How does the effect of pre-play suggestions vary with group size?
Experimental evidence from a threshold public-good game

Nick Feltovich[1]
Department of Economics, Monash University
Clayton, VIC 3800, Australia
nicholas.feltovich@monash.edu

Philip J. Grossman
Department of Economics, Monash University
Clayton, VIC 3800, Australia
philip.grossman@monash.edu

Abstract:
Numerous studies have examined factors influencing the likelihood of cooperative outcomes in nonzero-sum games, but there has been little study of the *interaction* between two of the most important: group size and pre-play cheap talk. We report results from an experiment in which groups of size between 2 and 15 play a one-shot multi-player threshold public-good game. In our *random leader* treatment, all group members select a suggestion (e.g., "Everyone should choose X"), with one randomly chosen to be broadcast to the group. In a *choice only* treatment, subjects choose suggestions but none is sent, and in a *baseline* treatment, there are no suggestions at all. We find a negative interaction between group size and this kind of communication: the beneficial effect of both suggestions overall and cooperative suggestions on cooperation, cooperative outcomes, and payoffs decreases sharply as the group size increases. We find a similar negative interaction in a follow-up treatment in which *all* group members' suggestions are broadcast to the group. Our results suggest that care should be taken in generalising conclusions from small-group experiments to large groups.

Keywords: pre-play communication; suggestions; coordination; cooperation; group decision making; equilibrium selection; leadership

JEL codes: C72, D80, C92

[1] Corresponding author. We gratefully acknowledge support from Monash University's economics department and from the Australian Research Council (DP130101695). We thank Brit Grosskopf, Mana Komai, participants at several conferences and seminars, and two anonymous referees for helpful suggestions and comments.

1. Introduction

Some of the most fundamental questions of economics involve the emergence of cooperative behaviour – actions that make others better off, usually with at least a chance of incurring a cost to oneself – in groups. Nearly all research into settings such as the Prisoners' Dilemma and linear public good games focuses on when and why people cooperate when standard theory predicts they will not. Studies involving a wide range of other environments (e.g., Stag Hunt, Chicken, minimum- and median-effort games, threshold public good games, or infinitely-repeated Prisoners' Dilemma games) are dominated by the question of why people cooperate in situations where standard theory may allow for cooperation, but allows equally for non-cooperative behaviour. A key component of all of this research consists of identifying factors that affect the likelihood and extent of cooperation. Two of the most important such factors are the size of the group and the presence of opportunities for communication among group members. A substantial body of experimental work, utilising a variety of strategic settings, has found that allowing communication tends to positively affect cooperation, while increasing group size tends to have a mixed effect. (See Section 2 for a review of this literature.)

While these two individual effects are well studied, there has been almost no systematic examination of the ***interaction*** between group size and communication. Understanding the nature of this interaction is manifestly important. Academic interest in this question was exemplified by John Ledyard in his chapter of the *Handbook of Experimental Economics*:

> "We see that communication increases contributions … with small (N < 15) groups. We do not know why. We also do not know what would happen in large groups."
> (Ledyard, 1995, p. 158)

Ledyard was speaking about public-good games, but his remark arguably holds true more generally. Beyond the academic interest, the relationship between group size and the effect of communication matters for reasons of external validity. The vast majority of experimental tests of communication involve small groups, with sizes typically between 2 and 6, and anything more than 10 quite unusual. In contrast, many of the real-life situations these experiments are meant to describe concern much larger groups; a collective-action problem like global climate change has at least 200 or so decision makers (if decisions are viewed as being made at the national level, more if not), coordination within a large firm can involve thousands of decision makers, and even a small group like an economics department will contain 20 or so. If the effect of communication turns out to vary substantially based on the size of the group, results from experiments with small groups may have little implication for real-world settings with larger groups.

The goal of this paper is to take a step toward understanding the interaction between group size and communication, utilising an experiment where we orthogonally vary both of these. The game we use is a multi-player threshold public-good game, where all players face a zero-one contribution decision, and the public good is provided if and only if everyone contributes. The threshold public-good game is useful to study as it models any situation where the group outcome depends on the effort of the least productive group member (Camerer and Knez, 1997; Weber et al., 2001). These "weak link" situations are most commonly found in organisations, but can also be seen in settings such as team sports, finance (e.g., bank runs), cyber-security, and international military alliances. The group-optimal outcome occurs when everyone contributes, but contribution is a best-response only when every other player contributes; otherwise, not contributing is optimal. (Thus, the game is also a multi-player Stag Hunt.) The close alignment of players' incentives in this game – and the likely difficulty of successful coordination when communication is not possible – mean that communication ought to be particularly effective here, relative to other frequently-studied games such as the Prisoners' Dilemma.

Following most but not all previous studies, the form of communication we examine is ***pre-play cheap talk***: costless, non-binding communication sent and received prior to the choice of actions.[2] Cheap talk communication generally takes one of three forms: unstructured communication (Isaac and Walker, 1988), a statement of the sender's (truthful or not) intended action choice (Cooper et al., 1992), or the sender's suggestion for the group members' action choices (Weber et al., 2001). Our experiment uses this last form of communication. Allowing communication to take the form of a suggested action for the group (a) may avoid the psychological or behavioural-economics implications of "lying", and (b) provides the sender an air of leadership, which might afford the received message greater authority than announcement of an intended action choice.[3] In many of the situations modelled by threshold public-good games (see previous paragraph), the group would typically be "managed" by a group leader suggesting how the group could improve performance.[4]

Groups of size 2, 3, 4, 5, 7 and 15 play this game a single time, under one of three message treatments. In our "random leader" treatment, each group member chooses a message to be sent to all group members, and exactly ***one*** group member's message is randomly chosen to be broadcast to the entire group prior to the choice of actions. (As noted above, messages are framed as suggestions to the group, and are cheap talk.) Two other treatments are controls. Our "choice only" treatment is designed to be as similar as possible to the random leader treatment, but without the communication itself. Specifically, subjects in this treatment still choose messages, but none is actually broadcast to the group, and actions are chosen based on no information beyond the game's strategic structure. A "baseline" treatment has no messages at all, so again, actions are chosen based only on the game's structure. In principle, our use of two control treatments should allow two distinct comparisons (leader vs. choice, leader vs. baseline) to assess the effect of allowing communication; however it turned out in the experiment that behaviour in those two controls was indistinguishable.

Our main result is a clear negative interaction between group size and communication. While overall we observe the standard positive effect of communication on cooperative choice frequencies, cooperative outcomes, and average payoffs, this effect is most pronounced in the smallest groups, declining significantly as the group size increases, and becoming statistically indistinguishable from zero for groups as small as 7 and negligible in our largest groups (size 15). This interaction is in addition to the negative effect of group size seen in the control treatments. Even in the best case – where group members receive a cooperative message – the beneficial effects of communication also decrease significantly as the group size rises. These decreases are not just statistically significant, they are substantial: on the order of a 40 percent decrease in individual cooperation, a 90 percent decrease for cooperative outcomes, and a decrease in payoffs of over 100 percent (i.e., the effect of cooperative messages on payoffs becomes negative at the highest group size).

---

[2] See Masclet et al. (2003) for an example of how ***post-play*** messages can improve outcomes.

[3] Weber et al. (2001) report that the leader's suggestion increases the number of both the least cooperative and the most cooperative responses. The leader's suggestion came after subjects had already played a number of rounds of the game.

[4] While our study contributes to the literature on communication, group size, and cooperation, it also speaks to the question of what constitutes a leader. Economics typically adopts a "minimal leader" paradigm, stripping leadership down to its most basic elements. Leadership style and personal and socioeconomic characteristics are assumed to be irrelevant. Leaders are typically assumed to be just average players whose authority arises merely by occupying the leadership position. Their legitimacy is derived from superior information about the group project (Hermalin, 1998; Vesterlund, 2003; Potters et al. 2005; Komai et al. 2007; and Komai and Stegeman, 2010). We have stripped down even further the definition of a leader; our leader has no superior information – only greater ability to make her voice heard by others. In our experiment, the "leader" is chosen randomly and does not differ from the other group members in terms of seniority, expertise, political skills, etc. This is admittedly a stylised notion of leadership, but is in the spirit of settings where leaders are chosen randomly (e.g., the ancient Athenian Boule), rotate through the eligible group members (e.g., the presidency of the Council of the European Union), or are chosen based on attributes unconnected to the issue at hand (perhaps, the non-permanent members of the United Nations Security Council). Subjects, when choosing their messages, do not know who will be chosen as their leader, or even whether anyone will, so that leaders' messages should not differ systematically from other group members' messages. The leader's authority comes from having her message chosen to be relayed to other members of the group.

While our random leader treatment has some clear advantages as an implementation of communication (most importantly, sending a single unequivocal signal to all group members), it also has disadvantages; for example, as the group size increases, a single message provides information about a smaller fraction of the group. To address this issue, we conduct an additional experimental treatment. In this "full communication" treatment, group members receive ***all*** of the messages of their fellow group members, so allowing the absolute amount of communication to increase in a natural way with the group size. However, this type of communication performs no better, and in some ways worse: either it exhibits the same decreasing effectiveness as the group size increases as in our random leader treatment, or it is ineffective even for small group sizes (in which case there is no systematic group-size effect). Our results therefore highlight the need for caution when extrapolating treatment effects (particularly for communication) observed in small-group experiments to larger groups.

2. Literature review

There are sizeable literatures on both cheap talk and the impact of group size in strategic situations; a full review is well beyond the scope of this paper. We limit our discussion to those papers most closely relevant to our research question.

2.1 Separate effects of group size or communication

Theoretical approaches to the efficacy of cheap talk have yielded mixed conclusions. Aumann (1990) argues that since non-cooperative games lack an external mechanism to enforce agreements between players, they will act on the basis of the incentives faced. Since cheap talk – by definition – does not change these incentives, it should therefore have no effect on the game's outcome. Farrell (1995) proposes an alternative view. If players hold heterogeneous priors on other players' likely play, then cheap talk by one player can prompt others to play strategies leading to a mutually beneficial outcome. Knowing this, the message sender is more likely to play consistently with the mutually beneficial outcome, thus resulting in cheap talk improving outcomes. Along the same lines, Crawford (1998) contends that cheap talk can affect behaviour, with the nature of its effects depending on the game being played.[5] In Battle-of-the-Sexes, a symmetric game whose pure-strategy equilibria are asymmetric, communication plays a role in "symmetry breaking", facilitating the play of different pure strategies by players in the same position. In the Stag Hunt game, by contrast, with pure-strategy equilibria that are symmetric but can be Pareto ranked, communication plays a "reassurance" role, allowing players to signal their understanding of the strategic environment and the alignment of their interests.

Aumann (1990) and Farrell and Rabin (1996) define two conditions that must hold if cheap talk is to facilitate coordination: *self-commitment* and *self-signalling*. *Self-commitment* requires that a sender's message, if believed (so that the receiver chooses a best-response), incentivises the sender to play the signalled action; that is, the message is part of a Nash equilibrium. *Self-signalling* requires the sender to prefer that the receiver play the best response to the signalled action if the sender actually intends to play the signalled action, but not otherwise. Farrell and Rabin (1996) argue that a signal that satisfies both conditions is highly credible and should be truthful and believed. In the threshold public-good game we use, both contribute and not-contribute messages are strictly self-committing and weakly self-signalling (see Figure 1 below). By way of contrast, cooperative messages in social dilemmas such as Prisoners' Dilemma and many public-good games are neither, suggesting that messages should be more credible in our game than in these other games.

[5] Ellingsen and Östling (2010) reach much the same conclusion using the level-k model of strategic thinking and assuming that indifferent players send truthful messages (or messages that reflect their intentions at the time the messages are sent). As they note, their model has little implication for settings like ours, where messages have the form "We should all choose X" rather than "I will choose X".

The impact of cheap talk on behaviour has been studied experimentally in various games. Many of these studies, including most of the earliest ones, involved social dilemmas such as the two-player Prisoners' Dilemma (e.g., Loomis, 1959; Deutsch, 1960; Swensson, 1967), multi-player versions of the PD (e.g., Jerdee and Rosen, 1974; Dawes et al., 1977), public-good games using a voluntary-contribution mechanism (VCM) (Isaac and Walker, 1988; Palfrey and Rosenthal, 1991; Brosig et al., 2003; Pogrebna et al., 2011; Koukoumelis et al., 2012; Oprea etal., 2014; Jack and Recalde, 2015), and trust games (Charness and Dufwenberg, 2006; Ben-Ner et al., 2011). In this literature, communication is usually found to increase cooperative behaviour, despite these messages being neither self-signalling nor self-committing, though the efficacy of communication depends partly on the form it takes (see Ledyard's (1995) survey, meta-analyses by Sally (1995) and Balliet (2010), and a controlled experiment by Bochet et al. (2006)).

