Not Just Babble: Opening the Black Box of Communication in a Voluntary Contribution Experiment

by Olivier Bochet and Louis Putterman*

Abstract

We let subjects in a voluntary contribution experiment make non-binding numerical announcements about their "possible" contributions and, in some treatments, send written promises to contribute specific amounts. We find that announcements were responded to both by others' announcements and by real play, for example by eliciting costly punishment when found to be misleading. We also find that adding pre-play announcements to treatments with punishment can increase efficiency by letting cost-free warnings substitute for costly punishment. The threat of punishing false announcements and promises helps reduce false signals, but only when promises are possible is the effect sufficient for achieving higher efficiency.

JEL Classification numbers: C91, H41, D23.
Keywords: Public goods, collective action, communication, punishment, cheap talk

Correponding author: Prof. Louis Putterman, Department of Economics, Box B, Brown University, Providence, RI 02912 USA. Tel: 401-863-3837; Fax: 401-863-1970;e-mail: Louis_Putterman@Brown.Edu.

---

* Maastricht University, CORE and Brown University. Send correspondence to o.bochet@algec.unimaas.nl. or to Louis_Putterman@Brown.Edu. Funding for the experiments reported here came from a grant from the MacArthur Foundation Network on Norms and Preferences, and from National Science Foundation Grant SES-0001769. We wish to thank Toby Page for collaborating on the design of the experiments. We wish to thank Arno Riedl for helpful discussions and comments. We thank again the many students who assisted in carrying out the experiments reported in Bochet, Page and Putterman (2006), as well as Dennis Zachary Schubert and Dmitri Lemmerman, who programmed the new treatments.

Not Just Babble: Opening the Black Box of Communication in a Voluntary Contribution Experiment

by Olivier Bochet and Louis Putterman

0. Introduction

Many experiments have been conducted to study to what degree and under what conditions individuals free ride in the voluntary provision of a public good. The question is much studied in economic theory and is relevant to problems ranging from the making and soliciting of charitable contributions to environmental protection, provision of effort in partnerships, and work in teams. In experiments, contributions typically begin at an average level well above predicted full free riding, in fact at more than 50% of subjects' endowments, but they decline steadily with repetition. Mechanisms that have been found to reduce free riding include taxing low contributors and rewarding high ones (Falkinger, Fehr, Gächter and Winter-Ebmer, 2000), allowing subjects to impose costly earnings reductions on one another (Fehr and Gächter, 2000a), and excluding free riders from playing with more cooperative subjects (Gunnthorsdottir, Houser, and McCabe, 2007; Önes and Putterman, 2007, Gächter and Thöni, 2005). However, in a review of 37 VCM, prisoners' dilemma, and other social dilemma studies, Sally (1995) found pre-play communication to be the single most effective way to promote cooperation, and in a direct comparison under controlled conditions, Bochet, Page and Putterman (hereafter BPP, 2006) found not only that pre-play face-to-face communication increased contributions and earnings far more than did opportunities to sanction, but also that it was so effective that adding sanction opportunities to it led to no further improvement in outcomes.[1]

BPP also reported VCM experiments with two other kinds of communication. First, we conducted chat room treatments in which subjects could communicate with the members of their group on line only, while maintaining anonymity as to who was in one's group. Second, we carried out "numerical cheap talk" treatments in which subjects (also anonymous to one another) could announce, by typing a number, a "possible amount" that

[1] Ostrom, Gardiner and Walker (1992) also provided either communication or punishment opportunities or both in common pool resource games. Their findings are similar except that the combination of communication and punishment achieves better results than communication alone, in their experiment.

they might contribute to the group account. Chat room treatments yielded higher levels of cooperation than baseline, despite the unavailability of facial expression, vocal intonation, and body language as means of conveying intention, emotion or other information. Unlike the treatments that allowed the exchange of written messages, however, the numerical announcement treatments did not enhance average cooperation relative to the no communication baseline. A similar experiment using numerical announcements, by Wilson and Sell (1997), also found them to be ineffective at engendering cooperation.

This paper attempts to shed further light on the difference of outcomes between numerical and verbal communication. We ask two main questions. First, is numerical communication truly cheap talk in the sense of being discounted by both senders and receivers, thus amounting to ineffective babble? Second, are Sally (1995) and BPP correct in their conjecture that a major reason for the efficacy of verbal communication is the ability to issue promises?

We explore the first question by carrying out a microanalysis of the data from BPP's numerical cheap talk treatment. We demonstrate that the patterns of numerical signals sent by subjects are far from random, and that the indifferent average results of numerical signaling mask a dispersion of outcomes that includes both groups that achieved greater cooperation than their most successful counterparts in treatments with no communication and groups that, due to opportunistic reliance on false signaling by some members, achieved even less cooperation than their least successful no communication counterparts. Both the coordination successes and the false signaling "disasters" indicate that subjects took numerical announcements as something other than cheap talk in a sense that we elaborate below.

We explore the second question by adding treatments in which subjects can elect to send non-binding promise statements as a follow-up to their numerical signals. Our results show that adding only a promise option had no impact on the level of average contribution compared to the baseline treatment. However, the option of sending promise messages significantly improved outcomes in those groups whose members were also given the opportunity to sanction one another with costly punishment. We perform the same type of microanalysis of promise treatment behaviors as of those in the simpler numerical cheap talk treatments. The data exhibit similar patterns, but we find that individuals paid more attention to promises than to announced contributions, and that when punishment was allowed, broken promises were severely punished and thus the percentage of false promises declined with repetition, allowing promises to grow in credibility.

The paper proceeds as follows. In Section 1, we discuss the literature on public goods games and communication. Section 2 lays out the design of the public goods experiments with and without numerical communication, sanction option, and promise option, and discusses theoretical predictions. Section 3 analyzes the experiments, emphasizing the impact of communication content on both others' communication and on binding decisions. Section 4 concludes the paper.

1. Voluntary Contribution Experiments and Communication

The voluntary contribution mechanism is an n-person linear public goods game with the following structure. In each of one or more periods (we focus on games of finite repetition), each of $n \geq 3$ individuals is endowed with a certain number of dollars, E, and must divide this between a private account and a group or public account. Money put in the group account is multiplied by a factor λ (where $n > \lambda > 1$) and divided equally among the n group members. The earnings of member i in a given period are

$$y_i = (E - C_i) + \lambda \sum_{\text{all } j}(C_j)/n \qquad (1)$$

where C_i ($0 \leq C_i \leq E$) is individual i's contribution to the group account and the summation is taken over all group members, i included. Equation (1) shows that all group members are better off if all contribute their full endowments to the group account than if they contribute nothing, but each individual is better off still if the others contribute but he does not. Efficiency, defined as the sum of earnings, is also highest when all contribute their full endowments. We focus on the symmetric case in which each has an equal endowment and information about endowments and payoff functions is common knowledge.

In a finitely repeated VCM game, the only subgame perfect equilibrium for rational individuals who care only to maximize their own payoffs and who have common knowledge of one another's preferences (including knowledge of one another's knowledge of this) is $C_i = 0$ for each i. While an outcome having $C_i = E$ for all players dominates it, there is no credible way to punish deviations from an agreement to contribute E, so communication cannot in principle alter the outcome.[2] Communication can change the outcome only if (a) payoffs can be altered in a manner external to the game, for example if any agreement reached by communication can be supported by the threat of penalties imposed by a third party, such as the state, or if (b) we drop the common knowledge and/or payoff maximization assumptions, allowing that some players' objective functions don't coincide with their material payoffs and/or that some players entertain the belief that players with such preferences might be present.

Isaac and Walker (1988) found that pre-play communication led their experimental subjects to contribute considerably more to a public good. Their study is one of 37 papers reporting 130 experimental treatments whose results Sally (1995) entered in multivariate regressions investigating which treatment variables best account for differing levels of cooperation and free riding. Sally concluded that face-to-face communication was the single most effective of the treatment variables, which also included the size of the monetary gain from cooperation, the number of repetitions, the discipline from which student participants were drawn, and the use of suggestive instructions by the experimenter.

To understand better what lies behind the effects of face-to-face communication, Brosig, Ockenfels and Weimann (hereafter BOW, 2003) and BPP conducted additional VCM experiments in which other forms of communication were substituted for face-to-face discussion. BOW 's comparison treatments included a no-communication baseline, a treatment with audio and visual communication from separated compartments, a treatment with only audio communication from separated compartments, and a treatment in which subjects could view one another on video terminals but could not communicate, prior to making their decisions. BPP's alternative treatments likewise included a no-

[2] A strategy of cooperating if others cooperate unravels because it is in no player's interest to cooperate in the last period.

communication baseline, but in addition, we conducted a treatment in which subjects could communicate text messages in a chat room and one in which subjects could relay non-binding possible choices in numerical form, with time for iterative reactions before each binding decision stage. For each of the four communication variants, BPP conducted two kinds of public goods game—a standard VCM experiment, and one like Fehr and Gächter's having a "punishment" stage[3]

While some of BOW and BPP's treatments achieved almost equally large efficiency gains as did their face-to-face communication treatments, BPP's numerical cheap talk (NCT) treatments performed no differently on average than did our no communication baseline. In this paper, we reanalyze those treatments at the level of individual subject behaviors. We demonstrate that numerical announcements were taken seriously by group members, affecting both subsequent announcements and subsequent binding play. For example, in the treatment with numerical communication and punishment stages, subjects announced larger punishments of those who announced smaller contributions, the same relationship as exhibited in binding play, and those at whom such announced punishments were targeted responded by raising their announced contributions, a reaction also seen in binding play. Actual contributions are significantly correlated with both own and others' announced contributions, and actually contributing less than the amount announced tended to elicit actual costly punishment. A micro-analysis of announcements and binding decisions thus provides evidence that subjects did *not* understand their messages to be "cheap talk" in the full sense suggested by a theory of rational, self-interested agents with common knowledge of one another's type.

When discussing why pre-play communication raises cooperation, contrary to standard economic theory, but why this effect is observed only when that communication has an open-ended written or oral component, we conjectured that the ability to make

[3] BPP's chat room communication treatment without punishment resembled that of Frohlich and Oppenheimer (1998), except that those authors used e-mail messages, which don't provide a continuing record of messages to all group members. BPP's non-binding numerical communication treatment without punishment, which we labeled "numerical cheap talk," resembled the numerical pre-announcement treatment of Wilson and Sell (1997), except that our subjects could react to one another's announcements with new nonbinding announcements for a period of a minute or longer before making binding decisions, whereas Wilson and Sell's subjects could send only one announcement before each binding decision.

promises plays a major role in raising rates of cooperation in treatments with face-to-face, audio-video, and chat room communication. To test this conjecture, we designed additional treatments identical to NCT, except that after iterative numerical communication and before each binding choice, we let subjects select, or not, a message promising to contribute a specific amount to the group account. We analyze the resulting treatments, finding that allowing promises significantly increased both contributions and earnings, but only when punishment opportunities were available to keep promise-breakers in line.

