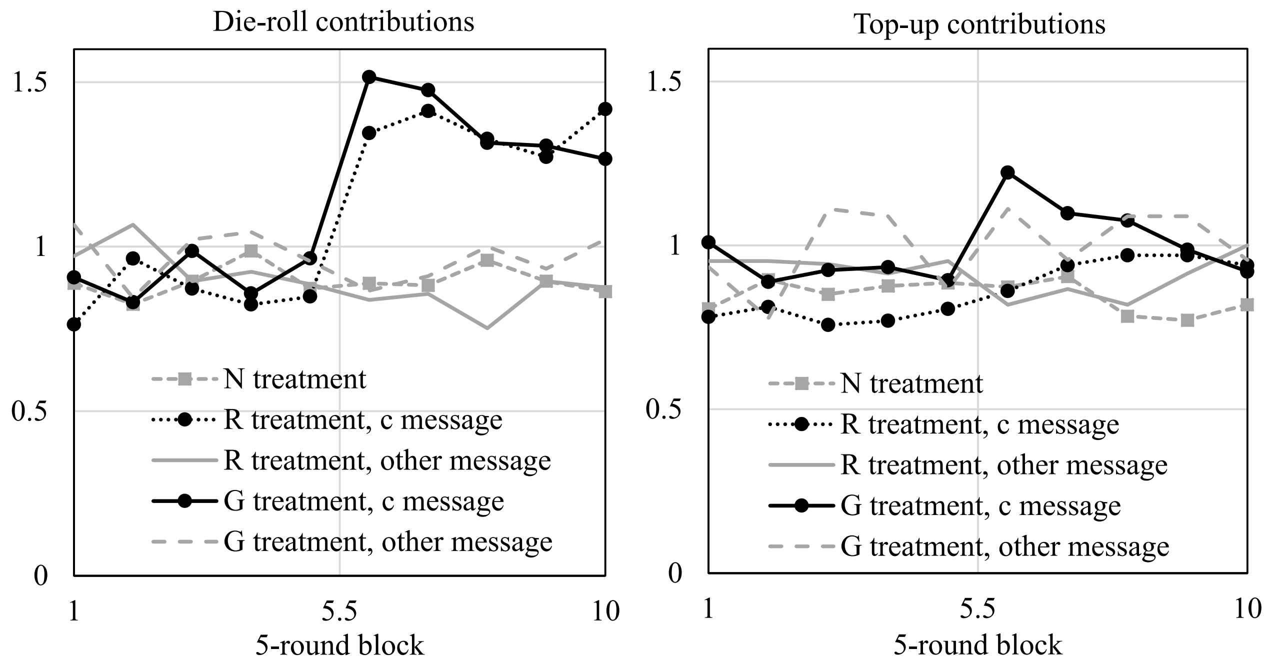



Figure E1: Die-roll and top-up contributions by 5-round block, treatment and leader's message

Figure E1 shows how contributions change over time. To reduce noise, the figure shows averages by 5-round blocks rather than every round. There are no systematic differences during the first half (blocks 1-5). In the sixth block, die-roll contributions rise by about half a token following a cooperate message in the R and G treatments, while remaining roughly the same in the N treatment and following any other message in the R and G treatments. There is some tendency for die-roll contributions to decline over time over blocks 6-10 following a cooperate message, but the difference from the other treatments persists until the end of the session. (Pooling the R and G treatments, die-roll contributions in the last 5-round block are significantly higher after a cooperate message than otherwise: two-tailed robust rank-order test, group-level data, $p \approx 0.009$). There are no apparent treatment effects in top-up contributions, and as with die-roll contributions, there is little overall time trend within either half of the session. This is in contrast to typical public good experiments, where contributions tend to decline substantially over time, and is very likely due to the lack of between-round feedback in our experiment.

Tables E1 and E2 present results of regressions similar to those in Tables 4 and 5 of the main text, but allowing for changes in behaviour over time. Rather than estimating Tobit models on the sample of subject-level averages as in Tables 4 and 5, here we use the (larger) sample of all individual subject contribution choices, and estimate panel Tobits. In addition to the variables in Tables 4 and 5, we include the round number, dummies for rounds 26 and 50 (the first and last rounds of the second half, to capture restart and endgame effects), the interactions of these three variables with our treatment dummies (and the leader dummy in Table E1 where it is present).

As in the main text, Table E1 presents second-half results for all subjects, while Table E2 focuses on followers in treatments R and G. In Table E1, the results for die-roll contributions are similar to those seen in Table 4. Second-half contributions are significantly higher in the R and G treatments compared to the baseline N treatment, and the differences are driven by the groups in which the leader sent a cooperate message. The effect of the round number is significantly negative, but treatment effects remain significant even in late rounds.[18] As in Table 4, there are no significant treatment effects on top-up contributions, suggesting that the gains in die-roll contributions are truly gains, rather than coming at the expense of top-up contributions.

[18] In round 49, estimated marginal effects for the R and G dummies in model E1 are +0.217* and +0.230**, with standard errors of 0.112 and 0.113 respectively, while the corresponding marginals for round 50 (where endgame effects may occur) are +0.384*** and +0.359*** with standard errors 0.145 and 0.145. The estimated marginal effects for the "R + c message" and "G + c message" dummies in model E2 are +0.470*** (standard error 0.110) and +0.243** (0.111) in round 49 and +0.580*** (0.140) and +0.338** (0.150).