There have been fewer experimental studies of communication in coordination games, but enough that some regularities have emerged. One-way communication (all group members receive the same message, sent by a single group member, as in our experiment) seems to be especially effective at increasing coordination and payoffs in games where the pure Nash equilibria cannot be Pareto-ranked (Cooper et al., 1989; Duffy and Feltovich, 2002), and performs well but not perfectly in games with Pareto-ranked pure-strategy Nash equilibria (Cooper et al., 1992; Charness, 2000; Weber et al., 2001; Duffy and Feltovich, 2002; Charness and Grosskopf, 2004). Multi-way communication (group members receive differing or multiple messages, with multiple message senders) performs very well when it happens that all of the messages carry the same implication, but are often little better than no communication at all otherwise (Cooper et al., 1989, 1992), though its effect sometimes depends on the game being played (e.g., Clark et al., 2001 who found communication's effectiveness was sensitive to payoffs and Blume and Ortmann, 2007 who found that multi-way communication did well in a minimum-effort game but not in a median-effort game).[6]

The effect of group size in games has also received some attention. In linear public-good games, individual studies have tended to find either no significant group-size effect or a positive effect when holding constant the marginal per-capita return to investment (Isaac et al., 1994; Goeree et al., 2002; Carpenter, 2007; Weimann et al., 2012; Nosenzo et al., 2015).[7] By contrast, systematic reviews and meta-analyses of social dilemmas have tended to find either a negative effect (Sally, 1995; Kollock, 1998) or no systematic effect (Ledyard, 1995). In coordination games, the effect of group size seems also to be either negative (Van Huyck et al., 1990) or insignificant (Heinemann et al., 2009).

2.2 Interaction between group size and communication

There has been little previous work jointly investigating group size and cheap talk, either theoretically or empirically.[8] We know of no theoretical work looking specifically at this interaction, and theoretical treatments of communication generally, such as Crawford's (1998) notion of "reassurance" and Farrell and Rabin's (1996) "self-commitment" and "self-signalling", arguably should apply equally irrespective of the group size. Thus, any differential effect of communication would seem to be an empirical question.

[6] Brandts et al. (2011) find that communication from a leader, elected more so than randomly appointed, improves group outcomes in a turnaround game.
[7] See Nosenzo et al. for a very nice review of the extant literature on this topic.
[8] There has been some study of ***costly*** signalling (Pogrebna et al., 2011 and Dannenberg, 2015) and of the interaction between group size and costly signalling (Komai and Grossman, 2009). Pogrebna et al. find similar effects in a VCM game for both costless and costly signalling, with both improving cooperation; Dannenberg finds a greater impact for costly signals. Komai and Grossman find that costly signals from a randomly selected leader are significantly more effective in inducing cooperation in groups of 3 than in groups of 9.

Isaac and Walker (1988) reported results from several VCM public-good experiments, including one with 4-person groups and another with 8-person groups (their "Design I" and "Design III" respectively), both with either no communication or free-form communication. However, they are careful to point out that other aspects of the environment were also changed between these two designs (e.g., in Design III the group optimum involved less than total contribution to the public good, while 100% contribution was group-optimal in Design I), making it difficult to discern the effect of the group size alone. Indeed, Isaac and Walker emphasised that they were not attempting "to separate out the relative importance of [group size and returns to the public good]" (p. 600).

Weber et al. (2001) look at "leadership" in groups of either 2 or 10 playing a 10-round minimum-effort game, where after the second round, the leader is randomly chosen to make a cheap-talk speech to the other subjects in the session, extolling the benefits of cooperation. Weber et al.'s results (see their Table 4) do suggest a negative interaction between group size and cheap talk, though they do not test this, and since the variation in communication is within-subject and not counter-balanced, it is not possible to distinguish between the effect of communication and that of other factors such as adaptation.[9] Indeed, Weber et al. note that their results are "consistent with the idea that subjects [initially] fail to anticipate the group-size effects" (p. 585), suggesting a learning explanation.

The closest to a systematic study of the interaction between group size and communication is Aflagah et al.'s (2014) study of the interaction between group size and N-way communication for two, relatively, large groups (10 and 20) of Senegalese farmers. They find no direct effect of communication on cooperation, but when interacted with group size it has a positive effect, but only for the larger group. One important difference between the Aflagah et al. experiment and our experiment is the nature of the message. Subjects in the Aflagah et al. experiment all send a message concerning their own cooperative intentions (i.e., how many tokens they intend to invest in the group); in our experiment, one subject is selected to send a more general message ("Everyone should choose [action]").

Finally, Balliet (2010), in his meta-analysis of 45 social-dilemma experiments, reports a ***positive*** interaction between group size and the effect of communication. However, this interaction becomes negligible and statistically insignificant after controlling for other relevant variables (whether communication is written or face-to-face, whether it happens once or before each round, and whether the game is played once or multiple times), and actually becomes negative (though still insignificant) if two outliers are removed.[10]

3. Theory

The game – a symmetric n-player threshold public-good game – is shown in Figure 1. Each player has two available actions, contribute (C) or not (N). N earns a sure payoff of 20, while the payoff from C is 30 if all of the other (n-1) players choose C, versus 5 if at least one of the others chooses N. For any n, there is a pure-strategy Nash equilibrium where all players choose C and one where all choose N; these are the payoff- and risk-dominant equilibria respectively. There is also a symmetric mixed-strategy equilibrium where each player chooses C with probability $(0.6)^{1/(n-1)}$ and earns an expected payoff of 20.

[9] The same is true for Fonseca and Normann's (2012) Bertrand oligopoly experiment with varying numbers of firms. Like Weber et al., they vary communication within-subject but without counter-balancing, so that their results (which also suggest a treatment effect that decreases with the group size) cannot distinguish between the effects of communication and those of other factors like learning. Also like Weber et al., they do not statistically test the apparent decreasing interaction effect between communication and group size.

[10] Our calculations, based on Balliet's data (see his Table 2), are available upon request.

		Rest of group choices	
		All C	***At least one N***
Player choice	***C[ontribute]***	30	5
	N[ot contribute]	20	20

Figure 1: Own payoffs in the symmetric n-player threshold public good game

In all treatments except our baseline, each player also chooses a message that might be broadcast to the entire group prior to the action choice. The available messages are "Everyone should choose C" and "Everyone should choose N"; it was also possible to send neither (we will call this a "blank" message).[11] Message choices (including the blank message) are cheap talk in the game-theoretic sense: none involves a monetary cost, nor do any of them constrain the subsequent action choice.[12] As noted in Section 2, theoretical treatments of communication in games tend to view this game as one where communication should work relatively well (Aumann 1990; Farrell and Rabin 1996; Crawford 1998), though they are mainly silent about whether this should vary with the group size.

4. The experiment

The main experiment has a 6x3 factorial design (though in Section 6, we discuss an additional treatment that makes this a 6x4 fractional factorial design). The two treatment variables are (1) the number of subjects in the group (n = 2, 3, 4, 5, 7 or 15), and (2) whether messages are sent and received. Our baseline ("Base") treatment is a standard n-player threshold public-good game with no messages. In our random leader ("Leader") treatment, all subjects have the opportunity to choose a message for the entire group after being informed of the group size; one randomly-chosen group member's message is actually broadcast. Our choice-only ("Choice") treatment is like the Leader treatment except no message is broadcast. The Base and Choice treatments thus represent two controls for assessing the effect of messages on subsequent actions. Comparison of the Leader treatment with the Base treatment identifies the total effect of all aspects of message sending and receiving, including the introspection involved in message sending, any potential "ownership" of an action a subjects chooses to send (even if not broadcast), and the increase in complexity of the decision-making environment, as well as the coordination-enhancing nature of all group members' receiving a common message or set of messages. By contrast, comparison with the Choice treatment isolates the effect of the message received.

4.1 Hypotheses

The three null hypotheses we will test are as follows.

Hypothesis 1: *Behaviour is not affected by whether a message is transmitted to group members, or by the content of any such message.*

Hypothesis 2: *Behaviour is not affected by the group size.*

Hypothesis 3: *The effect of transmitting messages, or any particular type of message, does not differ according to group size.*

[11] We call these "blank messages" rather than non-messages, since game-theoretic models of cheap talk allow these to have informational value, just as C or N messages can.

[12] See Section 2.1 for a discussion of theoretical predictions involving cheap talk.

For Hypothesis 1, the corresponding alternative hypothesis is directional; our prior is that if there actually is a treatment effect, it is that cooperative (C) choices and cooperative (payoff-dominant) outcomes are *more frequent* with than without communication – both in general and after a cooperative message is broadcast. Similarly, although in general the literature is split regarding the effect of communication in general, theory suggests a group-size effect in the particular game we consider. Many equilibrium-selection criteria – including risk-dominance, Harsanyi and Selten's (1988) criterion, and many dynamic theories – are based on the relative sizes of the pure-strategy equilibria's basins of attraction, with the boundary between them determined by the mixed-strategy equilibrium. Because the mixed-strategy equilibrium probability of C *increases* with the group size, the basin of attraction of the all-N equilibrium grows at the expense of the all-C equilibrium. This implies that C choices should *decrease* as the group size increases.

By contrast, there is no directional alternative hypothesis corresponding to Hypothesis 3, since as already noted, theory is either silent about any interaction between cheap talk and group size, or implies no interaction.

4.2 Experimental procedures

The experiment took place at MonLEE, Monash University's experimental economics lab. Subjects were primarily undergraduate students from Monash University, recruited using ORSEE (Greiner, 2015). At the beginning of a session, subjects were seated at partitioned work-stations in a single room and given written instructions for Part 1.[13] These instructions were also read aloud in an attempt to make the experimental environment common knowledge, and any questions were answered privately before Part 1 began. The experiment was programmed in z-Tree (Fischbacher, 2007). Subjects were visually isolated during the session, and asked not to communicate with each other except via the computer program.

Table 1: Session information

Treatment	# of sessions	# of subjects	# of groups of size					
			2	3	4	5	7	15
Base	10	166	7	8	5	4	4	4
Choice	14	198	12	8	8	6	4	4
Leader	17	258	14	9	9	7	6	6

At the beginning of each session, the group sizes were chosen by the experimenter, based on the number of subjects showing up and with an eye toward equalising the numbers of observations with each group size across our treatments. There were always at least two groups in every session, so that no subject could infer the choice made by any other individual subject. The experimenter also decided in advance whether a given session would be part of the Base treatment (which used a slightly different set of instructions) or either the Choice or Leader treatment. If the session was part of the Choice or Leader treatment, the decision between these was made by the computer program (rather than the experimenter) immediately before play began, and thus after group sizes were determined. Since we anticipated that not all subjects would send C messages, we set the probability

[13] The complete instructions are shown in Appendix A; the questionnaire from Part 3 is in Appendix B; and some sample screen-shots are in Appendix C. Other experimental materials and the raw data are available from the corresponding author upon request. As noted in the instructions, subjects were not given specific information about future parts of the experimental session in advance, nor were they given feedback about other group members' choices in the public-good game or nature's move in the lottery-choice task until the end of the experimental session.

of the Leader treatment at 0.6 in order to roughly equalise the number of observations in the Choice treatment with that in the Leader treatment conditional on a C message being broadcast. The resulting number of sessions in each treatment and the total number of groups of each size are shown in Table 1.

Each session was divided into four parts, with the instructions for each part being distributed immediately before that part began. In Part 1, subjects played the threshold public-good game. Each subject was first informed of the size of her group and, except in the Base treatment, prompted to choose a message to send. The messages were shown on buttons labelled "Everyone should choose [action]", and there was also a button labelled "No message". The contribute and not-contribute actions were given abstract labels ("Red" and "Green") in the instructions and on subjects' computer screens. After everyone had chosen a message (or no message), subjects in the Choice and Leader treatments were informed of whether a message from their group had been chosen to be transmitted, and if so, whether it was the subject's own message, and what the message was. After this, or at the beginning of the round in the Base treatment, subjects were prompted to choose their actions. After everyone had chosen an action, the session proceeded to Part 2 (no feedback was given until the end of the session).

Table 2: Lotteries available in Part 2 of the experiment (all payoffs in dollars)

	Lottery #				
	1	2	3	4	5
50% chance of	4	6	8	10	12
50% chance of	4	3	2	1	0

Parts 2-4 were identical in all treatments. Because our game involves a decision between a safe action and a risky one, one potential source of experimental noise is heterogeneity in subjects' risk attitudes. Our Parts 2 and 3 are attempts to control for such heterogeneity. Part 2 consisted of a lottery choice task adapted from Eckel and Grossman (2008). After the instructions for Part 2 were distributed and read by the experimenter, subjects were prompted to choose one of 5 lotteries, each with two equally-likely outcomes (see Table 2). The lotteries vary according to expected value and variance, from Lottery #1's certain gain of $4 to the riskiest but highest expected-value Lottery #5.