2. Experimental Design and Predictions

We discuss decisions made in eighteen experimental sessions, in each of which sixteen (in four of the sessions, twelve) undergraduate subjects (a total of 272 subjects) made a series of contribution decisions in randomly assigned and anonymous groups of four that stayed together for a total of ten periods of play. Each period involved simultaneous decisions by each subject on contributing to a group account versus a personal account, described by equation (1) above, with $E = 10$ experimental dollars (hereafter *E*\$10) and $\lambda = 1.6$, so that group members earned *E*\$16 per period if they perfectly cooperated and *E*\$10 if they all contributed nothing.[4] Subjects were drawn from the entire Brown University undergraduate population (numbering some 5800 students), sat at terminals in a large room, and were unable to read one another's screens or to communicate except in the treatments and manners indicated below. Three sessions were devoted to each of six different treatments (see Table 1), of which the first four are also discussed in BPP. In the baseline (**B**) treatment, the entire session consisted of ten such decisions, after each of which subjects learned of one another's individual contribution decisions. In the punishment or reduction (**R)** treatment, each contribution decision was followed by a stage in which subjects learned of the contributions of each of the others in their group and had an opportunity to reduce the earnings of one or more group members at a fixed cost of *E*\$0.25 to the punisher per *E*\$1 of earnings loss to the person punished. Individuals were informed only of the reductions they themselves received, without knowing the identities of the punishers. Subjects in all four treatments were identified to

[4] An experimental dollar exchanged for 0.13 real dollars at the end of the session, and total earnings averaged about \$25 for a 90 minute session, including a \$5 participation fee.

one another only by letters B, C and D which were randomly reassigned each period (as in Fehr and Gächter, 2000a), to prevent tracking of individual behaviors and thus reduce the tendency to carry out vendettas.

In the "numerical cheap talk" (**NCT)** and "numerical cheap talk with reduction opportunities" (**NCTwR**) treatments, so referred to because of the expectation of standard theory that announcements would amount to no more than cheap talk, each set of binding contribution and reduction decisions was preceded by a period of announcements and amended announcements. During these periods in the **NCT** treatment, subjects simply entered an amount in the group account assignment box of a screen identical to the binding decision interface (reproduced in BPP) but for the heading "Communication Stage" and a different background color. Once each had entered some number and the four numbers were displayed to each group's members, they were free to alter their announced numbers for up to 90 seconds (a smaller amount of time in later periods).[5] In the communication stages of the **NCTwR** treatment, subjects first entered possible contribution amounts, then, viewing the amounts entered by each group member, entered possible reduction amounts. Once each subject saw the four contribution announcements and the total reduction announcements from others, each was free to alter either her announced contributions or her announced reductions of others' earnings for up to 90 seconds (again, a smaller amount of time in later periods).[6]

After finding that average contributions and earnings in the **B** and **NCT** (the **R** and **NCTwR**) treatments were not significantly different, we designed a variation on the **NCT** and **NCTwR** treatments as a partial test of our conjecture that this might be attributable to unavailability of a way to make verbal commitments or promises. The new treatments are identical to the old ones, including a period of iterative numerical communication of non-binding "possible" choices. However, at the end of each period's numerical communication stage and before its binding contribution stage, subjects were asked to choose between two statements. The first option read: "I promise to contribute __ to the

[5] The instructions about this were: "At the beginning of each period there will be a communication stage. You will type in a possible amount for your group assignment. … at any time thereafter during this stage, you can adjust your possible assignment to the group account. … You are not committed to any of the numbers you type in during this stage.
[6] The full instructions given to the subjects, including practice problems, are provided in Bochet *et al*. (2005).

group account this period." and it required the choice of an integer in the 0 to 10 range, if selected. The other option read: "I do not wish to make a promise at this time." Depending upon the choice of the subject, the other group members would then be shown either the statement "A promises to contribute X to the group account." (where X = 0,1,…10) or "A chooses not to make a promise," and likewise for subjects B, C and D. The instructions given the subjects refer to "choosing a statement" rather than to "making a promise." We label the promise-including analogue of the **B** and **NCT** treatments **NCTwP**, and the promise-including analogue of the **R** and **NCTwR** treatments **NCTwP&R**.

A dilemma for us in designing these treatments was what, if anything, to tell the subjects about whether a promise was binding. If there were no statement about this and if a high proportion of subjects contributed the amounts typed into their promise statements, we would be unable to rule out the explanation that they adhered to their promises because they understood the rules of the experiment to *require* them to do so. To rule out the possibility that promises were effective because of such a misunderstanding, we included in the instructions about entering binding decisions the statement "If you have chosen to promise a specific amount, you can type that amount at this time, but the computer will not prevent you from typing in a different amount." This statement carried its own danger, because it may have been viewed as the granting of "permission to lie" by the experimenter; in fact, when instructions were being read aloud, there were a few chuckles or raised eyebrows among the subjects at this point in every session. Our results must be read, then, with an awareness of the downward bias against the efficacy of promises that we may have introduced by stating that promises were not binding.

[Table 1 about here]

Standard economic theory assuming strictly payoff-maximizing agents with common knowledge of this preference predicts that subjects will contribute nothing to the group account in the **B** treatment. As pointed out by Fehr and Gächter (2000a), standard theory also predicts that the opportunity to engage in costly punishment will not be made use of and will have no effect on the level of contributions, which will still be uniformly zero. The addition of an opportunity to enter a possible contribution or a possible contribution and possible reduction decisions into a non-binding communication field would also have no effect according to theory, assuming common knowledge of payoff maximizing type. Under that common knowledge assumption, each agent realizes that each other agent will contribute nothing to the group account and spend nothing on punishment, regardless of what numbers are communicated, and there is therefore no reason to pay any attention either to the numbers typed by others or to the numbers that one types oneself. A pure numerical babble is always an equilibrium.[7]

Suppose, instead, that our subjects believe that another type of agent, whose objectives include but are not limited to earning more money, is present in the subject pool with some non-negligible probability. Although unconditional altruists and individuals who experience a "warm glow" from contributing are among the potentially interesting possibilities (see Palfrey and Prisbrey, 1997, and references therein), we focus here on three possible "nonstandard" preferences: positive and negative reciprocity, and truth-telling. A positively reciprocating agent is one who prefers to cooperate if he believes that others are cooperating. A negatively reciprocating agent is willing to incur a cost to punish someone who exploits him by free riding.[8] An agent with a preference for truth-telling can be seen either as obtaining additional utility from adhering to her word, or as suffering a loss of utility if she breaks her word.[9] Finally, suppose that subjects begin with some prior beliefs about the proportions of such subjects who are present and adjust their choices during the course of play as they update those beliefs. Subjects thus enter into a Bayesian game of the type analyzed by Kreps *et al*. (1982) and Guttman (2003).

[7] As shown in Farrel-Rabin (1996), conflict of interest among agents means that messages cannot be self-signalling or self-committing. In such a case, if agents maximize their payoffs and types are common knowledge, it is always consistent to treat cheap talk as meaningless. The finding that the messages do not seem to be babble, in our experiment, implies that the subjects *do not* believe all to be rational payoff maximizers.

[8] See Fehr and Gächter (2000b) and Hoffman, McCabe and Smith (1998), who treat conditional cooperation and willingness to punish noncooperation as two sides of the same trait. Önes and Putterman (2007) consider that the relative degrees of positive and of negative reciprocity may differ from one reciprocator to another.

[9] Sánchez-Pagés and Vorsatz (forthcoming) find evidence for the presence of subjects with what they deem to be two distinct social preferences, lie-aversion and preference for truth-telling, in sender-receiver games with conflicting preferences. Charness and Dufwenberg (2006) suggest that players may keep their word due to a preference for living up to others' expectations. Bicchieri (2005) discusses the activation of a norm of promise-keeping.

If a group of reciprocators with optimistic expectations of one another's type are grouped together in a basic VCM experiment such as our **B** treatment, it is possible that they will contribute all or most of their endowments on the first decision and that, with their favorable beliefs thus supported, they will continue to contribute most of their endowments (Gunnthorsdottir *et al.*, 2007; Gächter and Thöni, 2005). Probably more typical is an encounter of subjects with differing degrees of reciprocity and differing initial beliefs. Upon seeing some low contributions, the reciprocators in such a group will begin to reduce their contributions to the group account, in the **B** treatment, leading to the gradual downward slide that is usually seen in finitely repeated VCM experiments. Allow the reciprocators to punish the free riders while maintaining their own high contributions, however, as in our **R** treatment, and contributions may stabilize or rise rather than fall, as is found by Fehr and Gächter (2000a), Masclet, Noussair, Tucker and Villeval (2003), BPP, Page, Putterman and Unel (2005), and Sefton, Shupp and Walker (2002).

Consider now what communication might add to the Bayesian story. The opportunity to send numerical signals wouldn't necessarily be viewed as useless by subjects who believe reciprocators or truth-tellers are common. Suppose, for example, that a substantial proportion of subjects are reciprocators and truth-tellers, and that all subjects know this to be so, although they don't know which individuals are and which are not of these types. Then subjects with the relevant preferences might, by entering a high number, try to signal intentions to contribute their endowments conditional on others doing so, and if others seem to signal a similar intention, they might proceed to contribute in fact and see whether the others follow through. If the game includes punishment opportunities, the reciprocators might signal intentions to punish low contributions, and some might follow through with actual punishment when they see evidence of attempts to mislead by contributing less than announced. Opportunistic subjects whose only goal is to maximize their payoffs might also signal and act cooperatively in some periods, with the intention of later exploiting the credulity of fellow players by signaling an intention to contribute but not following through.

The predictions of the Bayesian model which allows for nonstandard player types and of the standard model with common knowledge of universal payoff maximizing type are clearly quite different. The standard model implies that "numerical cheap talk" will be of no consequence, and might even be a stream of random numbers. The Bayesian model suggests that we should look for signs of attempts to coordinate, on the parts of some subjects, and of attempts to mislead, on the parts of others. BPP's analysis, which showed that outcomes in the **NCT** and **B** treatments, on the one hand, and in the **NCTwR** and **R** treatments, on the other, were on average indistinguishable, is apparently consistent with the standard prediction regarding communication; but it doesn't rule out the Bayesian one. A more micro-analytic look at messages, and at inter-group differences in outcomes, is required in order to see whether numerical cheap talk was really babble or was instead a flow of meaningful messages between subjects who viewed one another's preferences as an open question.

3. Results and Analysis

a. Contributions and earnings trends

Figure 1 shows the average number of dollars contributed to the public good in the six treatments, by period.[10] The pattern of contributions in the **B** treatment conforms well to expectations from the literature: a substantial average initial contribution followed by a generally declining trend.[11] In the **R** treatment, as in the similar treatment in Fehr and Gächter (2000a), contributions show no tendency to decay until the end of the session, a pattern which analysis shows to be attributable, at least in part, to the tendency of many subjects to impose costly punishment on low contributors.[12] This tendency is not significantly less in evidence in the last period, suggesting that it is indeed attributable to a taste, rather than being undertaken to raise future earnings.

Average contribution and its trend in the **NCT** and **NCTwP** treatments resemble closely those of their no communication and no promise counterpart, the **B** treatment, and

[10] Average contribution and earning trends in the **B**, **R**, **NCT**, and **NCTwR** treatments are shown in BPP and are reviewed here before our more detailed analysis, and for comparison with the new treatments, **NCTwP** and **NCTwP&R**.