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable (Tobit): | Die-roll contribution | | Top-up contribution | |
| R treatment | 0.194* | | −0.041 | |
| | (0.100) | | (0.103) | |
| R treatment + c message | | 0.440*** | | 0.023 |
| | | (0.096) | | (0.119) |
| R treatment + other message | | −0.229* | | −0.137 |
| | | (0.128) | | (0.129) |
| G treatment | 0.328*** | | 0.030 | |
| | (0.101) | | (0.106) | |
| G treatment + c message | | 0.384*** | | −0.007 |
| | | (0.095) | | (0.109) |
| G treatment + other message | | −0.057 | | 0.198 |
| | | (0.169) | | (0.204) |
| Leader | 0.218** | 0.226** | 0.165 | 0.167 |
| | (0.107) | (0.098) | (0.113) | (0.112) |
| Round number | −0.003* | −0.003** | −0.005*** | −0.005*** |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| Die-roll contribution (1$^{st}$-half avg.) | 0.053 | 0.029 | 0.235* | 0.251* |
| | (0.131) | (0.121) | (0.132) | (0.132) |
| Top-up contribution (1st-half avg.) | 0.513*** | 0.522*** | 1.019*** | 1.021*** |
| | (0.065) | (0.059) | (0.068) | (0.068) |
| Demographics? | Yes | Yes | Yes | Yes |
| Round-26 and round-50 dummies, interactions with treatments? | Yes | Yes | Yes | Yes |
| Observations | 4275 | 4275 | 4275 | 4275 |
| \|ln(L)\| | 4198.08 | 4174.80 | 4266.50 | 4361.66 |

Table E1: Factors affecting contributions (all subjects) – average marginal effects and standard errors (clustered by group)

In Table E2, the results for die-roll contributions are mostly similar to those seen in Table 5. Followers' second-half contributions are significantly higher in the R and G treatments when a cooperate message was sent by the leader, and the effect is still significant in late rounds (marginal effects of 0.505*** and 0.377* in rounds 49 and 50 respectively, with standard errors of 0.162 and 0.211). The effect of the round number is negative as in Table E1, but typically insignificant. The results for top-up contributions are the only place where allowing for time dependence affects the results; unlike in model 8 of Table 5, in model E8 the marginal effect of a cooperate message is insignificant (though still positive), suggesting that the significant positive effect seen in Table 5 is largely transitory. However, once again it is important to emphasise that even if this effect is not significant, it is certainly not negative, implying that the gains in die-roll contributions from cooperate messages are not crowding out top-up contributions.

| | (E5) | (E6) | (E7) | (E8) |
|---|---|---|---|---|
| Dependent variable (Tobit): | Die-roll contrib. (2$^{nd}$-half avg.) | | Top-up contrib. (2$^{nd}$-half avg.) | |
| Group-oriented leader | 0.058 | −0.110 | 0.107 | 0.169 |
| | (0.147) | (0.138) | (0.115) | (0.142) |
| c message | | 0.610*** | | 0.226 |
| | | (0.153) | | (0.165) |
| Leader die-roll contribution | −0.062 | −0.141 | 0.439* | 0.534 |
| (1$^{st}$-half avg.) | (0.282) | (0.334) | (0.225) | (0.335) |
| Round number | −0.003 | −0.003 | −0.003 | −0.004* |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Die-roll contribution | 0.022 | −0.266 | −0.067 | −0.073 |
| (1$^{st}$-half avg.) | (0.206) | (0.203) | (0.162) | (0.206) |
| Top-up contribution | 0.594*** | 0.523*** | 0.876*** | 0.972*** |
| (1$^{st}$-half avg.) | (0.104) | (0.101) | (0.061) | (0.111) |
| Demographics? | Yes | Yes | Yes | Yes |
| Observations | 1800 | 1800 | 1800 | 1800 |
| \|ln(L)\| | 1512.75 | 1499.80 | 1300.80 | 1624.05 |

Table E2: Factors affecting contributions (followers in leader treatments) – average marginal effects and standard errors (clustered by group)

## E2 A note on payoffs

Because the public good in our setting is linear, contributions and money earnings at the group level are closely related. From (1) in the main text, if we define $X_t$ and $\Pi_t$ as the group-level total contribution and money earnings in the t-th round, we have

$$\begin{aligned}
\Pi_t &= \sum_{i=1}^{3} \pi_{it} = \sum_{i=1}^{3} 0.08\left[ (4 - x_{it}) + \frac{1}{3}\sum_{j=1}^{3}\left(2x_{jt}\right)\right] \\
&= 0.08\left[ (12 - \sum_{i=1}^{3} x_{it}) + \frac{1}{3}\sum_{i=1}^{3}\sum_{j=1}^{3}\left(2x_{jt}\right)\right] \\
&= 0.08\left[12 + X_t\right]
\end{aligned}$$

so that group average profit is one-third of this. Hence, both total and average profit per round are affine functions of total per-round contributions at the group level. Also, the same relationship must hold for any superset of individual groups, such as 25-round group-level data, and treatment-level data. Thus, the effects we have seen at the aggregate level for contributions are preserved when we talk about earnings instead, as are significance results.

| Treatment | Earnings ($) | |
|---|---|---|
| | Rounds 1-25 | Rounds 26-50 |
| No leader | 11.81 | 11.68 |
| Random leader, c message | 11.40 | 13.88 |
| Random leader, h/blank message | 12.42 | 11.78 |
| Group-oriented leader, c message | 11.97 | 14.35 |
| Group-oriented leader, h/blank message | 12.41 | 12.44 |

Table E3: Per-person total earnings

Table E3 shows 25-round average subject earnings according to whether a cooperate message was sent, disaggregated as usual by treatment and 25-round block. Earnings in the first half of the session are similar across treatments and messages, ranging from the lowest average to the highest by about one dollar. By contrast, earnings in the second half vary more across treatments, with a cooperate message associated with about two dollars' additional earnings compared to other messages – holding the treatment constant – despite first-half earnings actually having been slightly lower in those groups where cooperative messages were sent.