After all subjects made their lottery choices, the instructions for Parts 3 and 4 were distributed and read aloud. Part 3 was an unincentivised questionnaire. It begins with 20 questions, on four screens of five questions each, adapted from the DOSPERT (Weber et al., 2002), about subjects' willingness to take risks. Ten of these questions are from the Social Risk scale (e.g., "Speaking your mind about an unpopular issue at a social occasion"); five each are from the Investment (e.g., "Investing 10% of your annual income in a very speculative stock") and Gambling (e.g., "Taking a day's income to play the pokies [slot machines] at a casino") subscales of the Financial Risk scale. Subjects answer each question on a 6-point Likert scale, from 1="extremely unlikely to engage in" to 6="extremely likely to engage in". Once these questions are answered, subjects complete the demographic part of the questionnaire, containing questions about age, sex, number of economics classes taken, and portion of life spent in Australia, where the experiment took place.

After everyone had finished all parts of the questionnaire, the session proceeded directly to Part 4 of the experiment, which involved no decisions. Subjects were first shown the results of the threshold public-good game (action chosen, how many group members had chosen each action, payoff). After all had clicked a button to continue, they were shown the results of the lottery choice task, including the lottery chosen and payoff, along with the realisation of a random integer between 1 and 100 inclusive, drawn i.i.d. for each subject; subjects received the high payoff from their chosen lottery if the random number was between 1 and 50 inclusive, and the low payoff otherwise. The next screen showed subjects their total earnings, after which they were paid, privately and individually. Each subject received – in Australian dollars – the sum of the amounts earned in Parts 1 and 2, plus $5 for Part 3; there was no show-up fee.[14] Total earnings ranged from the minimum possible of $10 to the maximum possible of $47, and averaged roughly $26.30 ($15.90, $5.40 and $5.00 in Parts 1, 2 and 3 respectively); sessions typically lasted about 45 minutes.

5. Experimental results

As indicated by Table 1, a total of 622 subjects in 125 groups of various sizes took part in these three treatments.[15] Table 3 shows message choices in the Choice and Leader treatments (recall that subjects in these treatments received identical treatment, including the same instructions, up to the point where all messages had been chosen) and broadcast messages in the Leader treatment.

Table 3: Message choice frequencies

		Group size					
		2	3	4	5	7	15
All messages, Choice and Leader treatments (individual-level data)	C[ontribute]	0.83	0.67	0.57	0.65	0.60	0.69
	N[ot]	0.08	0.20	0.29	0.25	0.30	0.23
	blank	0.10	0.14	0.13	0.11	0.10	0.08
Messages received, Leader treatment (group-level data)	C[ontribute]	0.93	0.33	0.67	0.71	0.50	0.67
	N[ot]	0.00	0.44	0.22	0.14	0.50	0.17
	blank	0.07	0.22	0.11	0.14	0.00	0.17

Overall, out of a total of 456 opportunities to send messages in the Choice and Leader treatments, 47 subjects (about 10%) chose to send a blank message. About two-thirds of subjects (67%) sent C messages (thus representing about 74% of non-blank messages), and just over one-fifth sent N messages (23% overall, or 26% of non-blank messages). C messages were somewhat more likely in 2-person groups than for other group sizes, and a chi-square test weakly rejects the null hypothesis that C messages are equally likely to be sent for all group sizes (χ^2_5 =10.67, p ≈ 0.058).[16]

Transmitted messages in the Leader treatment were not perfectly representative of the population of chosen messages in the combined Choice and Leader treatments, with C messages over-represented in groups of 2 and under-represented in groups of 3. As with the chosen messages themselves, differences across group size in these transmitted messages are weakly significant (χ^2_5 = 9.64, p ≈ 0.086), and suggests that our analysis will need to include regressions with controls for the message(s) received (see the end of this section).

Table 4 shows how the frequency of cooperative (C) action choices varies by treatment and group size, and within the Leader treatment, according to whether or not the group received a C message

[14] At the beginning of the experiment, the US and Australian dollars were roughly at parity, but by the end, an Australian dollar was worth only about 0.78 USD.

[15] One pilot session with 11 subjects was conducted to test the program and ensure the instructions were understandable to the subjects. This session is not included in the analysis. An additional 265 subjects in 59 groups participated in the full-communication treatment discussed in Section 6.

[16] See Siegel and Castellan (1988) for descriptions of the non-parametric statistical tests used in this paper, and Feltovich (2005) for critical values of the robust rank-order test statistic used below.

(that is, we pool N and blank messages). Since we find no important differences in action choices between the Base and Choice treatments, we pool their results.[17] Subjects in these two treatments choose the cooperative action a bit less than half the time overall, with the frequency of contribution decreasing significantly as the group size increases (Jonckheere test, $z = -2.17$, $p \approx 0.015$). We can similarly reject the corresponding null hypothesis for the Leader treatment (Jonckheere test, $z = -2.45$, $p \approx 0.007$).

Table 4: Frequency of C action choices

		Group size					
		2	3	4	5	7	15
Base treatment		0.71	0.46	0.35	0.35	0.25	0.40
Choice treatment		0.50	0.42	0.41	0.47	0.43	0.38
Base and Choice (pooled)		0.58	0.44	0.38	0.42	0.34	0.39
Leader treatment	total	0.82	0.41	0.53	0.37	0.40	0.49
	N/blank message	0.00	0.11	0.25	0.00	0.00	0.30
	C message	0.88	1.00	0.67	0.52	0.81	0.58
Difference in percentage points and significance							
Leader treatment vs. pooled Base/Choice		+0.24	−0.03	+0.14	−0.05	+0.07	+0.10
		p ≈ 0.037	*p ≈ 0.80*	*p ≈ 0.18*	*p ≈ 0.65*	*p ≈ 0.26*	*p ≈ 0.16*
Leader treatment/C message vs. pooled Base/Choice		+0.31	+0.56	+0.28	+0.10	+0.47	+0.19
		p ≈ 0.009	*p ≈ 0.002*	*p ≈ 0.022*	*p ≈ 0.41*	*p < 0.001*	*p ≈ 0.015*

Note: p-values from Jonckheere tests using individual-level data.

Within the Leader treatment, there are also substantial differences in behaviour according to the content of the information received, with subjects cooperating significantly more often after receiving a C message than after receiving an N or blank message (chi-square test, $p \approx 0.018$ for n=4, $p \approx 0.011$ for n=15, $p < 0.01$ for the other group sizes). Even conditional on receiving a C message, however, cooperation decreases with the group size, and we can reject the null of no difference across groups (Jonckheere test, $z = -2.91$, $p \approx 0.002$).

The bottom two rows of Table 4 provide information about the effect of communication. For this, two comparisons are particularly important. Comparison between the Leader treatment and the pooled Base and Choice treatments sheds light on the unconditional effect of allowing one of the group members' messages to be broadcast. A similar comparison, but restricting consideration in the Leader treatment to those groups that had received a C message, yields the effect of receiving cooperative messages in particular. For both comparisons, the bottom portion of the table shows effects that are typically positive but trending downward (albeit noisily) with the group size. That is, while communication in general, and C messages specifically, lead to more cooperative action choices, their ability to increase cooperation decreases as the group size increases.

Table 5 shows how the frequency of the efficient all-C outcome varies by treatment, group size, and within the Leader treatment, the message (again, with N and blank messages pooled). There is an apparent group-size effect, with efficient outcomes becoming less common as the group size increases. The decrease with the group size is significant in the pooled Base and Choice treatments (Jonckheere test, $z = -1.93$, $p \approx 0.027$), in the Leader treatment overall ($z = -3.25$, $p < 0.001$), and in that treatment conditional on receiving a C message ($z = -3.14$, $p < 0.001$). The bottom portion of the

[17] Even without making any correction for multiple comparisons, chi-square tests fail to reject the null hypothesis that action choice frequencies are equal between the Base and Choice treatments for each individual group size ($p \approx 0.16$ for n=7, $p \approx 0.20$ for n=2, $p > 0.3$ otherwise).

table shows that, as with C action choices, either the Leader treatment or a C message in this treatment has a strong positive effect on all-C outcomes when groups are small, but this effect is smaller for larger group sizes.

Table 5: Frequency of efficient (all-C) outcome

		Group size					
		2	3	4	5	7	15
Base and Choice treatments (pooled)		0.37	0.06	0.00	0.00	0.00	0.00
Leader treatment	total	0.79	0.33	0.11	0.14	0.17	0.00
	no C message	0.00	0.00	0.00	0.00	0.00	0.00
	C message	0.85	1.00	0.17	0.20	0.33	0.00
Difference in percentage points and significance							
Leader treatment vs. pooled Base/Choice		+0.42 *p ≈ 0.017*	+0.27 *p ≈ 0.076*	+0.11 *p ≈ 0.22*	+0.14 *p ≈ 0.22*	+0.17 *p ≈ 0.23*	0.00
Leader treatment/C message vs. pooled Base/Choice		+0.48 *p ≈ 0.007*	+0.94 *p < 0.001*	+0.17 *p ≈ 0.13*	+0.20 *p ≈ 0.14*	+0.33 *p ≈ 0.087*	0.00

Note: p-values from Jonckheere tests using group-level data.

Table 6 shows (per-subject) game payoffs, again according to the group size, treatment, and within the Leader treatment, whether a C message was broadcast. The effects here are similar to those seen in the previous two tables. Payoffs decrease significantly with the group size (Jonckheere tests, $p \approx 0.046$ for pooled Base and Choice, $p < 0.001$ both for Leader overall and for Leader with a C message received) The benefit of communication also apparently decreases as the group size increases: from a significant increase for groups of 2 and 3, to no significant difference for groups of 4-7, to a significant ***decrease*** for groups of 15.

Table 6: Per-subject game payoffs in dollars

		Group size					
		2	3	4	5	7	15
Base and Choice treatments (pooled)		20.5	15.0	14.2	13.7	14.9	14.1
Leader treatment	total	27.3	22.2	14.9	18.0	18.1	12.7
	no C message	20.0	18.3	16.2	20.0	20.0	15.5
	C message	27.9	30.0	14.2	17.2	16.2	11.2
Difference in dollars and significance							
Leader treatment vs. pooled Base/Choice		+6.8 *p ≈ 0.012*	+7.2 *p ≈ 0.004*	+0.6 *p > 0.20*	+4.3 *p ≈ 0.14*	+3.2 *p > 0.20*	−1.5 *p ≈ 0.079*
Leader treatment/C message vs. pooled Base/Choice		+7.4 *p ≈ 0.009*	+15.0 *p ≈ 0.004*	−0.1 *p > 0.20*	+3.5 *p > 0.20*	+1.3 *p > 0.20*	−2.9 *p ≈ 0.004*

Note: p-values from Jonckheere tests using group-level data.

While Tables 4-6 provide suggestive evidence of a negative interaction between group size and communication, we next test for such interactions rigorously by estimating several regression models. Our three dependent variables are the fraction of C choices in the group (using a Tobit model), an indicator for an all-C outcome (probit), and payoff efficiency, defined as the mean payoff of the group, normalised so that the minimum and maximum possible payoffs map to 0 and 1 (Tobit). We capture communication in two ways: either as an indicator for the Leader treatment (which will yield the effect of allowing communication, without conditioning on what the message is) or as an indicator for the group receiving a C message (which will yield the effect of a cooperative message in the Leader treatment). For the former, we use all group-level observations in the Base, Choice and Leader treatments, while for the latter, we drop the observations in the Leader treatment in which a message other than C was broadcast.

There are several ways to account for the group size, each raising potential issues. One way would be to use a set of group-size dummies. This has the advantage of making no parametric assumption about the effect of group size – or the interaction between group size and communication – but is unsuitable for our purposes because (1) it ignores information from similar group sizes (e.g., the effect of communication in groups of 4 ought to be in between the effects in groups of 3 and 5), and (2) the results obtained cannot be extrapolated to other group sizes than those considered in the experiment. Another common choice, using a group-size variable, avoids these disadvantages but forces a linear effect (which would lead to nonsensical predicted probabilities when extrapolated to large group sizes). Instead of either of these, we use the reciprocal of the group size, which imposes an effect that decreases in magnitude as the group size increases (though all of our main results are robust to using the group size instead of its reciprocal). We also include the product of this (1/group-size) variable with either the Leader indicator or the C-message indicator.