[11] Past results are surveyed in Davis and Holt (1993) and in Ledyard (1995).

[12] The failure of contributions to rise as steeply in the **R** treatment as they do in Fehr and Gächter's punishment condition could be due to minor differences in design. A similar tendency for the introduction of a punishment stage to stem the usual decaying trend but without significant upward trend is also found in other replications of Fehr and Gächter, for example Carpenter and Matthews (2002). See also Nikifarakos and Normann (2005), who find that the trend of contributions depends on the effectiveness of punishment (that is, cost to the person punished).

the same holds when comparing the **NCTwR** treatment to its counterpart treatment, **R**. Mann-Whitney tests confirm that average contributions over the ten periods as a whole do not differ significantly as between the **NCT, NCTwP** and **B** treatments, or as between the **NCTwR** and **R** treatments (see Table 2).[13] It appears on first inspection, then, that giving subjects the opportunity to announce possible decisions before each set of binding decisions made no difference to contributions. However, the combination of punishment opportunities, opportunities to make numerical announcements, and opportunities to send promise statements raised contributions in the **NCTwP&R** significantly above those in the **NCTwR** and **R** treatments, hence also above those of the other three treatments.

Result 1. Apart from the ***NCTwP&R*** *treatment, average contributions and their trend over time differ significantly only between treatments without punishment opportunities and those with such opportunities. Contributions are higher on average and are more sustained in treatments with punishment. But among treatments with punishment, contributions are significantly higher in the* ***NCTwP&R*** *treatment.*

Figure 2 shows average earnings in the six treatments and their trend over time. Neither availability of punishment nor that of numerical communication significantly affect average earnings, nor does adding promise opportunities alone do so.[14] Even though contributions are more sustained in treatments with punishment, earnings are not higher in those treatments, because the costly punishment that induces the higher contributions cancels out the associated earnings gain.[15] Earnings are, however, higher in the **NCTwP&R** treatment, where higher contributions are attained with less punishment (as will be shown later).

[13] Since group members had no knowledge of what was occurring in any of the other groups in their session or other sessions, and group behaviors are thus statistically independent, our tests use the group average contribution or the group average earning level of a subject, averaged over all 10 periods, as observations.

[14] Earnings in **B** exceed those in **R** but only at the 10% level in a two-tailed test. Apart from this, the tests find no difference in earnings among any of the five treatments.

[15] Cinyabuguma, Page and Putterman (2004) show that about 20% of earnings reductions in the **R** treatment were aimed at high rather than low contributors, and that absent these reductions and the immediate lowering of contributions that they led to, earnings would have been higher in the **R** than in the **B** treatment. Ertan, Page and Putterman (2005) show that earnings rise unambiguously compared with baseline treatments if only punishment of low contributors is permitted (as is the case in their experiments when subjects vote on what types of punishment to permit).

*Result 2. Apart from the **NCTwP&R** treatment, average earnings do not differ significantly between treatments. Average earnings are higher in **NCTwP&R** than in all other treatments.*

[Figures 1 and 2 about here]
[Table 2 about here]

b. Differing dispersions of group outcomes

Although there is little difference between the **B** and **NCT** (**R** and **NCTwR**) treatments with respect to average behaviors, we noticed that there was greater variation among groups in treatments with **NCT**. Figure 3 shows the variance of average contribution among groups in each treatment, by period. In the left panel, containing the treatments without reductions, the variances are noticeably higher for the two treatments that include **NCT**, in all periods. In the right panel, containing the treatments with reductions, variances are higher for **NCTwR** than for **R** in all but one period, but the variance in contributions is lowest in **NCTwP&R** except in two periods. Table 3 confirms that on average, the variance among groups is higher in the **NCT** and **NCTwP** treatments than in the **B** treatment, and higher in the **NCTwR** treatment than in the **R** treatment. Formal statistical tests are ruled out unless one considers the variance in each period to be independent of that in other periods; if we assume that they are, the variances are found to be significantly different.[16]

[16] We performed both Mann-Whitney tests, which treat each of the ten variances for each treatment symmetrically, and Wilcoxon tests, which pair the variances of two treatments to be compared for period 1, for period 2, etc. As Table 3 reports, both tests conclude that variances are greater in **NCT** compared to **B,** in **NCTwP** compared to **B,** and in **NCTwR** compared to **R**, significant at the 1% level. As an additional check, we graphed the average contributions of the three highest-contributing and of the three lowest-contributing groups in each pair of treatments, and found high (low) performers contributing more (less) in **NCT** and **NCTwP** than in **B** and in **NCTwR** than in **R** in almost every period. Using either average contribution or average deviation from treatment average contribution or both, the differences are mildly statistically significant (despite having only 3 observations per group) for high groups in **NCT** vs. **B** and **NCTwR** vs. **R**, and for low groups in **NCTwR** vs. **R**. Both the highest and lowest performing groups in **NCTwP&R** contributed more than (for highest groups: more than or the same as) their **R** and **NCTwR** counterparts in every period, however, with these differences also being significant. See Figures A1 and A2 and test details in the Appendix, which is available at http://www.econ.brown.edu/fac/Louis%5FPutterman/working/pdfs/AppendixNotJustBabble%204-24-07.pdf.

Result 3: Except when combined with promises and reductions, the introduction of numerical announcements increased the variance of average contributions per period across groups.

[Table 3 about here]

c. Subjects' announcements responded to others' announcements

What can account for the greater variance of outcomes in groups with numerical communication? One possibility is that groups having members more inclined towards truth-telling and reciprocity achieved and sustained greater cooperation, whereas groups with higher proportions of opportunistic individuals had worse outcomes when announcements were possible than when they were not. For this conjecture to hold true, subjects would have to take announcements as more than random babble. The next three results will demonstrate this, beginning by showing that subjects responded to one another's announced plans by changing their announcements, in much the same way as they responded to one another's binding decisions by changing their subsequent binding decisions, in treatments *without* communication.

Tables 4a – 4d report regression and non-parametric correlation tests of the proposition that subjects who found themselves to be contributing more (less) than their group's average tended to reduce (increase) their contribution—whether it be the actual binding contribution, changing from period to period in the **B** and **R** treatments, or the announced possible contribution, changing from iteration to iteration in a given period's announcement stage. Tables 4a and 4b report the results for changes over time in binding contributions, while Tables 4c and 4d report the results for changes over communication iterations in announced contributions. Tables 4a and 4c report panel corrected standard error regressions, with and without individual and period fixed effects. The change in (announced) contribution is the dependent variable, and in addition to the difference between own and others' previous (announced) contribution, the subject's own previous (announced) contribution is included.[17] For both actual contributions (Table 4a) and

[17] Subjects who entered only one contribution announcement during a given period's communication stage are treated as having an announcement change of zero.

announcements (Table 4c), the coefficients on the contribution difference are always negative, supporting the proposition, significant at the 5% level or better in all but one specification of each table. Adjustment of contributions toward group average may be a reflection of conditionally cooperative behavioral tendencies.[18]

Because the contributions of members of a given group cease to be fully independent variables after the first period, due to their revelation to the group members, the assumptions supporting the significance tests on regression coefficients may not be fully valid, despite the use of fixed effects and corrected standard errors. Therefore, in Tables 4b and 4d we present results of non-parametric Spearman correlation tests of the correlations between the key variables: changes in each subject's (announced) contribution, and difference between own and others' contribution in the previous period (announcement). The observations of a given period are tested separately, with those of period 10 left out due to possible end-game effects. All but one test for the actual contributions (Table 4b) are supportive at the 5% level of significance, while all of the tests for the announcements (Table 4d) are supportive at the 1% level.

Result 4. During communication periods, subjects in the **NCT** *treatment adjusted their announced contributions in the direction of the average announced contributions of other group members.*

[Tables 4a, 4b, 4c and 4d about here]

A second common response seen in binding play is that when a group member contributes a substantially smaller amount than others, he or she tends to be targeted for punishment.[19] We next check whether a parallel phenomenon is found in the interactions between announcements. The first column of Table 5 reports a random effects tobit regression in which the amount of announced reductions aimed at subject *j* is the dependent variable, and the absolute negative and positive deviations of *j*'s announced contribution

[18] Fischbacher, Gächter and Fehr (2001) find that 50% of their subjects contribute more when they expect others also to do so, with a further 14% also doing this so long as average contributions remain below half of the endowment. Similar results are reported by Kurzban and Houser (2001).
[19] See Table 5 in Fehr and Gächter (2000a), Table 3 in Cinyabuguma, Page and Putterman (2006), and Table 2 in Önes and Putterman (2007).

from the average of other group members, and that average itself, are independent variables.[20] The second column reports an otherwise identical tobit regression with group fixed effects. The coefficient on absolute negative deviation is positive and significant at the 1% level, while that on absolute positive deviation is negative and significant at the 10% level, indicating that subjects were assigned more (less) announced punishment the further below (above) the average was their announced contribution. For comparison, the table's third and fourth columns show parallel regressions, also using data from the **NCTwR** treatment, but this time data on actual, as opposed to announced, contributions and reductions. We find that the interactions of announced decisions follow a closely similar pattern to the interactions of actual decisions. The parallelism of the two regressions defies the idea that numbers communicated would amount to random noise. Due to the potential problem of non-independence of observations, we also conducted period by period non-parametric tests of the relationship between announced contribution and announced punishment, the results of which appear in the top row of Table 5b. These tests find a correlation between the absolute negative deviation of the announced contribution, and announced punishment, significant at the 1% level.[21] Corresponding tests of the correlation between binding punishment and the absolute negative deviation of *i*'s binding contribution appear in the second row of Table 5b; five of the nine correlations are significant at the 1% level and a sixth at the 5% level.

Result 5. The relationship between announced punishment and announced contribution replicates that between actual punishment and actual contribution: the less one contributes relative to others in one's group, the more one is targeted for punishment.

[Table 5a, 5b and 5c about here]

---

[20] Tobit estimation is used because there are numerous cases of zero punishment which constitute potentially censored observations. In particular, 251 of the 440 observations in the announced punishment regression and 282 of the 440 observations in the actual punishment regressions have zero values of the dependent variable. Following Fehr and Gächter, the negative deviation variable is assigned a value of zero if *j* contributed more than the average of other group members, and likewise for the positive deviation variable if *j* contributed less than the others' average.

[21] The last two columns of Table 5a are discussed in conjunction with Result 8.