Table 7: Tobit results for C choice fraction and group payoff efficiency, probit results for cooperative (all-C) Nash equilibrium outcomes, group-level data – estimated average marginal effects (unless otherwise stated) and standard errors

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	C choice		Cooperative outcome		Payoff efficiency	
1/(group size)	0.718***	0.616***	1.198***	1.176***	0.964***	0.916***
	(0.186)	(0.171)	(0.175)	(0.187)	(0.126)	(0.136)
Leader treatment…	0.108*	---	0.250***	---	0.192***	---
	(0.059)		(0.061)		(0.041)	
Group receives C message…	---	0.323***	---	0.376***	---	0.209***
		(0.052)		(0.076)		(0.045)
…when group size = 2	0.256***	0.389***	0.452***	0.557***	0.331***	0.371***
	(0.095)	(0.070)	(0.147)	(0.131)	(0.058)	(0.055)
… when group size = 3	0.150**	0.364***	0.354***	0.535***	0.253***	0.296***
	(0.066)	(0.053)	(0.089)	(0.108)	(0.046)	(0.052)
… when group size = 4	0.082	0.323***	0.207***	0.355***	0.177***	0.191***
	(0.064)	(0.058)	(0.072)	(0.103)	(0.046)	(0.056)
… when group size = 5	0.039	0.290***	0.133**	0.247**	0.126**	0.118
	(0.071)	(0.068)	(0.063)	(0.099)	(0.052)	(0.063)
… when group size = 7	−0.008	0.246***	0.072	0.146*	0.066	0.033
	(0.084)	(0.085)	(0.048)	(0.086)	(0.061)	(0.073)
… when group size = 15	−0.066	0.179	0.027	0.061	−0.008	−0.068
	(0.101)	(0.112)	(0.027)	(0.057)	(0.074)	(0.082)
Significance of difference:	*$p < 0.001$*	*$p < 0.001$*	*$p < 0.001$*	*$p < 0.001$*	*$p < 0.001$*	*$p < 0.001$*
(excluding n=2)	*$p \approx 0.16$*	*$p < 0.001$*	*$p \approx 0.001$*	*$p < 0.001$*	*$p < 0.001$*	*$p < 0.001$*
Observations	125	108	125	108	125	108
|ln(L)|	94.67	54.91	35.69	29.17	29.75	29.02

* (**, ***): Significant at 10% (5%, 1%) level.
Additional explanatory variables: lottery choice, S question average, I question average, G question average (all group means).

Additional controls are: the group average lottery choice (between 1 and 5, with higher numbers more risky), three variables capturing questionnaire responses ("S average", the group mean of responses to the S questions, and "I average" and "G average" defined similarly), and a constant term. All of the models are estimated using Stata (version 12.1). Table 7 reports the results.[18]

The effects of group size and communication individually are broadly consistent with the alternative hypotheses corresponding to Hypotheses 1 and 2. The reciprocal-group-size variable is positive and significant in all six models, meaning that the group size itself has a negative effect on C choices, all-C outcomes, and payoffs. The marginal effect of the Leader-treatment indicator is positive and at least weakly significant, while that of the group receiving a C message is positive and highly significant.

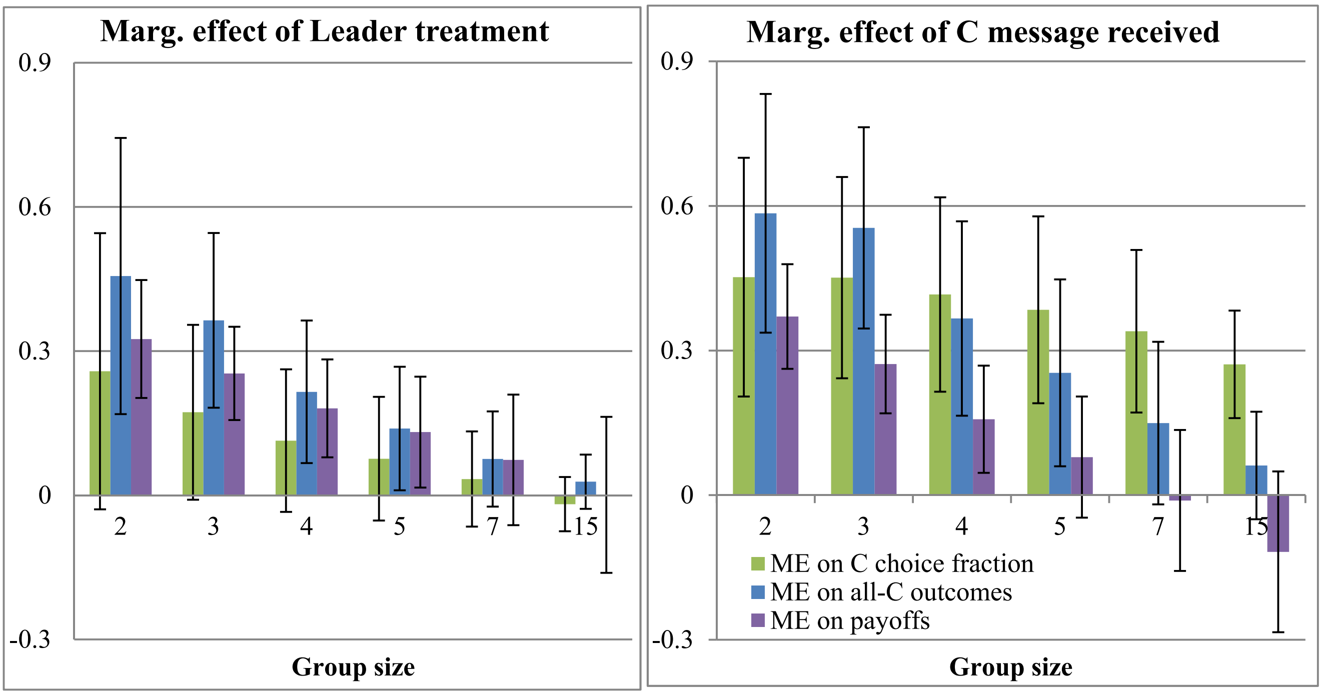


Figure 2: Marginal effects on C choices, all-C outcomes, and payoffs

The interaction between group size and communication is shown in Table 7 as the marginal effects of the Leader-treatment and C-message variables at each particular group size, and graphically in Figure 2. This interaction is stark, and similar in all six cases: both the Leader treatment and a broadcast C message have fairly large positive effects on cooperative (C) choice frequency, efficient (all-C) outcomes, and average payoffs when the group size is small, but these effects decrease sharply as the group size increases, in all but one case becoming insignificant and either close to zero or negative at the largest group size. In all six cases, we can reject the null hypothesis of equal marginal effects

[18] We have conducted many robustness checks, including: replacing the reciprocal group-size variable with a linear one; using the reciprocal group-size variable with an additional dummy for n=2 (in case groups of two are "special"); and using a set of group-size dummies instead of a single variable. We find that our original model has the best overall fit to the data: it out-performs the other models in all six cases (corresponding to the six columns of Table 7) according to the Bayesian Information Criterion (Schwarz, 1978), and is best in four of the six according to the Akaike (1974) Information Criterion with the correction for finite samples. Also, our main results are qualitatively robust to using these alternative models; the only qualitative change to a result is for the interaction between group size and the Leader treatment on C choices (the specification with an extra dummy for n=2 implies no interaction for group sizes larger than 2). Further regressions with only the Base and Choice treatments confirm that there are no significant differences in any of our outcome variables between these treatments. Some of these results can be found in Appendix E; the others, along with the estimation program and raw results, are available from the corresponding author upon request.

across group sizes at the 1% level or better; combined with the decrease in point estimates as group size increase, this indicates a significant negative interaction between group size and both the Leader treatment and C messages in this treatment.

6. Follow-up experiment: multiple messages

We chose to implement within-group communication by means of our random leader treatment because of its simplicity and because of the straightforward implication of the signals subjects receive: either a single C message or a single N message has a clear implication about the sender's intention and her view about what actions should be chosen. This implementation also has disadvantages, however. For example, as the group size increases, a single message provides less information about the preferences of the group as a whole. To address this issue, in this section we report results from a new treatment, called "full communication", in which group members receive ***all*** of the messages of their fellow group members, so allowing the absolute amount of communication to increase in a natural way with the group size.

Experimental procedures were largely similar to those from the original experiment. A total of 13 sessions, with 265 subjects, of the full-communication treatment were conducted, with group sizes of 2, 3, 5 and 15 (we left groups of 4 and 7 out of the design in hopes of increasing statistical power for the remaining groups). Subjects were told in the instructions that all messages from their group would be broadcast, in the form of the number of each type of message that was chosen (i.e., "Including you, X group members chose to send the message Y" for the three possible choices of message). Hence subjects knew for sure when choosing their message that it would be broadcast, in contrast to the Leader treatment where subjects did not know at this point whether their message (or indeed that of any group member) would be broadcast. The threshold public-good game, the structure of an experimental session (in particular the tasks used in parts 2 and 3 of a session), the location of the sessions, and the subject pool were identical to the original experiment.

Table 8: Aggregate results from Full treatment

	Group size			
	2	3	5	15
Number of groups	22	17	13	7
C messages	0.66	0.82	0.74	0.70
Groups receiving only C messages	0.41	0.47	0.31	0.00
C actions	0.57	0.63	0.55	0.40
difference from pooled Base/Choice	*−0.01*	*+0.19**	*+0.13*	*+0.01*
All-C outcomes	0.45	0.29	0.31	0.00
difference from pooled Base/Choice	*+0.09*	*+0.23**	*+0.31**	*0.00*
Payoffs	22.8	17.9	19.4	14.0
difference from pooled Base/Choice	*+2.3*	*+2.9*	*+6.7*	*−0.1*

**: Significantly different at 10% level*

Table 8 shows aggregate results: fractions of C messages sent, C actions chosen and efficient outcomes, and average payoffs (corresponding to Tables 3-6 for the other treatments). C messages are chosen 73% of the time (80% of non-blank messages), and while there is some variation across group sizes (fewer C messages when n=2 and more when n=3), these differences are not significant (χ^2_3=3.73, p ≈ 0.29).

The table shows that support for our alternative hypotheses is weaker in the Full treatment than in the other treatments. The effect of group size is fairly weak: while there is a rough decrease in C action frequencies with increasing group size, it is not significant (Jonckheere test, $z = -0.87$, $p \approx 0.19$), though the corresponding decreases for all-C outcomes and payoffs are at least weakly significant ($z = -1.59$, $p \approx 0.056$ and $z = -2.31$, $p \approx 0.010$ respectively). The beneficial effect of communication is also much weaker than we had seen for the Leader treatment, even for groups of 2: none of C action frequencies, all-C outcomes or payoffs is significantly different in the Full treatment compared to the pooled Base and Choice treatments for n=2, and overall only three of the twelve comparisons is significant at the 10% level, and none are at the 5% level.

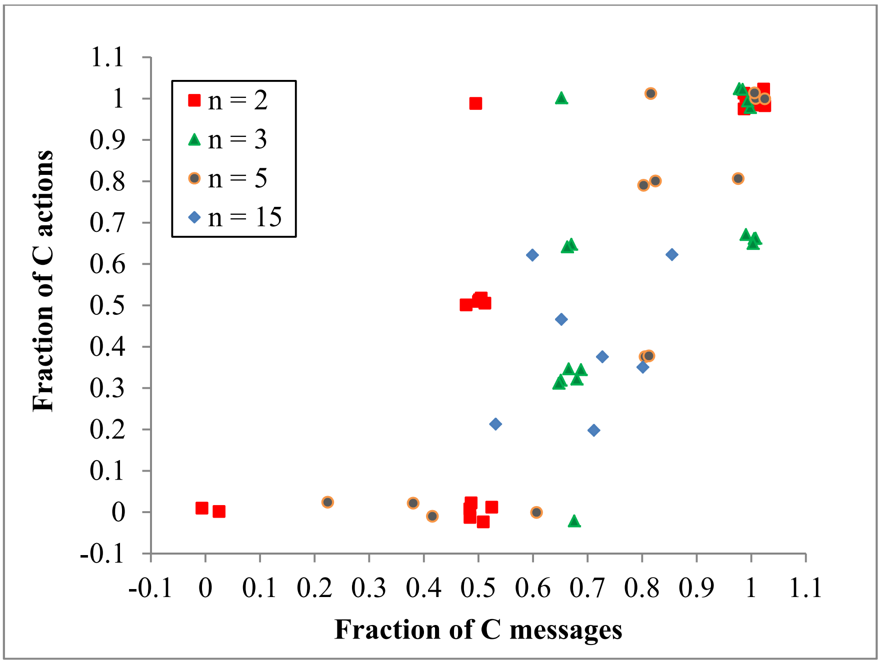


Figure 3: Scatter-plot of individual groups' fraction of C messages and actions (perturbed with uniform[-0.025,0.025] noise to make individual points visible).

Figure 3 shows the relationship between the fraction of C messages sent by members of a group and the fraction of C actions subsequently chosen by that group. The figure shows a positive, and fairly linear, relationship.[19] This relationship suggests that part of the explanation for the minimal benefit of this kind of communication lies in the actual messages that were sent. Even though the overall frequency of C messages in this treatment (73%) is comparable to that in the Leader treatment (67%), the broadcasting of all messages rather than just one of them means that a group will typically receive mutually-inconsistent suggestions, making successful coordination difficult. However, groups whose members all (or perhaps nearly all) sent C messages are likely to benefit from the communication.

We proceed as before by estimating a Tobit for the group fraction of C choices, a probit for all-C outcomes, and a Tobit for payoff efficiency, using the observations in the Base, Choice and Full treatments. We include two variables to capture communication: an indicator for the Full treatment and a continuous variable for the fraction of C messages sent by the group. We capture group size with the reciprocal-group-size variable as before, and we use the same additional controls (group average lottery choice, group averages of S, I and G responses) and a constant term. Table 9 reports the results.

These new regressions re-affirm the negative effect of group size (i.e., positive effect of the reciprocal of group size) on all three outcome variables. The Full-treatment dummy has a negative marginal effect, implying that communication makes outcomes worse when no C messages are sent, but the fraction-of-C-messages variable has the expected positive effect, indicating that outcomes

[19] We had anticipated that C action frequencies would be nearly zero except when every group member sent a C message; clearly, this did not happen.

improve as the group sends (and receives) a larger fraction of C messages. The middle block of the table shows the overall effect of the Full treatment for each group size.[20] Since this kind of communication did not increase C action choices in 2-person groups, it is not surprising that we see no significant interaction with group size here, though the sign of the relationship is decreasing as expected. In the case of efficient outcomes, the relationship is broadly decreasing (the exception being the increase going from n=2 to n=3) but not significant. In the case of payoffs, the marginal effect decreases with the group size as expected, and a test of equal effects across group sizes can be rejected.

Table 9: Tobit results for C choice fraction and group payoff efficiency, probit results for efficient outcomes, group-level data from Base, Choice and Full treatments – estimated average marginal effects (unless otherwise stated) and standard errors

Dependent variable:	(7) C choice	(8) Cooperative outcome	(9) Payoff efficiency
1/(group size)	0.488***	0.836***	0.725***
	(0.151)	(0.214)	(0.136)
Full treatment	−0.440***	−0.308***	−0.144
	(0.024)	(0.076)	(0.110)
Fraction of C messages	1.505***	0.800***	0.416***
	(0.172)	(0.121)	(0.151)
Full treatment, 72.8% C messages…			
…when group size = 2	0.142	0.102	0.185***
	(0.091)	(0.182)	(0.070)
… when group size = 3	0.111*	0.161*	0.162***
	(0.058)	(0.092)	(0.047)
… when group size = 5	0.081	0.081	0.130**
	(0.068)	(0.077)	(0.057)
… when group size = 15	0.048	0.027	0.091
	(0.101)	(0.047)	(0.086)
Significance of difference:	*$p \approx 0.42$*	*$p \approx 0.26$*	*$p \approx 0.015$*
Full treatment, 100% C messages…			
…when group size = 2	0.416***	0.560***	0.315***
	(0.063)	(0.143)	(0.067)
… when group size = 3	0.463***	0.714***	0.282***
	(0.044)	(0.103)	(0.056)
… when group size = 5	0.471***	0.578**	0.206**
	(0.069)	(0.188)	(0.087)
… when group size = 15	0.445***	0.387	0.109
	(0.144)	(0.307)	(0.136)
Significance of difference:	*$p \approx 0.22$*	*$p \approx 0.046$*	*$p \approx 0.010$*
Observations	133	133	133
|ln(L)|	64.33	33.57	40.64

* (**, ***): Significant at 10% (5%, 1%) level.
Additional explanatory variables: lottery choice, S question average, I question average, G question average (all group means).

The results are somewhat stronger in the lowest block of the table, which shows the estimated effect of the Full treatment when all messages are C, for each group size. (Note that this requires extrapolation in the case of n=15, since this outcome was never observed there.) Marginal effects decline with the group size in two of the three models (again, sometimes excepting the change from

[20] We calculate this by estimating the difference between the predicted dependent-variable value given a Full-treatment indicator equal to one and a fraction of C messages equal to the Full-treatment average (approx. 0.73), and the prediction given both of these equal to zero. Our calculation of the effect of all C messages in the Full treatment (for the bottom section of the table) is similar, except that the fraction of C messages is set to one.

groups of two to groups of three), and tests of equal effects across group sizes can be rejected in the cases of efficient outcomes and payoffs.[21]

So in a broad sense, the results from our new Full treatment show that our conclusions are robust to changing the type of communication from one-way to all-way. The main difference from the Leader treatment is that in the Full treatment, communication may not even improve outcomes in the smallest groups, in which case the interaction between group size and communication is typically insignificant. However, in cases where communication does improve outcomes in small groups (notably, payoffs when everyone sends a C message), we tend to see the same negative interaction as before. Either way, the benefit in large groups is minimal.

7. Discussion

With rare exceptions, the effects of group size and the availability of communication in strategic settings have previously been studied in isolation. In this paper, we orthogonally vary both of them, allowing examination of their ***joint*** as well as separate effects. Groups of size varying from 2 up to 15 play a multi-player threshold public-good game one time, either with or without communication in the form of a single pre-play cheap-talk message – framed as a suggestion – sent by one group member selected at random (the "leader"). We find that the beneficial effect of such communication on cooperative behaviour, either overall or conditional on the group receiving a "cooperate" message, falls significantly as the group size increases. Even more striking are the corresponding decreases in successful coordination and payoffs: from a strong positive effect in very small groups, they fall to statistical insignificance and to magnitudes near zero or even less than zero at the largest group sizes we consider. Qualitatively similar results are found in a follow-up experiment where not just one, but all, group members' messages are broadcast.

Our main result – this negative interaction between group size and communication – leads us to what we believe is an important point about interpreting results from other experiments. We conjecture that ***any*** particular mode of communication, held constant, will become less and less effective as group sizes increase. Obviously, this is an empirical question, and one that will require substantial additional research to address. We have only considered a very specific kind of communication, and other modes that have previously been found to work particularly well – such as free-form pre-play messages, or post-play messages that can shame bad behaviour – should receive particular attention in future research. Other strategic settings should also be considered, such as linear public good games or coordination games with payoff asymmetry. However, our results ought at least to suggest some caution is warranted when assuming benefits arising from communication found in a small group automatically extends to much larger groups.[22]

Our findings also highlight a broader issue: caution should be used in extrapolating treatment effects found in laboratory experiments to the outside world. This should be true especially in strategic environments where group size has been found to matter, such as coordination games, social

---

[21] The increase in marginal effects as we move from n=2 to n=3 in the 100%-C-messages case is due to the restriction of frequencies to [0, 1]. In the pooled Base and Choice treatments, the regression models predict that the fractions of C choices and efficient outcomes would be 0.56 and 0.35 respectively when n=2, meaning that the marginal effect of the Full treatment is bounded above at 0.44 and 0.65, less in both cases than the corresponding marginal effects when n=3.

[22] A related criticism of our conclusions is based on the recognition that our experiment's largest group size of 15 is still orders of magnitude below the outside-world group sizes we mention; because of this, one might question whether our results themselves have any relevance for such large groups. Our response to such questioning is that our approach of determining point estimates for a range of group sizes, thus yielding a trend that can be extrapolated beyond the range we consider, is superior to approaches that find a point estimate from a single group size and assume there is no trend. We acknowledge that extrapolation is always risky, and that we cannot rule out the possibility that the observed decline in effectiveness of communication reverses itself at some group size larger than what we have considered, so that further increases in the group size would ***improve*** the benefits of communication. Such a reversal seems implausible to us, but this is ultimately an empirical question, and we encourage others to test their conjectures in future research.

dilemmas, and oligopoly experiments. Questions about the external validity of lab experiments are of course not new, and attention has been given to many differences between the lab and the field (see Harrison and List, 2004, for a thorough discussion of potential issues), but the difference between the small groups typically studied in the lab and the large groups common in the field has received, to our knowledge, little attention. We encourage other researchers to probe the robustness of treatment effects they discover by varying not only their main treatment variable, but also the size of the group.

References

Aflagah, F.K.D., T. Bernard, and A. Viceisza. 2014. Communication and coordination: Experimental evidence from farmer groups in Senegal. Working Paper.

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716-723.

Aumann, R. 1990. Nash equilibria are not self-enforcing. In J.J. Gabszewicz, J.-F. Richard, and L.A. Wolsey (eds.) *Economic Decision Making: Games, Econometrics and Optimisation.* North Holland, Amsterdam: 201-206.

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

Ben-Ner, A., L. Putterman, and T. Ren. 2011. Lavish returns on cheap talk: Two-way communication in trust games. *The Journal of Socio-Economics* 40: 1-13.

Blume, A. and A. Ortmann. 2007. The effects of costless pre-play communication: experimental evidence from games with Pareto-ranked equilibria. *Journal of Economic Theory* 132: 274-290.

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

Brandts, J., D.J. Cooper, and R. Weber. 2011. Legitimacy, social distance, and leadership in the turnaround game. Florida State University Working Paper.

Brosig, J., A. Ockenfels, and J. Weimann (2003). The effect of communication media on cooperation. German Economic Review 4: 217-241.

Camerer, C.F. and M. Knez. 1997. Coordination in organizations: A game-theoretic perspective. In Z. Shapira (ed), *Organizational Decision Making.* Cambridge Series on Judgement and Decision Making, Cambridge, UK: 158-188.

Carpenter, J. 2007. Punishing free-riders: how group size affects mutual monitoring and the provision of public goods. *Games and Economic Behavior* 60: 31-51.

Charness, G. 2000. Self-serving cheap talk: a test of Aumann's conjecture. *Games and Economic Behavior* 33: 177-194.

Charness, G., and M. Dufwenberg. 2006. Promises and partnership. *Econometrica* 74: 1579-1601.

Charness, G. and B. Grosskopf. 2004. What makes cheap talk effective? Experimental evidence. *Economics Letters* 83: 383-389.

Clark, K., S. Kay, and M. Sefton. 2001. When are Nash equilibria self-enforcing? An experimental analysis. *International Journal of Game Theory* 29: 495-515.

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross. 1989. Communication in the battle of the sexes game: some experimental results. *RAND Journal of Economics* 20: 568-587.

Cooper, R., D.V. DeJong, R. Forsythe, and T.W. Ross. 1992. Communication in coordination games. *Quarterly Journal of Economics* 107: 739-771.

Crawford, V. 1998. A survey of experiments on communication via cheap talk. *Journal of Economic Theory* 78: 286-298.

Dannenberg, A. 2015. Leading by example versus leading by words in voluntary contribution experiments. *Social Choice and Welfare* 44: 71-85.

Dawes, R.M., J. McTavish, and H. Shaklee. 1977. Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology* 35: 1-11.

Deutsch, M. 1960. The effect of motivational orientation upon trust and suspicion. *Human Relations* 13: 123-139.

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.

Eckel, C.C. and P.J. Grossman. 2008. Forecasting risk attitudes: an experimental study of actual and forecast risk attitudes of women and men. *Journal of Economic Behavior and Organization* 68: 1-17.

Ellingsen, T. and R. Östling. 2010. When does communication improve cooperation? *American Economic Review* 100: 1695-1724.

Farrell, J. 1995. Talk is cheap. *American Economic Review* 85: 186-190.

Farrell, J. and M. Rabin. 1996. Cheap talk. *The Journal of Economic Perspectives* 10: 103-118.

Feltovich, N. 2005. Critical values for the robust rank-order test. *Communications in Statistics - Simulation and Computation* 34: 525-547.

Fonseca, M.A. and H.-T. Normann. 2012. Explicit vs. tacit collusion - the impact of communication in oligopoly experiments. *European Economic Review* 56: 1759-1772.

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

Goeree, J.K., C.A. Holt, and S.K. Laury 2002. Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *Journal of Public Economics* 83: 255-276.

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

Grossman, P.J. 2013. Holding fast: the persistence and dominance of gender stereotypes. *Economic Inquiry* 51: 747-763.

Harrison, G.W. and J.A. List. 2004. Field experiments. *Journal of Economic Literature* 42: 1009-1055.

Harsayni, J.C. and R. Selten. 1988. *A general theory of equilibrium selection in games*. MIT Press, Cambridge, MA.

Heinemann, F., R. Nagel, and P. Ockenfels. 2009. Measuring strategic uncertainty in coordination games. *Review of Economic Studies* 76: 181-221.

Hermalin, B.E. 1998. Toward an economic theory of leadership: leading by example. *American Economic Review 88*: 1188-1206.

Isaac, R.M., and J.M. Walker. 1988. Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry* 26: 585-608.

Isaac, R.M., J.M. Walker, and A.W. Williams. 1994. Group-size and the voluntary provision of public goods – experimental evidence utilizing large groups. *Journal of Public Economics* 54: 1-36.

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

Jerdee, T.H. and B. Rosen. 1974. Effects of opportunity to communicate and visibility of individual decisions on behavior in the common interest. *Journal of Applied Psychology* 59: 712-716.

Kollock, P. 1998. Social dilemmas: the anatomy of cooperation. *Annual Review of Sociology* 24: 183-214.

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

Komai, M. and M. Stegeman. 2010. Leadership based on asymmetric information. *Rand Journal of Economics* 41: 35-63.

Komai, M., M. Stegeman, and B.E. Hermalin. 2007. Leadership and information. *American Economic Review* 97: 944-947.

Koukoumelis, A., M.V. Levati, and J. Weisser. 2012. Leading by words: A voluntary contribution experiment with one-way communication. *Journal of Economic Behavior & Organization* 8: 379-390.

Ledyard, J.O. 1995. Public goods: a survey of experimental research. In J.H. Kagel and A.E. Roth (eds.), *Handbook of experimental economics*. Princeton University Press, Princeton, NJ: 111-194.