Our last demonstration that subjects adjusted their announcements in response to one another's announcements in a fashion paralleling changes in binding decisions involves the influence of announced punishment on subsequent announced contribution. Cinyabuguma, Page and Putterman (2004, Table 2) and Önes and Putterman (2007, Table 3) show that in finitely repeated partner treatments like those studied here, low (high) contributing subjects who receive punishment tend to increase (reduce) their contributions in the next period. Table 6a shows panel corrected standard errors regressions in which the dependent variable is subject *i*'s initial change of contribution announcement during each communication period. The independent variables are the number of dollars by which others announce that they might reduce *i*'s earnings, interacted with dummy variables to distinguish those announcing their group's highest contribution from others.[22] Three specifications are estimated for the **NCTwR** treatment, three for **NCTwP&R**, differing in inclusion of fixed effects for individuals and/or periods. All of the first coefficients are positive and highly significant, suggesting that those who announced a less-than-maximum contribution increased their announcement by an average of around 25 (16) cents per dollar of announced punishment "received" in the **NCTwR** (**NCTwP&R**) treatment, a reaction qualitatively identical and quantitatively similar to that found for actual contributions following actual punishment. The second coefficient is consistently negative, suggesting that a targeted high contributor might slightly reduce her announcement, as also found for contributions in the studies cited above, but this coefficient is less statistically significant in **NCTwR**, and insignificant in **NCTwP&R**. Corresponding non-parametric tests of the correlation between the first change of *i*'s announced contribution and the amount of punishment aimed at *i* if *i* is not the group's highest contributor are shown in Table 6b. For the **NCTwR** treatment, the correlations for all but one period are statistically significant, mostly at the 1% level. For the **NCTwP&R** treatment, only periods 1 – 3 and 9 show significant positive correlations between the two variables; the low and in one case negative correlations in other periods may in part reflect the high contributions and low incidence of punishment in this treatment.

[22] Hence, the first (second) variable is 0 if *i* is (not) the highest announced contributor.

Result 6. In response to announced possible punishment, subjects who announced low contributions tended to increase their announced possible contributions.

[Table 6a and 6b about here]

d. Subjects' binding decisions also responded to others' announcements

Did announcements influence other announcements only, or did they influence binding, costly decisions? We demonstrate that the latter is the case, beginning by showing that binding contributions are significantly influenced by announced possible contributions.

Table 7a shows panel corrected standard error regression estimates in which subject *i*'s contribution in period *t*, $t = 2,\ldots,9$, is the dependent variable, and the independent variables are the average contribution by the others in *i*'s group in period t - 1, the average last announced contribution of the others in the period *t* communication stage, and *i*'s last announced contribution in that stage.[23] As before, three specifications are estimated for each of two treatments, varying with regard to what fixed effects are included. Own announcement is significantly related to own actual contribution for both **NCT** and **NCTwR** subjects, suggesting a tendency by most subjects to make more-or-less truthful announcements (whether out of genuine aversion to lying or to contribute to others' beliefs—which may redound to one's own benefit—that honest types may be present. In both treatments, actual last period average contribution by others significantly and positively affects own contribution. Finally, average of others' most recent announced contribution has a positive coefficient in all of the regressions, is significant at the 10% level in one of the **NCT** and one of the **NCTwR** regressions, and is significant at the 1% level in the other two **NCTwR** regressions. Corresponding non-parametric tests for correlation between binding contribution of subject *i* and others' average announcements, and between binding contribution of subject *i* and the own announcement of subject *i*, are shown in Tables 7b and 7c, respectively. Almost all of the correlations are statistically significant, most of them at high levels.

[23] Period 1 must be excluded to allow for the lagged average contribution term, and period 10 is left out to exclude potential end-game effects.

Result 7. In the ***NCT*** and ***NCTwR*** *treatments, actual contributions in a period are positively related to own announced contribution and to the average of others' announced contributions.*

[Tables 7a, 7b, 7c about here]

A still more striking indication that "cheap talk" announcements had impacts on costly binding decisions is the finding that many subjects incurred real monetary costs to punish "lying" about contribution "intentions."[24] The last two columns of Table 5a report tobit regressions with the same specification as the middle two, except that the difference between *j*'s last announced contribution and his/her actual contribution in the same period is added as an independent variable. The new variable has a positive coefficient significant at the 10% level in the random effects tobit estimate and at the 5% level in the tobit estimate with group fixed effects. These estimates imply that for every one dollar of difference between announced and actual contribution, a subject received on average about 12 to 17 cents of punishment. This amount of punishment may not have sufficed to induce much more truth-telling, but it is important for our purposes because it demonstrates that rather than treating one another's announcements as noise, many subjects predicated costly decisions on them. Table 5c provides corresponding non-parametric test results for early and later periods of the **NCTwR** treatment. In all periods, the correlations are positive, and in five of them, they are statistically significant at the 10% level or better.

Result 8. Subjects received costly punishment for contributing less than their announced "possible" contribution.

[Table 8a and 8b about here]

A likely reason why the **NCTwP&R** treatment succeeded where the **NCTwP** treatment did not is that more subjects were deterred from using the promise option opportunistically in **NCTwP&R** since other group members had the possibility of inflicting monetary damage on them were they to contribute less than the promised amount. We've

[24] Recall that the instructions referred to each number communicated only as a "possible choice," which need not have suggested commitment or intention.

just seen that false announcements were punished in the **NCTwR** treatment, yet contributions and earnings in that treatment didn't exceed those in treatment **R**. Punishment of false promises in **NCTwP&R** could have made a decisive difference only if its magnitude was greater, which could conceivably be the case if broken promise statements engendered greater anger or indignation.

Table 8a reports a series of tobit regressions[25] resembling the last specification in Table 5, with the addition of a dummy variable controlling for the choice of a promise, and a few changes made necessary by unusually high correlations among certain variables.[26] Columns 1 and 2 contain the basic results paralleling Fehr and Gächter's, indicating that subjects received more punishment the further below others' average was their contribution.[27] In column 3 and 4, we add the dummy variable for contributing less than promised and find it to have highly significant positive coefficients, implying that a subject received an average of between 4.69 and 4.99 experimental dollars of punishment if she broke a promise. In columns 5 and 6, we use instead a dummy variable which takes the value of 1 if the subject contributed less than her last "numerical cheap talk" announcement; this term also obtains highly significant positive coefficients, although the indicated average punishment is smaller, between *E*$2.44 and *E*$3.48. Table 8b provides non-parametric tests of the correlations between punishment received by *i* and (a) the negative deviation of *i*'s contribution from the group average, (b) the dollars, if any, by which *i* underfulfilled the amount he/she promised,[28] and (c) the corresponding difference for underfulfilled announced contributions. Almost all correlations for all three variables are statistically significant at the 1% level.

In the NCTwP treatment, subjects chose the promise statement in 78.4% of opportunities to do so; for NCTwP&R, the corresponding figure is 86.8%. Most subjects

---

[25] As with Table 5, a tobit is used because of the large number of zero cases; here 336 of the 440 observations involve zero punishment and are thus potentially left-censored.

[26] To reduce multicolinearity, we substitute for the size of the deviations between actual and announced or promised contribution dummy variables equaling 1 if the individual contributed less than announced or promised, and zero otherwise.

[27] In two specifications of this table, absolute positive deviation also obtains significant positive coefficients. These are suggestive of the presence of some "perverse punishment," punishing of higher contributors by low ones that is found in other studies (see footnote 15 and the sources cited there). However, the corresponding coefficients are not significant in other estimates, including all of the table's fixed effects estimates, or alternative estimates in which we drop the "others' average contribution" term.

[28] Treated as 0 if *i* did not select a promise statement.

who broke a promise were simultaneously failing to fulfill an announcement, so tests including only one or the other variable are poor at pin-pointing which of these two was more important. Returning to Table 8a, in columns 7 and 8 both dummy variables are included. In the random effects specification of column 7, both variables have significant coefficients, but in the fixed effects specification of column 8, only the promise dummy is significant. In both cases, the coefficients suggest at least twice as much punishment for failing to fulfill a promise as for not doing so with an announcement.[29] These results strongly suggests that breaking a promise was likely to attract more punishment than was failing to fulfill an announcement. The magnitude of *E*$4.69 to *E*$4.99 per promise broken, estimated in columns 3 and 4, can be compared to the monetary gain from breaking a promise, which averaged only *E*$2.87 per episode,[30] suggesting that a promise-breaker usually achieved no net gain.

[Table 8a and 8b about here]

Result 9. Contributing less than the amount specified in a promise statement drew actual costly punishment in even larger amount than did contributing less than indicated as a "possible" amount.

Results 4 – 9 suggest that many subjects attempted to use nonbinding numerical announcements to coordinate on a more rewarding cooperative strategy. One reason why outcomes were not on average better in the treatments with numerical communication than in their counterpart treatments may be that in addition to such cooperation-seekers, there were also subjects who intentionally used misleading signals to improve their individual returns from free riding. We test this conjecture by testing whether groups in which there was less opportunistic "lying" about intentions had better outcomes than those in which there was more "lying". Our tests show that the abuse of announcements to mislead other subjects did indeed have a detrimental effect on cooperation.

---

[29] To be sure, there is a high correlation (.700) between the two dummy variables, reducing the reliability of the Column 7 and 8 estimates. Comparison of columns 3 and 4 with columns 5 and 6 provides some independent evidence that a broken promise's effect is the larger of the two.
[30] The overall average difference between amount promised and amount contributed, in those cases in which a promise wasn't fulfilled, was *E*$4.78, and the subject saved *E*$0.60 for each *E*$1 not contributed.

Let "lie" denote the difference between a subject's last announced contribution and her binding contribution in a given period, and calculate a group's "average lie" by dividing the sum of "lies" by the number of group members.[31] In Table 9a, we present a series of panel corrected standard errors regressions for the **NCT** and **NCTwR** treatments, with one observation per group and period, using average contribution in period *t* as dependent variable. Independent variables are the average "lie" last period (period *t*-1), the average last announcement in the current period (*t*), and in some specifications group or period fixed effects. The estimates support the idea that a previous period's "lying" reduced the current period's contributions, in both treatments. Non-parametric tests in Table 9b yield negative correlation coefficients in all periods, but most are insignificant, perhaps in part due to the small sample size.[32]

[Tables 9a and 9b about here]

*Result 10. Contributing less than announced led to lower future contributions in groups in the **NCT** and **NCTwR** treatments.*

A similar analysis can be done for the **NCTwP** and **NCTwP&R** treatments, except that here effects of both the average gap between announced and actual contributions, and the average gap between promised and actual contributions, can be studied. Table 10a reports panel corrected standard error regressions at group level that parallel those in Table 9a but include both lagged variables and both the average last announced contribution and the average last promised contribution.[33] The estimates for the **NCTwP** treatment support the idea that both failure to fulfill announcements and lying on promises significantly reduced subsequent contributions. Those for the **NCTwP&R** treatment indicate even

[31] Although the extent of an individual's "lie" might be defined as being equal to zero whenever his actual contribution exceeded his last announced contribution in the communication round, we let "lie" (in the few cases of this type) take negative values.
[32] Unlike tables 4b and 4d, 5b and 5c, 6b, 7b and 7c, and 8b, the test in Table 9b uses one observation per group rather than per subject. Observations from different periods are not pooled since those of a given group in different periods may not be statistically independent.
[33] The table's row headings denote a "lie" on announcement by '"lie" on A' and a "lie" on promise by '"lie" on P,' the variables being measured as contribution amount announced or promised minus actual amount contributed. '"Lie" on promise' is defined as 0 for subjects not choosing the promise statement, and 'average last promised contribution' is the average choice of those in a group selecting the promise statement, only, and is treated as 0 when no member of a group selected that statement.

stronger negative effects of lying on promises, but the coefficients for lie on announcement are inexplicably positive. The corresponding non-parametric correlation tests in Tables 10b and 10c show negative correlation coefficients between both lie variables and next period contribution in every period of both treatments, with the same problem of small sample size and limited significance as with Table 9b, but with a larger number of significant correlations for each variable in the **NCTwP&R** treatment (Table 10c).