Leibbrandt, A. and L. Sääksvuori. 2012. Communication in intergroup conflicts. *European Economic Review*: 1136-1147.

Loomis, J.L. 1959. Communication, the development of trust, and cooperative behavior. *Human Relations* 12: 305-315.

Masclet, D., Noussair, C., Tucker, S., Villeval, M-C. 2003. Monetary and nonmonetary punishment in the voluntary contributions mechanism. *American Economic Review* 93: 366-380.

Nosenzo, D., S. Quercia, and M. Sefton. 2015. Cooperation in small groups: the effect of group size. *Experimental Economics* 18: 4-14.

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

Palfrey, T.R. and H. Rosenthal. 1991. Testing for effects of cheap talk in a public goods game with private information. *Games and Economic Behavior* 3: 183-220.

Pogrebna, G., D.H. Krantz, C. Schade, and C. Keser. 2011. Words versus actions as a means to influence cooperation in social dilemma situations. *Theory and Decision* 71: 473-502.

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

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

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

Schwarz, G., 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461-464.

Swensson, R.G. 1967. Cooperation in the prisoner's dilemma game I: The effects of asymmetric payoff information and explicit communication. *Behavioral Science* 12: 314-322.

Van Huyck, J.B., R.C. Battalio, and R.O. Beil. 1990. Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review* 80: 234-248.

Vesterlund, L. 2003. The informational value of sequential fund-raising. *Journal of Public Economics* 87: 627-657.

Weber, E. U., A.-R. Blais, and N. Betz. 2002. A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making* 15: 263-290.

Weber, R., C. Camerer, Y. Rottenstreich, and M. Knez. 2001. The illusion of leadership: misattribution of cause in coordination games. *Organization Science* 12: 582-598.

Weimann, J., J. Brosig-Koch, H. Hennig-Schmidt, C. Keser, and C. Stahr. 2012. Public-good experiments with large groups. Working Paper n. 120009, Otto-von-Guericke University Magdeburg FEMM.

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, we use horizontal lines where page breaks would have appeared; (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 (on one double-sided page), the part-2 instructions before the risky-choice task, and the part-3-and-4 instructions before the survey.

General instructions [all treatments]

You are about to participate in a decision making experiment. Please read these instructions carefully, as how much money you earn may depend on how well you understand them. If you have a question at any time, please feel free to ask the experimenter. We ask that you not talk with the other participants during the experiment.

This experiment consists of four parts. The first two parts are decision tasks. The amount you are paid for these tasks will depend on the decisions you make, other participants' decisions, and chance. You will receive detailed instructions for these tasks before you begin them. The third part of the experiment is a questionnaire; you will receive a flat payment of $5 for completing this part. In the fourth and final part of the experiment, you will learn the results of the first two parts, as well as the total amount you will be paid for the experiment (i.e., the sum of your payments for parts 1-3). After this, you will be paid, privately and in cash.

Instructions: Part 1 [Base treatment]

In this part of the experiment, you will take part in an Interactive Decision Task. You and the other participants will be randomly assigned to groups by the computer. Your group will have at least two members, including you. You will be told the size of your group, but you will not be told the identities of the other members – even after the experiment ends.

The task: The task is as follows. You and the other members of your group will be simultaneously asked to choose a colour: either Red or Green. Each of you makes your choice without knowing anyone else's choice. Your earnings from the task – and the other group members' earnings – depend on the colours that were chosen. If everyone in the group has chosen Red, then each group member earns $30. If one or more group members have chosen Green, then each member who chose Green earns $20, and any member who chose Red earns $5.

		Other group members' colours			
		All Red		At least one Green	
Your colour	Red	You earn:	$30	You earn:	$5
	Green	You earn:	$20	You earn:	$20

Sequence of play: The sequence of play for this task is as follows.
(1) Your computer screen shows the number of members in your group.
(2) You choose a colour.
(3) The task ends. You go on to the next task (part 2 of the experiment).

In part 4 of the experiment, you will be informed of the results of this task: your colour, the number of group members choosing each colour, and your earnings.

Instructions: Part 1 [Choice/Leader treatments]

In this part of the experiment, you will take part in an Interactive Decision Task. You and the other participants will be randomly assigned to groups by the computer. Your group will have at least two members, including you. You will be told the size of your group, but you will not be told the identities of the other members – even after the experiment ends.

The task: The task is as follows. You and the other members of your group will be simultaneously asked to choose a colour: either Red or Green. Each of you makes your choice without knowing anyone else's choice. Your earnings from the task – and the other group members' earnings – depend on the colours that were chosen. If everyone in the group has chosen Red, then each group member earns $30. If one or more group members have chosen Green, then each member who chose Green earns $20, and any member who chose Red earns $5.

		Other group members' colours			
		All Red		At least one Green	
Your colour	Red	You earn:	$30	You earn:	$5
	Green	You earn:	$20	You earn:	$20

Messages: Before performing the task, you will have the opportunity to choose a message that might be sent to the members of your group. There are two possible messages: "Everyone should choose Red", and "Everyone should choose Green". It is also possible to send no message. To choose one of the messages, or to choose no message, click the corresponding button on your screen.

After everyone has made their message choice, the computer randomly chooses whether or not a message from your group will be sent to all group members; it is possible that exactly one message will be sent, or that none are. If a message is sent from your group, the computer randomly chooses one of the group members, and everyone in the group is shown this member's message (or non-message), before making their colour choices. This means that if you are shown a message, then everyone in your group is shown the same message.

You are told at this time whether your message was chosen to be sent, but you receive no other information about the sender's identity. If the computer determines that no message will be sent from your group, then you are informed of this before making your colour choice.

Sequence of play: The sequence of play for this task is as follows.

(1) Your computer screen shows the number of members in your group. You choose the message you wish to send, or you choose not to send a message.
(2) If your group is chosen to receive a message, then all of you view the message sent by the same randomly chosen group member.
(3) You choose a colour.
(4) The task ends. You go on to the next task (part 2 of the experiment).

In part 4 of the experiment, you will be informed of the results of this task: your colour, the number of group members choosing each colour, and your earnings.

Instructions: Part 1 [Full treatment]

In this part of the experiment, you will take part in an Interactive Decision Task. You and the other participants will be randomly assigned to groups by the computer. Your group will have at least two members, including you. You will be told the size of your group, but you will not be told the identities of the other members – even after the experiment ends.

The task: The task is as follows. You and the other members of your group will be simultaneously asked to choose a colour: either Red or Green. Each of you makes your choice without knowing anyone else's choice. Your earnings from the task – and the other group members' earnings – depend on the colours that were chosen. If everyone in the group has chosen Red, then each group member earns $30. If one or more group members have chosen Green, then each member who chose Green earns $20, and any member who chose Red earns $5.

		Other group members' colours	
		All Red	At least one Green
Your colour	Red	You earn: $30	You earn: $5
	Green	You earn: $20	You earn: $20

Messages: Before performing the task, you will have the opportunity to choose a message that will be sent to the members of your group. There are two possible messages: "Everyone should choose Red", and "Everyone should choose Green". It is also possible to send no message. To choose one of the messages, or to choose no message, click the corresponding button on your screen.

After everyone has made their message choice, you will be shown all of the messages sent by your group. You will see how many group members (including yourself) sent the message "Everyone should choose Red", how many sent "Everyone should choose Green", and how many chose not to send a message. Notice that everyone in your group observes the same information at this stage.

Sequence of play: The sequence of play for this task is as follows.
(1) Your computer screen shows the number of members in your group. You choose the message you wish to send, or you choose not to send a message.
(2) All members of your group view how many messages of each type were sent.
(3) You choose a colour.
(4) The task ends. You go on to the next task (part 2 of the experiment).

In part 4 of the experiment, you will be informed of the results of this task: your colour, the number of group members choosing each colour, and your earnings.

Instructions: Part 2 [all treatments]

In this part of the experiment, you will take part in a Gamble Selection Task. You will be shown five gambles, and will be asked to choose the one you prefer. Each gamble has two possible outcomes, both with equal (50%) chance of occurring. Your earnings from this task will depend on which gamble you choose, and which outcome occurs.

The gambles are as follows:

Gamble	Random numbers 1-50 (50% chance)	Random numbers 51-100 (50% chance)
1	You earn $4	You earn $4
2	You earn $6	You earn $3
3	You earn $8	You earn $2
4	You earn $10	You earn $1
5	You earn $12	You earn $0

After you have chosen one of these gambles, the computer will randomly draw a whole number between 1 and 100 (inclusive). If the random number is 50 or less, your earnings from this task are as shown in the middle column of the table. If the random number is 51 or more, your earnings from this task are as shown in the right column of the table. The random number drawn for you may be different from the ones drawn for other participants.

Once everyone has chosen a gamble, you will proceed to part 3 of the experiment. In part 4 of the experiment, you will be informed of the results of this task: your choice of gamble, your random number, and your earnings.

Instructions: Parts 3 and 4 [all treatments]

In part 3 of the experiment, you will be asked to fill out a questionnaire. The first part of this questionnaire lists 20 activities (on four screens of 5 activities each). For each activity, please indicate your likelihood of engaging in it, using the following scale:

O	O	O	O	O	O
1 **Extremely Unlikely**	**2** **Unlikely**	**3** **Somewhat Unlikely**	**4** **Somewhat Likely**	**5** **Likely**	**6** **Extremely Likely**

After entering this information, you will be shown another screen containing some demographic questions. You will receive $5 for answering all of the questions. Once everyone has done so, you will proceed to part 4 of the experiment, where you will be shown the results of parts 1 and 2, along with your total earnings from all parts of the experiment.

After everyone has been shown their total earnings, you will be called, one by one, to receive your payment.

Appendix B: Risk attitude and demographic questions from Part 3 of the experiment

Risk attitude questions

The 20 questions are listed below, along with the scale (S = social, I = investment, G = Gamble) each one belongs to. Each question was answered on a 6-point Likert scale.

Asking your boss for a pay rise. (S)
Investing 10% of your annual income in government bonds. (I)
Investing in a business that has a good chance of failing. (I)
Deciding to share an apartment with someone you don't know well. (S)
Wearing unconventional clothes. (S)

Arguing with a friend, who has a very different opinion on an issue. (S)
Speaking your mind about an unpopular issue at a social occasion. (S)
Investing 10% of your annual income in blue-chip stock. (I)
Spending money impulsively without thinking about the consequences. (G)
Admitting that your tastes are different from those of your friends. (S)

Moving to a new city. (S)
Investing 10% of your annual income in a very speculative stock. (I)
Taking a job where you get paid exclusively on a commission basis. (G)
Lending a friend an amount of money equivalent to one month's income. (I)
Disagreeing with your father on a major issue. (S)

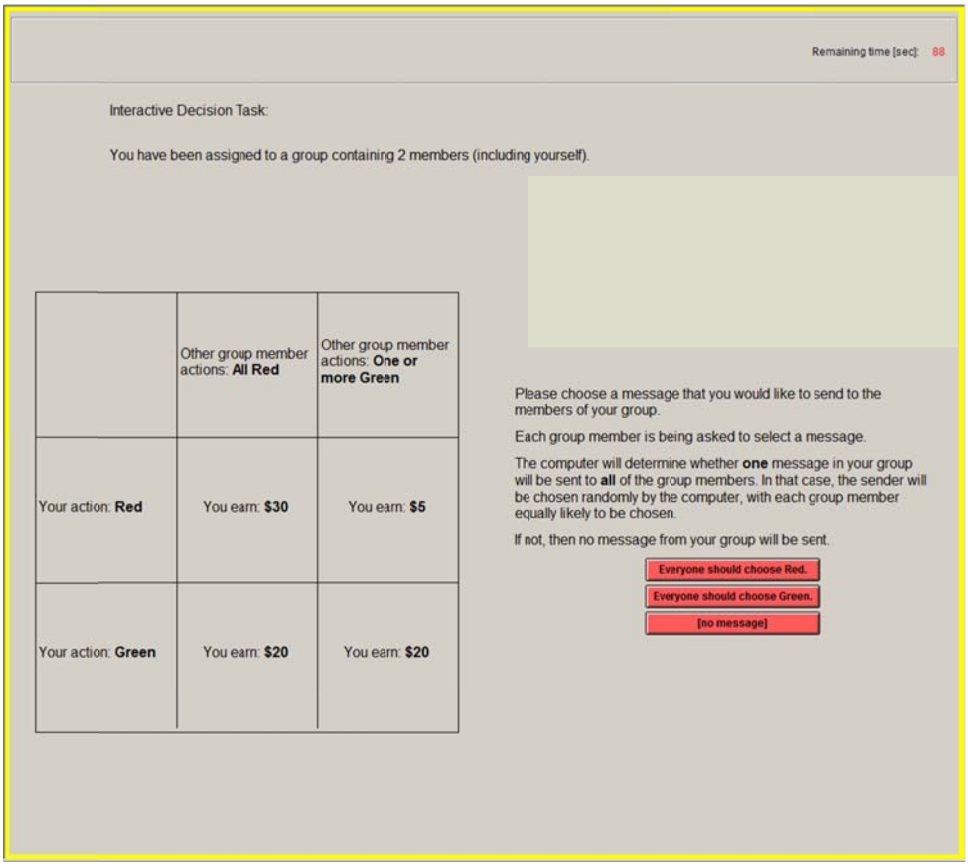
Openly disagreeing with your boss in front of your coworkers. (S)
Co-signing a new car loan for a friend. (G)
Taking a day's income to play the pokies (poker machines) at a casino. (G)
Dating someone that you are working with. (S)
Betting a day's income at the horse races. (G)