Result 11. The greater the degree to which members contributed less than announced or promised, the smaller were average subsequent contributions in groups.

[Tables 10, 10b and 10c about here]

Together, results 10 and 11 suggest that differences in the extent of false announcements and promises help to account for differences in achieved cooperation. The "flip side" of the finding that "lies" undermined cooperation is that honesty promoted it.

*e. Punishment of promises kept "lying" in check in the **NCTwP&R** treatment*

The facts that false promises were heavily punished (Result 9) and that fulfilled promises led to greater cooperation (Result 11) can explain the higher contributions in the **NCTwP&R** treatment provided that subjects responded to the punishment of false announcements and promises by being more truthful in their communications. By making communications more credible, this could make more credible the threat that free riding would be met by actual punishment. If free-riding could thus be deterred by cost-free announcements rather than costly punishment, contributions would rise without offsetting punishment costs. And with assurance that others would contribute more, conditionally-cooperative subjects would also contribute more.

A series of figures demonstrate that the punishment of "lying" seen in tables 9 and 10 led to a closer correspondence between announcements and binding decisions on contributions in the **NCTwR** and especially in the **NCTwP&R** treatment. We also show that costly punishment declined with repetition even as the degree of "lying" fell, in the **NCTwP&R** treatment, which suggests that "cheap talk" warnings were largely sufficient to induce high contributions, causing earnings to rise along with contributions.

[Figure 4 about here]

Figure 4 has four quadrants, each showing, for one of the four treatments with numerical communication, the evolution during the 10 periods of play of the average contribution initially announced in a period, the average final announced contribution in the same period, the average actual contribution, and for treatments with promise option, the average contribution amount promised by those choosing the promise statement. For the **NCT** and **NCTwP** treatments, which lack punishment opportunities, we see announced possible contribution (and in **NCTwP**, promised contribution) tending to increase from first to last announcement (and promise) of a period's communication phase, but average actual contribution falling further and further below the announcements and promises as the experiment progresses. In the **NCTwR** treatment, we see no systematic tendency for promised amounts to either increase or decrease during a given communication stage, and the gap between promised and actual contribution grows only slightly larger with time, except for a last period jump. In the **NCTwP&R** treatment, finally, first and last announced and promised amounts tend to converge as the experiment progresses, and actual contributions show no tendency to diverge from promised and announced amounts until period 10 (or periods 9 and 10).

Using data at group level, we run MW tests to check whether the gap between average last announced and average actual contributions differs across treatments. We find that the average gaps are significantly lower for groups in **NCTwP&R** than for those in **NCT** and **NCTwP** treatments (2-tailed test *p*-values below 1%) but that the difference is less significant for groups in the **NCTwR** treatment (2-tailed test *p*-value equals 19%, hence 1-tailed test p-value a little under 10%).[34] As for the gap between average promise and average actual contribution, it is significantly greater in **NCTwP** than in **NCTwP&R** (*p*-value in 2-tailed group level MW test equals .001). The results are summarized in Table 11, below. Because gaps may not be entirely independent across periods, we also ran the same type tests but for each period separately. We find similar results –the tables are

[34] In absolute terms, the gap averages 0.62 in **NCTwP&R** versus 1.11 in **NCTwR**

available upon request but not reproduced here. Interestingly, for each pair of treatments in which there is a significant difference, the associated per period tests are significant starting in Period 3 but not before.

[Table 11 about here]

[Figure 5a about here]

Figure 5a graphs by period, for the **NCTwP** and **NCTwP&R** treatments, the proportion of subjects who failed to contribute as much as their last announcement and the proportion (of those choosing a promise statement) who failed to contribute as much as promised. It shows a dramatic difference between the two treatments, with the proportion of subjects not fulfilling their announcements and promises starting slightly higher and rising steeply toward about 70% in the **NCTwP** treatment, whereas those proportions stay in the neighborhood of around 20% in the **NCTwP&R** treatment, with a small last period up-tick. We run two MW tests to compare the average proportion of non-fulfillment for each case—announcements and promises—at group level over the whole 10 periods in the two treatments. The tests both report a *p*-value of less than 1%. The incidence of false announcements and broken promises is thus statistically significantly larger in **NCTwP** than in **NCTwP&R**. In fact, while on the whole 53.4% of promises were not fulfilled in **NCTwP**, the corresponding figure for the **NCTwP&R** treatment is only 17.5%. The results of the MW tests are shown in Table 12.

[Table 12 about here]

[Figure 5b about here]

Finally, Figure 5b graphs the average cost of punishment per period due both to cost to punishers and money lost by punishees in the **NCTwR** and the **NCTwP&R** treatments. The figure shows total cost of punishment to be higher in all but one period in the **NCTwR** treatment with the gap tending to widen after period 3. A MW test finds the average cost significantly higher in **NCTwR**, with a *p*-value of 9%. The result is shown in table 13a below.[35]

[Table 13a, 13b about here]

The fact that the roughly 80% honest announcement and promise rates in the treatment (Figure 4a) were achieved at modest social cost of around *E*$1 per subject per period, plus the effectiveness of punishment threats in boosting the contributions of less cooperative subjects and that of high contribution announcements in boosting the contributions of conditionally cooperative subjects, explains why earnings were considerably higher in the **NCTwP&R** treatment.

4. Discussion and Conclusions

Because no cooperative equilibrium is possible in a finitely repeated public goods game with rational payoff maximizing agents having common knowledge of their types, standard economic theory implies that the addition of opportunities to announce possible contributions in a non-binding fashion before costly play will have no effect. Being devoid of potential efficacy, any numerical messages sent could just as well be meaningless babble.

Our earlier experiment with non-binding numerical communication appeared to confirm the expectation that such communication has no effect on play, insofar as average binding behaviors followed approximately the same patterns with as without announcement stages. However, by disaggregating the results to the level of individual groups and to within-group interactions, we've discovered that non-binding announcements helped some groups to cooperate, while leading to a more complete break-down of cooperation in others. The increase in dispersion of group outcomes is evidently explained by inter-group differences in the extent to which subjects misled other group members with false announcements. Subjects seemed to take their messages seriously, as evidenced by the fact that mutual adjustments of announced choices display the same qualitative patterns as does

[35] Interestingly, the average expenditure on punishment for the three groups that spend the least on it is *E*$3.83 per period in the **NCTwR** treatment versus 0 in the NCTwP&R treatment

real play, and by the fact that real contributions and, in treatments with punishment opportunities, costly punishments are influenced by message content. It makes sense for opportunists to try to "set up" others so as to free ride on their contributions, but only if opportunists believe that their signals may be taken seriously. To such opportunists, "talk is cheap" in the lay person's common sense of that phrase, but not in the more demanding sense of the kind of economic theory that assumes common knowledge and payoff maximization. That theory would have talk be uniformly ignored, and individuals, knowing that their talk would be ignored, would waste no effort on issuing meaningful signals.

In BPP (2006), we had speculated that one reason why numerical cheap talk was less effective overall than was verbal communication is that NCT prevented subjects from framing their announcements in the moral language of explicit promises. As a partial test of that conjecture, we conducted new experiments in which, in addition to typing "possible" decisions into the message space used in the "numerical cheap talk" treatments, subjects could select, or not, a statement promising to contribute a specific amount to the public good. This test was imperfect, because we explicitly told subjects that promises were not binding (rather than risk the possibility that promises would be fulfilled due to a misunderstanding of the experiment's rules), stirring up cynicism of a kind less likely to arise when subjects make promises in a more spontaneous fashion. Nevertheless, the outcome supported the conjecture in one treatment, in which subjects could impose costly punishments. Many subjects punished "lying" on promise statements, and accordingly promises became more truthful and more credible, permitting many groups to achieve high levels of cooperation partly through cost-free threats and using less costly punishment.

The goal of our research has been to shed light on *why* communication aids cooperation, despite the predictions of standard economic theory. Our experiments add weight to the evidence suggesting that (a) many decision-makers behave as if they were maximizing something other than their monetary payoff alone, and that (b) most decision-makers act as if they assume this to be the case. At least three "extended" or "non-standard" preferences may underlie the results of our own and similar experiments. The efficacy of written promises to contribute even in treatments without punishment suggests that many subjects get disutility from breaking their word and/or believe this to be true of others, in which case the exchanging of promises alters expectations about one another's behaviors. Many may also get higher subjective payoffs from cooperating provided that others cooperate, so that what are prisoners' dilemma payoffs in pecuniary terms are assurance game payoffs in the space of utilities (Guttman, 2003; Page, Putterman and Unel, 2005). Finally, many subjects display a willingness to incur monetary costs in order to penalize free riders and those who deliberately mislead in their announcements and promises.

In the real world, people frequently do cooperate in matters of common interest. A cynical view is that when businessmen, partners in political coalitions, and others get together to find common ground, they simply bargain over the terms of agreements and the penalties and other mechanisms they will put in place to make those agreements self-enforcing for rational, self-interested agents. A more natural interpretation, however, is that such communication also allows parties to assess one another's trustworthiness, or in the language of economic theory, the content of their utility functions. Giving one's word alters subsequent play in part because some individuals can be counted on to penalize themselves, psychically, should they break such a bond, and because the promiser, knowing human nature, knows that retaliation for betrayal may go beyond what is in the pecuniary interest of the punisher.

References

Bicchieri, C., 2005. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press, New York.

Bochet, O., Page, T., Putterman, L., 2005. Communication and punishment in voluntary contribution experiments. Working Paper 2005-09, Department of Economics, Brown University. www.brown.edu/Departments/Economics/Papers/2005/2005-09_paper.pdf

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

Brosig, J., Ockenfels, A., Weimann, J., 2003. The effect of communication media on cooperation. German Economic Review 4, 217-42.

Carpenter, J., Matthews, P., 2002. Social reciprocity. Middlebury College Department of Economics Working Paper #29.

Charness, G., Dufwenberg, M., 2006. Promises and partnerships. Econometrica 74, 1579-1601.

Cinyabuguma, M., Page, T., Putterman, L., 2004. On perverse and second-order punishment in public goods experiments with decentralized sanctioning. Brown University Department of Economics Working Paper 2004-12.

Cinyabuguma, M., Page, T., Putterman, L., 2006. Can second-order punishment deter perverse punishment? Experimental Economics 9, 265-79.

Davis, D., Holt, C., 1993. Experimental Economics. Princeton University Press, Princeton.

Arhan, E., Page, T., Putterman, L., 2005. Can endogenous institutions mitigate the free rider problem and reduce perverse punishment? Brown University Dept. of Economics Working Paper 2005-13.

Falkinger, J., Fehr, E., Gächter, S., Winter-Ebmer, R., 2000. A simple mechanism for the efficient provision of public goods: experimental evidence. American Economic Review 90, 247-264.

Farrel, J., Rabin, M., 1996. Cheap talk. Journal of Economic Perspectives 10, 103-118.

Fehr, E., Gächter, S., 2000a, Cooperation and punishment. American Economic Review 90, 980-94.

Fehr, E., Gächter, S., 2000b. Fairness and retaliation: the economics of reciprocity. Journal of Economic Perspectives 14, 159-81.

Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? evidence from a public goods experiment. Economics Letters 71, 397-404.

Frohlich, N., Oppenheimer, J., 1998. Some consequences of e-mail vs. face-to-face communication in experiment. Journal of Economic Behavior and Organization 35, 389-403.

Gächter, S., Thöni, C., 2005. Social learning and voluntary cooperation among like-minded people. Journal of the European Economic Association 3, 303-314.

Gunnthorsdottir, A., Houser, D., McCabe, K., 2007. Disposition, history and contributions in public goods experiments. Journal of Economic Behavior and Organization 62, 304-15.

Guttman, J., 2003. Repeated interaction and the evolution of preferences for reciprocity. Economic Journal 113, no. 489: 631-656.

Hoffman, E., McCabe, K., Smith, V., 1998. Behavioral foundations of reciprocity: experimental economics and evolutionary psychology. Economic Inquiry 36, 335-52.

Isaac, R., Walker, J., 1988. Communication and free-riding behavior: the voluntary contributions mechanism. Economic Inquiry 26, 585-608.

Kurzban, R., Houser, D., 2001. Individual differences in cooperation in a circular public goods game. European Journal of Personality 15 (S1): S37-S52.

Kreps, D., Milgrom, P., Roberts, J., Wilson, R., 1982. Rational cooperation in finitely repeated prisoners' dilemma. Journal of Economic Theory 27, 245-52.

Ledyard, J., 1995. Public goods: a survey of experimental research. in Kagel, J., Roth, A. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 111-94.

Masclet, D., Noussair, C., Tucker, S., Villeval, M.-C., 2003. Monetary and non-monetary punishment in the VCM. American Economic Review 93, 366-80.

Nikiforakis, N., Normann, H.-T., 2005. A comparative statics analysis of punishment in public goods experiments. Royal Holloway University of London Discussion Papers in Economics, June.

Önes, U., Putterman, L., 2007. The ecology of collective action: a public goods and sanctions experiment with controlled group formation. Journal of Economic Behavior and Organization 62, 495-521.

Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without a sword: self governance is possible. American Political Science Review 86, 404-416.

Page, T., Putterman, L., Unel, B., 2005. Voluntary association in public goods experiments: reciprocity, mimicry, and efficiency. Economic Journal 115, 1032-53.

Palfrey, T., Prisbrey, J., 1997. Anomalous behavior in public goods games: how much and why? American Economic Review 5, 829-846.

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

Sánchez-Pagés, S., Vorsatz, M., forthcoming. Enjoy the silence: an experiment on truth-telling. Games and Economic Behavior (in press).

Sefton, M., Shupp, R., Walker, J., 2002. The effect of rewards and sanctions in provision of public goods. Working Paper, University of Nottingham and Indiana University.

Wilson, R., Sell, J., 1997. "Liar, liar...": cheap talk and reputation in repeated public goods settings. Journal of Conflict Resolution 41, 695-717.

Figure 1: Average contribution per period, by treatment

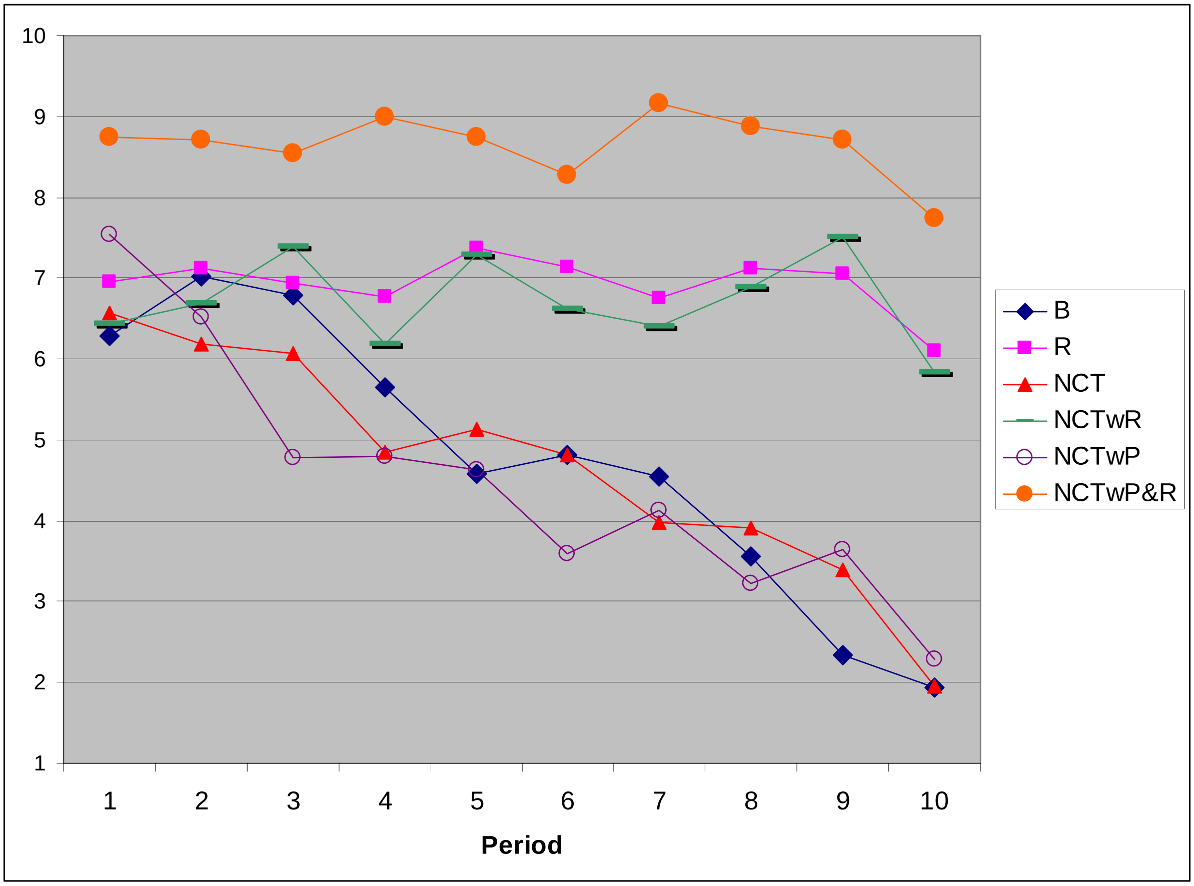


Figure 2: Average earnings per period, by treatment

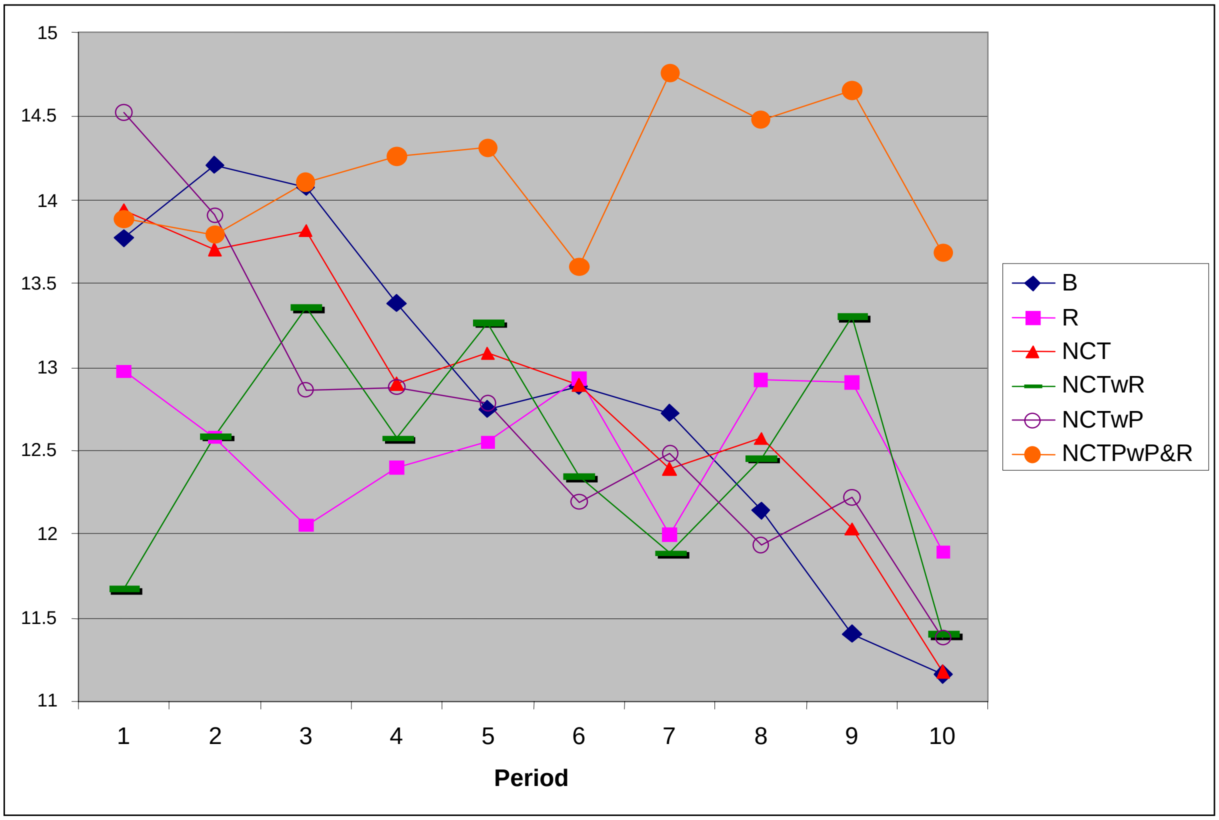


Table 1. Summary of Treatments.

Communication \ Reduction Option	No	Yes
None	Baseline (**B**) – 12 groups (48 subjects)	Reduction (**R**) – 12 groups (48 subjects)
Numerical announcements	Numerical Cheap Talk (**NCT**) – 11 groups (44 subjects)	Numerical Cheap Talk with Reduction Option (**NCTwR**) 11 groups (44 subjects)
Numerical announcements and promise statements	Numerical Cheap Talk with Promise Option (**NCTwP**) 11 groups (44 subjects)	Numerical Cheap Talk with Promise and Reduction Options (**NCTwP&R**) 11 groups (44 subjects)

<u>Table 2a. *p*-values of two-tailed Mann-Whitney tests of differences in group average contributions and group average earnings</u>

	B	R	NCT	NCTwR	NCTwP	NCTwP&R
B	.	**0.002** **C(R)>C(B)**	0.740	**0.096** **C(NCTwR)>C(B)**	0.235	**0.001** **C(NCTwP&R)>C(B)**
R	0.977	.	**0.016** **C(R)>C(NCT)**	1	**0.016** **C(R)>C(NCTwP)**	**0.023** **C(NCTwP&R)>C(R)**
NCT	0.740	0.786	.	**0.065** **C(NCTwR)>C(NCT)**	0.650	**0.001** **C(NCTwP&R)>C(NCT**
NCTwR	0.695 (0.451)	0.786 (0.462)	0.847 (0.485)	.	**0.087** **C(NCTwR)>C(NCTwP)**	0.087 **C(NCTwP&R)>C(NCTwR**
NCTwP	0.235 (0.348)	0.832 (0.501)	0.699 (0.446)	0.898 (0.521)	.	0.001 (0.087)
NCTwP &R	**0.037** **C(NCTwP&R)> C(B)**	**0.051** **E(NCTwP&R)> E(R)**	**0.023** **E(NCTwP &R)>E(NCT)**	**0.040** **E(NCTwP&R)> E(NCTwR)**	**0.013** **E(NCTwP&R)> E(NCTwP)**	.