Demographic questions

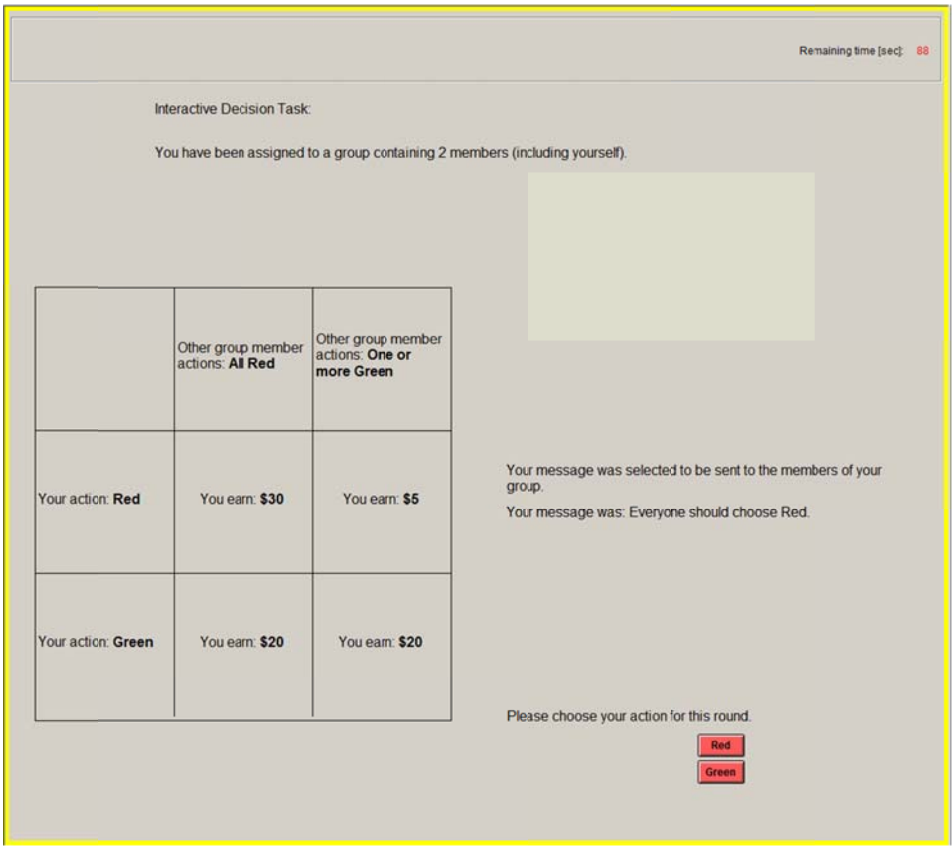
What is your age, to the nearest year? [positive integer]
What is your gender? [female/male]
Were you born in Australia? [yes/no]
How many years have you lived in Australia? [positive integer]
What is your major? [economics/other business/other]
Since you began studying at university, how many Economics classes have you completed? [0/1/2/3/4/more than 4]

Appendix C: Sample screenshots

Message choice screen (Choice/Leader):



Action choice screen (Leader):



Message choice screen (Full):

Remaining time [sec]: 89

Interactive Decision Task:

You have been assigned to a group containing 5 members (including yourself).

	Other group member actions: **All Red**	Other group member actions: **One or more Green**
Your action: **Red**	You earn: **$30**	You earn: **$5**
Your action: **Green**	You earn: **$20**	You earn: **$20**

Please choose a message that you would like to send to the members of your group.

Each group member is being asked to select a message.

Your message, and the other group members' messages, will be sent to **all** of the group members.

Everyone should choose Red.

Everyone should choose Green.

[no message]

Action choice screen (Full):

Remaining time [sec]: 85

Interactive Decision Task:

You have been assigned to a group containing 5 members (including yourself).

	Other group member actions: **All Red**	Other group member actions: **One or more Green**
Your action: **Red**	You earn: **$30**	You earn: **$5**
Your action: **Green**	You earn: **$20**	You earn: **$20**

Including your message, the messages sent by members of your group were:

3 sent the message "Everyone should choose Red".

1 sent the message "Everyone should choose Green".

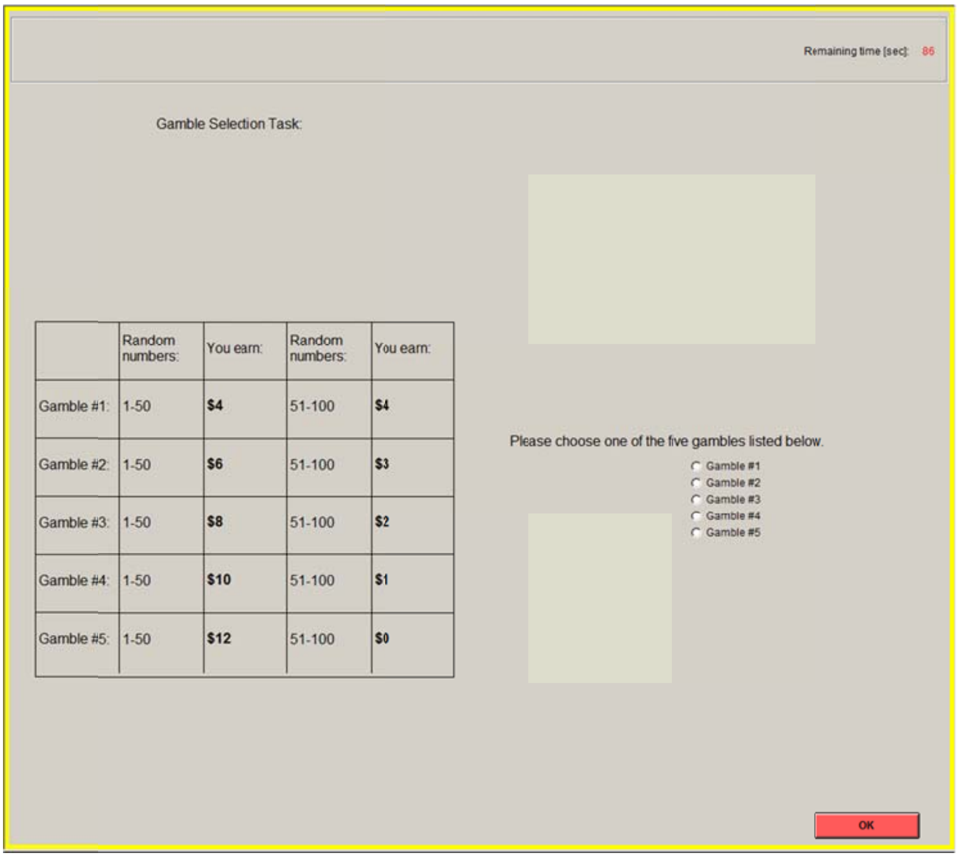
1 sent no message.

Please choose your action for this round.

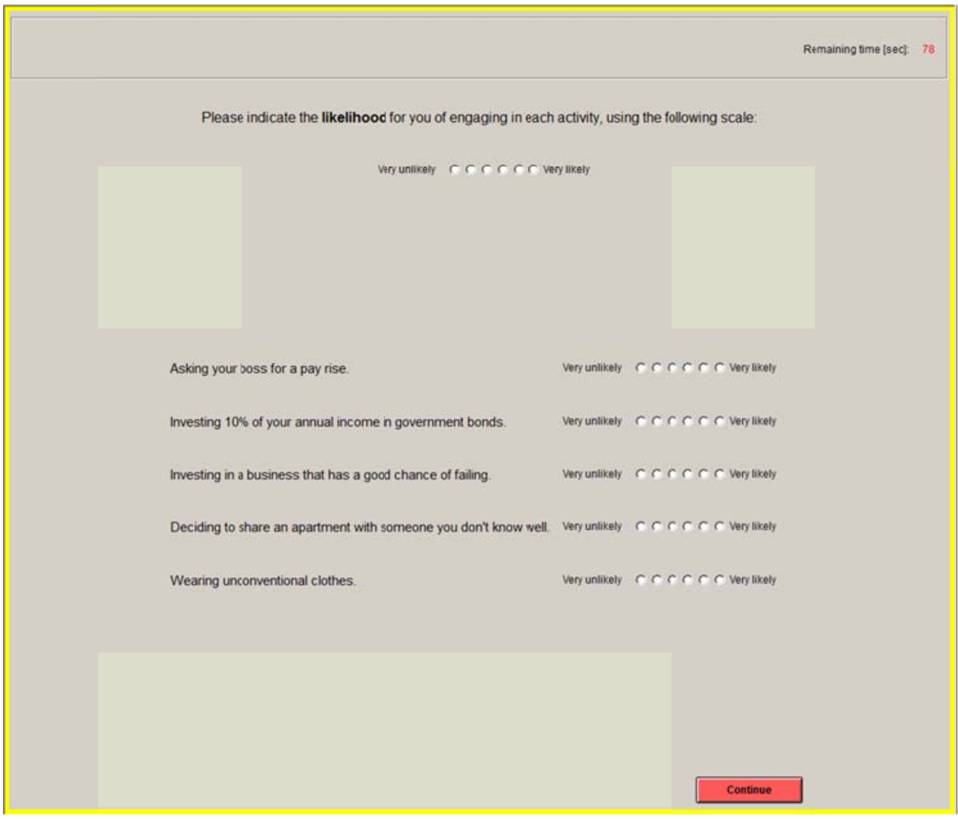
Red

Green

Lottery choice screen:



Questionnaire screen:



Appendix D: Other notable results

Beyond the main treatment effects discussed above, our data show a number of regularities involving aspects of behaviour such as message sending, lottery choices and questionnaire responses. Even though we did not begin our study with explicit hypotheses concerning these aspects, they may be relevant to future research, and thus worth cataloguing.

Table D1: Lottery choices and questionnaire responses

		Lottery choice				
		1 (4, 4)	2 (6, 3)	3 (8, 2)	4 (10, 1)	5 (12, 0)
Percent choosing		7.6	13.5	20.4	17.7	40.8
	Percent of all questionnaire responses					
	1	2	3	4	5	6
10 S questions	8.4	13.6	16.1	22.3	23.8	15.8
5 G questions	23.9	20.8	17.7	18.2	13.1	6.4
5 I questions	40.0	22.2	14.3	12.9	7.5	3.1

Descriptive statistics for the lottery choices and questionnaire responses (all subjects and treatments) are shown in Table D1. Two features are evident. First, nearly half of our subjects choose the riskiest (highest expected value) lottery, far more than the 24% making the corresponding choice in Eckel and Grossman (2008).[23] On the other hand, our subjects' responses to the risk attitude questions are quite similar to those of Grossman's (2013) subjects. Additionally, we observe a positive albeit weak relationship between lottery choice and mean questionnaire response for each of the three domains (see Figure D1).

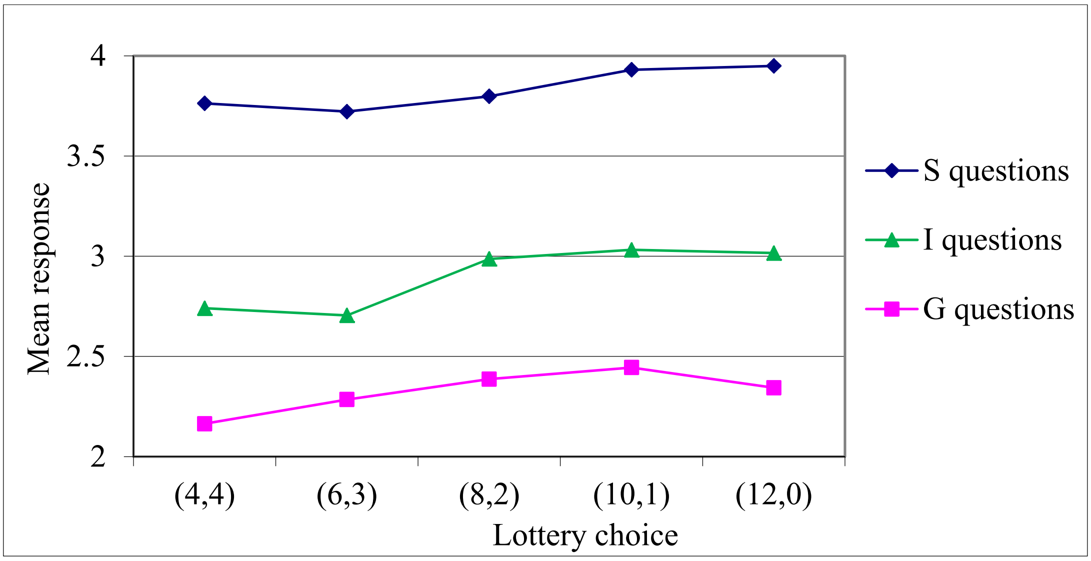


Figure D1: Association between lottery choices and questionnaire responses

[23] This could be due to subjects knowing that regardless of the lottery outcome, they would have earnings of at least \$10 (minimum earnings in Eckel and Grossman's experiment were \$0) or because the stakes were relatively low (one-fourth those of Eckel and Grossman's). More prosaic explanations include differences in the subject pool or in the wording of instructions. It is also worth noting that one common criticism of computerised random draws – the claim that they are viewed as "rigged" by subjects – seems not to apply to our subject pool.