Note: Numbers to the right and above the diagonal are for tests of differences in contributions. Numbers to the left and below the diagonal are for tests of differences in earnings. The number of observations per treatment equals the number of groups (see Table 1).

<u>Table 2b. Average contributions and average earnings across groups ~~and periods~~</u>

Treatment	Average Contribution	Average Earnings
B	4.752	12.851
R	6.935	12.520
NCT	4.684	12.850
NCTwR	6.722	12.485
NCTwP	4.515	12.717
NCTwP&R	8.641	14.153

<u>Table 3a. *p*-values of two-tailed Mann-Whitney and Wilcoxon tests of differences in per period group variances</u>

	B vs NCT	B vs. NCTwP	R vs NCTwR	R vs NCTwP&R
Mann-Whitney test	**0.001** **var(C(NCT))> var(C(B))**	**0.001** **var(C(NCTwP))> var(C(B))**	**0.001** **var(C(NCTwR))> var(C(R))**	0.218
Wilcoxon test	**0.005** **var(C(NCT))> var(C(B))**	**0.005** **var(C(NCTwP))> var(C(B))**	**0.007** **var(C(NCTwR))> var(C(R))**	0.139

	NCT vs NCTwP	NCTwR vs. NCTwP&R
Mann-Whitney test	**0.049** **var(C(NCTwP))> var(C(NCT))**	**0.001** **var(C(NCTwR))> var(C(NCTwP&R))**
Wilcoxon test	**0.093** **var(C(NCTwP))> var(C(NCT))**	**0.007** **var(C(NCTwR))> var(C(NCTwP&R))**

Note: The number in each cell is the *p*-value of the respective test. One observation per group.

Table 3b. Variance of average contribution, by treatment.

Treatment	Average variance of average contribution across groups
B	3.245
NCT	7.251
NCTwP	9.541
R	5.128
NCTwR	9.833
NCTwP&R	4.213

Figure 3
Variance of average contribution across groups by period and treatment

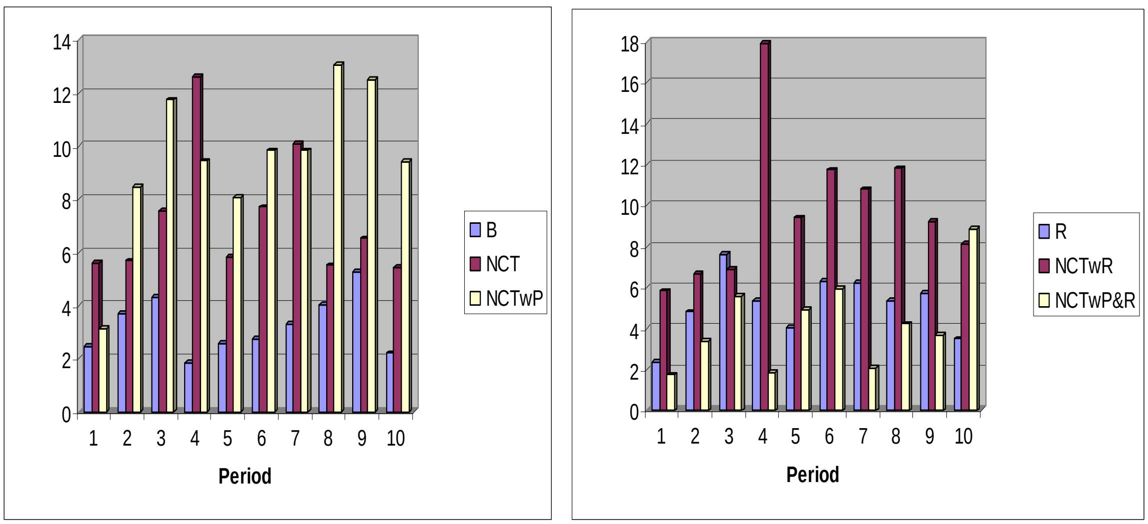


Table 4a. Adjustment of contributions in response to differences from means, B and R treatments

Dependent variable: (contribution by subject *i* in period *t*)-(contribution by subject *i* in period *t*-1).

	B			R		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
(Contribution by *i* in period *t*-1) – (Average contribution by others in *i*'s group in period *t*-1)	-0.371 (0.067) [0.001]	-0.445 (0.077) [0.001]	-0.237 (0.072) [0.001]	-0.331 (0.059) [0.001]	-0.227 (0.287) [0.010]	-0.329 (0.058) [0.001]
Contribution by *i* in period *t*-1	-0.233 (0.066) [0.001]	-0.294 (0.075) [0.001]	-0.424 (0.080) [0.001]	-0.179 (0.054) [0.001]	-0.668 (0.107) [0.001]	-0.163 (0.052) [0.002]
Constant	0.703 (0.365) [0.054]	0.217 (0.805) [0.787]	3.396 (0.668) [0.001]	1.165 (0.401) [0.004]	1.216 (0.862) [0.158]	1.303 (0.494) [0.008]
Individual Fixed effects	No	Yes [0.004]	No	No	Yes [0.001]	No
Period fixed effects	No	No	Yes [0.001]	No	No	Yes [0.100]
R^2	0.303	0.370	0.359	0.246	0.433	0.261

Note: Panel corrected standard errors regression (using the xtpcse command in Stata). N = 432, all regressions. In all the regressions, numbers in parentheses are standard-errors and numbers in square brackets are *p*-values. In the individual and period fixed effects rows, number in square brackets are *p*-values of tests of the joint significance of individual fixed effects, and period fixed effects, respectively.

Table 4b. Correlation between (change in *i*'s contribution between period *t* and *t*-1) and (difference between contribution of *i* in *t*-1 and contribution of others in *i*'s group in *t*-1)

Period	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
B	$\rho = -0.562$ [0.001]	$\rho = -0.295$ [0.041]	$\rho = -0.601$ [0.001]	$\rho = -0.459$ [0.001]	$\rho = -0.526$ [0.001]	$\rho = -0.505$ [0.001]	$\rho = -0.715$ [0.001]	$\rho = -0.714$ [0.001]
R	$\rho = -0.492$ [0.001]	$\rho = -0.416$ [0.003]	$\rho = -0.68$ [0.001]	$\rho = -0.321$ [0.025]	$\rho = -0.546$ [0.001]	$\rho = -0.123$ [0.425]	$\rho = -0.548$ [0.001]	$\rho = -0.510$ [0.001]

Note: Non-parametric Spearman correlation test. N = 48, all tests. Number **ρ** is Spearman's correlation coefficient. Numbers in square brackets are *p*-values and numbers N refer to the number of paired observations. All the subsequent tables following regressions show Spearman correlation tests.

Table 4c. Adjustment of announced contributions in response to differences from means, NCT and NCTwR treatments

Dependent variable: first change of announced contribution by subject *i*.

	NCT			NCTwR		
	(1)	(2)	(3)	(4)	(5)	(6)
(1st announced contribution by *i*) – (Average 1st announced contribution by others in *i*'s group)	-0.366 (0.073) [0.001]	-0.311 (0.089) [0.001]	-0.375 (0.073) [0.001]	-0.232 (0.075) [0.002]	-0.079 (0.098) [0.418]	-0.243 (0.076) [0.001]
1st announced contribution by *I*	-0.343 (0.078) [0.001]	-0.498 (0.102) [0.001]	-0.332 (0.080) [0.001]	-0.376 (0.084) [0.001]	-0.692 (0.115) [0.001]	-0.361 (0.085) [0.001]
Constant	2.348 (0.635) [0.001]	1.306 (1.311) [0.319]	1.817 (0.774) [0.019]	2.812 (0.727) [0.001]	4.994 (1.274) [0.001]	1.911 (0.819) [0.020]
Individual fixed effects	No	Yes [0.001]	No	No	Yes [0.001]	No
Period fixed effects	No	No	Yes [0.691]	No	No	Yes [0.270]
R^2	0.339	0.434	0.348	0.242	0.374	0.260

Note: Panel corrected standard errors regression. N = 440, all regressions.

Table 4d: Correlation between (first change of announced contribution by subject *i*) and (first announced contribution by *i* – average first announced contribution by others in *i*'s group)

Period	2	3	4	5	6	7	8	9
NCT	$\rho = -0.601$ [0.001]	$\rho = -0.370$ [0.013]	$\rho = -0.608$ [0.001]	$\rho = -0.663$ [0.001]	$\rho = -0.551$ [0.001]	$\rho = -0.511$ [0.001]	$\rho = -0.461$ [0.002]	$\rho = -0.479$ [0.001]
NCTwR	$\rho = -0.440$ [0.003]	$\rho = -0.397$ [0.008]	$\rho = -0.465$ [0.001]	$\rho = -0.572$ [0.001]	$\rho = -0.534$ [0.001]	$\rho = -0.424$ [0.004]	$\rho = -0.432$ [0.003]	$\rho = -0.595$ [0.001]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

<u>Table 5a: Announced and actual reductions as a function of announced and actual contribution deviations and the deviation of actual from announced contribution, NCTwR treatment.</u>

Dependent variable:

	announced pun. "received" by *j*		actual punishment received by *j*		actual punishment received by *j*	
	(1)	(2)	(3)	(4)	(5)	(6)
Abs. neg. dev.	1.741 (0.161) [0.001]	1.761 (0.150) [0.001]	1.097 (0.104) [0.001]	1.113 (0.105) [0.001]	1.096 (0.104) [0.001]	1.106 (0.105) [0.001]
Abs. Pos. dev.	-0.679 (0.362) [0.061]	-0.868 (0.332) [0.009]	-0.496 (0.175) [0.005]	-0.558 (0.180) [0.002]	-0.482 (0.175) [0.006]	-0.543 (0.180) [0.003]
Avg. contrib. (*j* excluded)	-1.152 (0.318) [0.001]	-1.393 (0.336) [0.001]	-0.088 (0.086) [0.306]	-0.159 (0.113) [0.160]	-0.077 (0.086) [0.373]	-0.153 (0.112) [0.174]
Difference btw. announced and actual contrib.					0.122 (0.064) [0.058]	0.169 (0.083) [0.043]
Constant	4.619 (3.056) [0.131]	2.535 (3.447) [0.462]	-2.293 (1.116) [0.040]	-2.421 (1.364) [0.077]	-2.418 (1.113) [0.030]	-2.948 (1.390) [0.035]
Random Effects Tobit	Yes		Yes		Yes	
Tobit with group fixed effects		Yes		Yes		Yes
LR test p-value	0.001		1.000		1.000	
Log Likelihood	-770.33	-705.33	-569.36	-770.33	-567.57	-560.23

Note: Tobit regressions and random effects tobit regressions. N = 440, all regressions.