Table D2 shows results from additional regressions. We estimate a probit model (model 10) with an indicator for a C message sent as the dependent variable, and a Tobit model (model 11) with lottery choice (1 = safest, 5 = riskiest) as the dependent variable, both using individual-level data. For the C-message probit, our right-hand-side variables are the group size, the lottery choice, various demographic variables (listed in the table), and the three questionnaire-average variables that were used in the previous probits (see Table 7), and all individual-level observations except those in the Base treatment are used. The list of right-hand-side variables for the lottery-choice Tobit was similar, except without the group size and (of course) the lottery choice, and with all individual-level observations used.

Table D2: Regression results – estimated marginal effects (at variables' means) and standard errors

Dependent variable:	(10) C message sent (probit)	(11) Lottery choice (Tobit)
Group size	0.000	---
	(0.003)	
Lottery choice (1=safe,	0.041***	---
5=highest EV)	(0.013)	
S question average	−0.041	0.295*
(1=least risk, 6=most risk)	(0.028)	(0.152)
I question average	−0.027	0.239*
(1=least risk, 6=most risk)	(0.025)	(0.134)
G question average	0.053**	−0.062
(1=least risk, 6=most risk)	(0.025)	(0.124)
Female	0.051	−0.907***
	(0.036)	(0.179)
Born in Australia	0.047	−0.608
	(0.075)	(0.402)
Fraction of life in	−0.085	0.715
Australia	(0.089)	(0.436)
Econ classes taken	−0.011	−0.092
(up to 4)	(0.016)	(0.077)
Age (yrs.)	0.002	−0.049***
	(0.003)	(0.016)
Econ major	−0.071	0.318
	(0.075)	(0.308)
Business (non-econ) major	0.043	0.144
	(0.045)	(0.219)
Constant term?	Yes	Yes
N	721	887
|ln(L)|	435.86	1383.53

* (**, ***): Significant at 10% (5%, 1%) level.

Since we have no ex-ante hypotheses concerning these regressions, we simply mention a few of the interesting results. There is no group-size effect on message choice; subjects in larger groups are no more or less likely to send C messages. Subjects choosing riskier lotteries are more likely to send a C message, as are those expressing more preference for gambling-type risk. On the other hand, subjects expressing more preference for social-type risk may be ***less*** likely to send a C message, though the effect just misses significance ($p \approx 0.14$).[24] Females are nearly significantly ($p \approx 0.15$) more likely than males to send a C message, even after controlling for risk attitudes; they also make safer lottery choices. Older subjects make safer lottery choices.

[24] The different direction of effects of the S and G averages is not counter-intuitive. Sending a C message – to the extent that it implies a greater likelihood of choosing a C action – can be thought of as taking on a gamble. On the other hand, it can also be thought of as proposing the attractive joint-payoff-maximising outcome (in contrast to an N or blank message); hence it might be ***less*** socially risky to do so.

Appendix E: Additional regression results

Replication of Table 7, but using group size instead of its reciprocal on the right-hand side:

Table E1: Tobit results for C choice fraction and group payoffs, probit results for cooperative (all-C) Nash equilibrium outcomes, group-level data – estimated average marginal effects (unless otherwise stated) and standard errors

	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)
Dependent variable:	C choice		Cooperative outcome		Payoffs	
Group size	−0.013*	−0.015**	−0.133***	−0.138***	−0.024***	−0.024***
	(0.008)	(0.008)	(0.032)	(0.036)	(0.005)	(0.006)
Leader treatment…	0.104*	---	0.259***	---	0.199***	---
	(0.063)		(0.066)		(0.045)	
Group receives C message…	---	0.337***	---	0.400***	---	0.246***
		(0.052)		(0.077)		(0.050)
…when group size = 2	0.145*	0.387***	0.329**	0.491***	0.282***	0.352***
	(0.080)	(0.060)	(0.152)	(0.142)	(0.055)	(0.057)
… when group size = 3	0.132*	0.375***	0.442***	0.637***	0.258***	0.324***
	(0.072)	(0.056)	(0.097)	(0.107)	(0.051)	(0.055)
… when group size = 4	0.118*	0.362***	0.326***	0.510***	0.232***	0.291***
	(0.066)	(0.054)	(0.080)	(0.103)	(0.048)	(0.054)
… when group size = 5	0.105*	0.347***	0.191**	0.331***	0.204***	0.254****
	(0.064)	(0.054)	(0.081)	(0.119)	(0.047)	(0.055)
… when group size = 7	0.077	0.311***	0.042	0.086	0.143**	0.168**
	(0.070)	(0.063)	(0.047)	(0.090)	(0.053)	(0.064)
… when group size = 15	−0.032	0.117	0.0000002	0.0000004	−0.095	−0.164
	(0.158)	(0.167)	(0.004)	(0.0000032)	(0.111)	(0.116)
Significance of difference:	*p ≈ 0.65*	*p < 0.001*	*p < 0.001*	*p < 0.001*	*p ≈ 0.002*	*p < 0.001*
Observations	125	108	112	112	112	112
|ln(L)|	100.91	59.80	37.86	30.18	43.33	42.38

* (**, ***): Significant at 10% (5%, 1%) level.
Additional explanatory variables: lottery choice, S question average, I question average, G question average (all group means).

Replication of Table 7, but including a dummy for n = 2 (along with the reciprocal of n):

Table E2: Tobit results for C choice fraction and group payoff efficiency, probit results for cooperative (all-C) Nash equilibrium outcomes, group-level data – estimated average marginal effects (unless otherwise stated) and standard errors

	(1b)	(2b)	(3b)	(4b)	(5b)	(6b)
Dependent variable:	C choice		Cooperative outcome		Payoffs	
1/(group size)	−0.048*	0.315	3.578	4.577	0.420*	0.478*
	(0.357)	(0.340)	(192.4)	(443.1)	(0.250)	(0.281)
group size = 2	0.300	0.121	−0.015	−0.169	0.223**	0.176*
	(0.107)	(0.116)	(0.107)	(0.067)	(0.086)	(0.097)
Leader treatment…	0.091	---	0.242***	---	0.184***	---
	(0.058)		(0.061)		(0.040)	
Group receives C	---	0.317***	---	0.393***	---	0.202***
message…		(0.053)		(0.081)		(0.047)
…when group size = 2	0.301***	0.354***	0.488***	0.552	0.296***	0.314***
	(0.093)	(0.075)	(0.147)		(0.061)	(0.060)
… when group size = 3	−0.004	0.408***	0.249	0.636	0.273***	0.383***
	(0.115)	(0.109)	(0.180)		(0.080)	(0.102)
… when group size = 4	0.010	0.338***	0.198***	0.408	0.180***	0.225***
	(0.075)	(0.079)	(0.070)		(0.054)	(0.074)
… when group size = 5	0.018	0.290***	0.145*	0.248	0.122***	0.121*
	(0.071)	(0.071)	(0.062)		(0.051)	(0.064)
… when group size = 7	0.028	0.229***	0.098	0.117	0.056**	0.001
	(0.089)	(0.089)	(12.96)		(0.064)	(0.075)
… when group size = 15	0.040	0.143	0.054	0.031	−0.029	−0.144
	(0.133)	(0.142)	(0.062)		(0.094)	(0.105)
Significance of difference:	*p ≈ 0.16*	*p < 0.001*	*p ≈ 0.005*		*p ≈ 0.042*	*p < 0.001*
(excluding n=2)	*p ≈ 0.94*	*p < 0.001*	*p ≈ 0.12*		*p ≈ 0.21*	*p ≈ 0.002*
Observations	125	108	125	108	125	108
|ln(L)|	90.40	54.26	35.19	28.83	26.54	26.79

* (**, ***): Significant at 10% (5%, 1%) level.
Additional explanatory variables: lottery choice, S question average, I question average, G question average (all group means).

Replication of Table 7, but including dummies for all group sizes:

Table E3: Tobit results for C choice fraction and group payoff efficiency, probit results for cooperative (all-C) Nash equilibrium outcomes, group-level data – estimated average marginal effects (unless otherwise stated) and standard errors

	(1c)	(2c)	(3c)	(4c)	(5c)	(6c)
Dependent variable:	C choice		Cooperative outcome		Payoffs	
group size = 3	−0.272***	−0.038			−0.248***	−0.051
	(0.068)	(0.068)			(0.050)	(0.047)
group size = 4	−0.259***	−0.233***			−0.357***	−0.321***
	(0.071)	(0.071)			(0.047)	(0.051)
group size = 5	−0.260***	−0.237***			−0.320***	−0.300***
	(0.077)	(0.077)			(0.052)	(0.055)
group size = 7	−0.292***	−0.206**			−0.296***	−0.283***
	(0.080)	(0.080)			(0.058)	(0.061)
group size = 15	−0.244***	−0.242***			−0.376***	−0.336***
	(0.082)	(0.081)			(0.051)	(0.058)
Leader treatment…	0.091	---			0.185***	---
	(0.057)				(0.039)	
Group receives C message…	---	0.333***			---	0.218***
		(0.048)				(0.040)
…when group size = 2	0.301***	0.354***			0.296***	0.314***
	(0.093)	(0.074)			(0.061)	(0.058)
… when group size = 3	−0.042	0.572***			0.312***	0.598***
	(0.137)	(0.070)			(0.093)	(0.058)
… when group size = 4	0.148	0.292**			0.079	0.054
	(0.142)	(0.134)			(0.102)	(0.117)
… when group size = 5	−0.101	0.087			0.147	0.102
	(0.156)	(0.157)			(0.114)	(0.130)
… when group size = 7	0.009	0.461***			0.160	0.103
	(0.175)	(0.152)			(0.126)	(0.162)
… when group size = 15	0.055	0.161			−0.068	−0.111
	(0.171)	(0.170)			(0.118)	(0.131)
Significance of difference:	$p \approx 0.17$	$p \approx 0.022$			$p \approx 0.053$	$p < 0.001$
(n=2 vs. n=15 only)	$p \approx 0.20$	$p \approx 0.30$			$p < 0.001$	$p < 0.001$
Observations	125	108			125	108
|ln(L)|	89.48	50.02			24.10	21.01

* (**, ***): Significant at 10% (5%, 1%) level.
Additional explanatory variables: lottery choice, S question average, I question average, G question average (all group means).

Table E4: Goodness-of-fit of regression models in main text and this appendix

			|ln(L)|	N	k	BIC	AIC
C Actions	Unconditional	Main	94.674	125	8	***227.98***	206.59
		Linear	100.91	125	8	240.44	219.05
		Dummies	89.478	125	16	256.21	215.99
		Main + n=2 dummies	90.403	125	10	229.09	***202.74***
	C message	Main	54.91	108	8	***147.28***	***127.27***
		Linear	59.798	108	8	157.05	137.05
		Dummies	50.024	108	16	174.96	138.03
		Main + n=2 dummies	54.262	108	10	155.35	130.79
All-C outcomes	Unconditional	Main	35.69	125	8	***110.01***	***88.622***
		Linear	37.864	125	8	114.35	92.97
		Dummies					
		Main + n=2 dummies	35.187	125	10	118.66	92.303
	C message	Main	29.17	108	8	***95.797***	***75.795***
		Linear	30.185	108	8	97.826	77.824
		Dummies					
		Main + n=2 dummies	28.831	108	10	104.48	79.931
Payoff efficiency	Unconditional	Main	29.754	125	8	***98.134***	76.749
		Linear	43.33	125	8	125.29	103.9
		Dummies	24.105	125	16	125.46	85.246
		Main + n=2 dummies	26.542	125	10	101.37	***75.015***
	C message	Main	29.019	108	8	***95.496***	***75.493***
		Linear	42.378	108	8	122.21	102.21
		Dummies	21.005	108	16	116.92	79.988
		Main + n=2 dummies	26.786	108	10	100.39	75.84

Notes: N = number of observations, k = number of free parameters, BIC = Bayesian information criterion (Swartz, 1978), AIC = Akaike (1974) information criterion. Emphasised values for BIC and AIC correspond to best-fitting model according to that criterion.