<u>Table 5b.</u> Correlation between the announced (respectively binding) punishment received by *i* and the absolute negative deviation by *i* from the average announced (respectively binding) contribution by others in *i*'s group, NCTwR treatment.

Period	1	2	3	4	5	6	7	8	9
Announced	$\rho = 0.483$ [0.001]	$\rho = 0.592$ [0.001]	$\rho = 0.490$ [0.001]	$\rho = 0.362$ [0.016]	$\rho = 0.549$ [0.001]	$\rho = 0.430$ [0.004]	$\rho = 0.738$ [0.001]	$\rho = 0.407$ [0.006]	$\rho = 0.537$ [0.001]
Binding	$\rho = 0.648$ [0.001]	$\rho = 0.592$ [0.001]	$\rho = 0.667$ [0.001]	$\rho = 0.237$ [0.122]	$\rho = 0.090$ [0.560]	$\rho = 0.757$ [0.001]	$\rho = 0.217$ [0.158]	$\rho = 0.578$ [0.001]	$\rho = 0.328$ [0.030]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

<u>Table 5c.</u> Correlation between the binding punishment received by *i* and the difference between announced and actual contribution made by *i*, NCTwR treatment.

Period	1	2	3	4	5	6	7	8	9
Difference between announced and actual contribution	$\rho = 0.280$ [0.066]	$\rho = 0.094$ [0.542]	$\rho = 0.490$ [0.001]	$\rho = 0.034$ [0.812]	$\rho = 0.337$ [0.025]	$\rho = 0.295$ [0.052]	$\rho = 0.292$ [0.054]	$\rho = 0.232$ [0.129]	$\rho = 0.029$ [0.851]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

<u>Table 6a. Responses to announced reduction threats</u>

Dependent variable: first change in *j*'s announced possible contribution

	NCTwR			NCRwP&R		
	(1)	(2)	(3)	(4)	(5)	(6)
Initial announced reductions, if *j* is not the maximum announced contributor	0.244 (0.042) [0.001]	0.283 (0.048) [0.001]	0.247 (0.041) [0.001]	0.168 (0.043) [0.001]	0.140 (0.045) [0.002]	0.169 (0.045) [0.001]
Initial announced reductions, if *j* is the maximum announced contributor	-0.163 (0.071) [0.023]	-0.155 (0.077) [0.045]	-0.135 (0.069) [0.052]	-0.190 (0.154) [0.215]	-0.212 (0.158) [0.180]	-0.196 (0.157) [0.213]
Constant	-0.604 (0.178) [0.001]	-0.468 (0.478) [0.327]	-1.000 (0.513) [0.052]	-0.610 (0.116) [0.001]	0.085 (0.088) [0.336]	-0.659 (0.135) [0.001]
Individual fixed effects	No	Yes [0.004]	No	No	Yes [0.075]	No
Period fixed effects	No	No	Yes [0.322]	No	No	Yes [0.010]
R^2	0.132	0.268	0.152	0.040	0.220	0.054

Note: Panel corrected standard errors regression. N = 440, all regressions.

<u>Table 6b.</u> Correlation between first change of announced contribution by subject *i* and announced reduction threat (*i* not highest contributor)

Period	1	2	3	4	5	6	7	8	9
NCTwR	$\rho = 0.270$ [0.076]	$\rho = 0.459$ [0.002]	$\rho = 0.107$ [0.491]	$\rho = 0.454$ [0.002]	$\rho = 0.448$ [0.002]	$\rho = 0.458$ [0.002]	$\rho = 0.696$ [0.001]	$\rho = 0.414$ [0.005]	$\rho = 0.458$ [0.002]
NCTwP&R	$\rho = 0.493$ [0.001]	$\rho = 0.332$ [0.028]	$\rho = 0.486$ [0.001]	$\rho = 0.139$ [0.369]	$\rho = 0.223$ [0.145]	$\rho = -0.084$ [0.589]	$\rho = 0.020$ [0.900]	$\rho = 0.031$ [0.840]	$\rho = 0.381$ [0.011]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

Table 7a. Actual contribution as a function of others' past and announced contributions, and own announced contribution, NCT and NCTwR treatments.

Dependent variable: period *t* contribution by subject *i*

	NCT			NCTwR		
	(1)	(2)	(3)	(4)	(5)	(6)
Average *binding* contribution of group members other than *i* in period *t*-1	0.434 (0.072) [0.001]	0.298 (0.087) [0.001]	0.366 (0.755) [0.001]	0.431 (0.063) [0.001]	0.140 (0.079) [0.078]	0.448 (0.065) [0.001]
Average last *announced* contribution of group members other than *i* in period *t*	0.142 (0.095) [0.135]	0.124 (0.113) [0.273]	0.168 (0.094) [0.074]	0.267 (0.077) [0.001]	0.146 (0.078) [0.063]	0.257 (0.077) [0.001]
i's last announced contribution in period *t*	0.265 (0.076) [0.001]	0.298 (0.087) [0.001]	0273 (0.073) [0.001]	0.397 (0.052) [0.001]	0.255 (0.064) [0.001]	0.397 (0.051) [0.001]
Constant	-0.684 (0.426) [0.109]	-2.588 (1.262) [0.040]	0.255 (0.658) [0.698]	-1.378 (0.496) [0.005]	4.440 (1.446) [0.001]	-1.035 (0.617) [0.093]
Individual fixed effects	No	Yes [0.001]	No	No	Yes [0.001]	No
Period fixed effects	No	No	Yes [0.135]	No	No	Yes [0.359]
R^2	0.245	0.589	0.268	0.500	0.684	0.510

Note: Panel corrected standard errors regression. N = 352, all regressions.

Table 7b. Correlation between binding contribution by subject *i* in period *t* and average last announced contribution of others in *i*'s group in period *t*

Period	2	3	4	5	6	7	8	9
NCT N=44	$\rho = 0.276$ [0.070]	$\rho = 0.466$ [0.001]	$\rho = 0.475$ [0.001]	$\rho = 0.382$ [0.011]	$\rho = 0.292$ [0.054]	$\rho = 0.508$ [0.001]	$\rho = 0.225$ [0.142]	$\rho = -0.157$ [0.308]
NCTwR N=44	$\rho = 0.483$ [0.001]	$\rho = 0.529$ [0.001]	$\rho = 0.741$ [0.001]	$\rho = 0.650$ [0.001]	$\rho = 0.649$ [0.001]	$\rho = 0.525$ [0.001]	$\rho = 0.626$ [0.001]	$\rho = 0.213$ [0.165]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

Table 7c. Correlation between binding contribution by subject *i* in period *t* and *i*'s last announced contribution in period *t*

Period	2	3	4	5	6	7	8	9
NCT N=44	$\rho = 0.343$ [0.023]	$\rho = 0.448$ [0.002]	$\rho = 0.407$ [0.006]	$\rho = 0.230$ [0.134	$\rho = 0.468$ [0.001]	$\rho = 0.488$ [0.001]	$\rho = 0.277$ [0.069]	$\rho = 0.247$ [0.106]
NCTwR N=44	$\rho = 0.520$ [0.001]	$\rho = 0.491$ [0.001]	$\rho = 0.624$ [0.001]	$\rho = 0.704$ [0.001]	$\rho = 0.526$ [0.001]	$\rho = 0.557$ [0.001]	$\rho = 0.624$ [0.001]	$\rho = 0.485$ [0.001]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

Table 8a. Punishment received as a function of broken promises and other variables, NCTwP&R treatment

	Dependent variable:							
	actual pun. received by *j*		actual pun. received by *j*		actual pun. Received by *j*		actual pun. received by *j*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Abs. Neg. dev.	1.355 (0.114) [0.001]	1.530 (0.135) [0.001]	0.951 (0.124) [0.001]	1.147 (0.143) [0.001]	1.052 (0.123) [0.001]	1.311 (0.145) [0.001]	0.882 (0.126) [0.001]	1.111 (0.145) [0.001]
Abs. Pos. dev.	0.118 (0.297) [0.691]	-0.154 (0.343) [0.654]	0.640 (0.312) [0.040]	0.388 (0.351) [0.270]	0.477 (0.307) [0.120]	0.156 (0.349) [0.654]	0.706 (0.311) [0.024]	0.440 (0.353) [0.214]
Avg. contribution (j excluded)	0.157 (0.200) [0.433]	-0.299 (0.264) [0.257]	0.665 (0.221) [0.003]	0.138 (0.274) [0.614]	0.542 (0.218) [0.013]	-0.015 (0.276) [0.954]	0.751 (0.224) [0.001]	0.193 (0.279) [0.489]
Dummy = 1 if promise was made, 0 otherwise	-0.161 (0.763) [0.833]	0.583 (0.871) [0.504]	-2.378 (0.825) [0.004]	-1.908 (0.952) [0.046]	-0.351 (0.740) [0.635]	0.326 (0.838) [0.697]	-2.013 (0.837) [0.016]	-1.728 (0.963) [0.074]
Dummy = 1 if promise was broken, 0 otherwise			4.993 (0.903) [0.001]	4.692 (1.005) [0.001]			3.913 (1.021) [0.001]	4.165 (1.121) [0.001]
Dummy =1 if contributed less than last announced, 0 otherwise					3.475 (0.752) [0.001]	2.437 (0.820) [0.003]	1.777 (0.844) [0.035]	0.893 (0.891) [0.316]
Constant	-3.293 (2.006) [0.101]	-2.467 (2.885) [0.393]	-6.275 (2.100) [0.003]	-4.228 (2.869) [0.141]	-7.215 (2.216) [0.001]	-5.233 (3.023) [0.084]	-7.578 (2.198) [0.001]	-5.062 (2.991) [0.091]
Random effects Tobit	Yes		Yes		Yes		Yes	
Tobit with group fixed effects		Yes		Yes		Yes		Yes
LR p-value	0.001		0.001		0.001		0.001	
LogLikelihood	-409.5	-346.8	-394.2	-336.1	-399.8	-342.5	-392.1	-335.5

Note: Tobit regressions and random effects tobit regressions. N = 440, all regressions.

<u>Table 8b</u>: NCTwP&R. Correlation between binding punishment received by *i* and,

Period	1	2	3	4	5	6	7	8	9
Negative deviation	$\rho = 0.529$ [0.001]	$\rho = 0.627$ [0.001]	$\rho = 0.579$ [0.001]	$\rho = 0.617$ [0.001]	$\rho = 0.484$ [0.001]	$\rho = 0.574$ [0.001]	$\rho = 0.683$ [0.001]	$\rho = 0.643$ [0.001]	$\rho = 0.434$ [0.003]
Broken promise	$\rho = 0.119$ [0.438]	$\rho = 0.486$ [0.001]	$\rho = 0.588$ [0.001]	$\rho = 0.645$ [0.001]	$\rho = 0.707$ [0.001]	$\rho = 0.546$ [0.001]	$\rho = 0.677$ [0.001]	$\rho = 0.637$ [0.001]	$\rho = 0.221$ [0.149]
Unfulfilled announcement	$\rho = 0.432$ [0.003]	$\rho = 0.467$ [0.001]	$\rho = 0.493$ [0.001]	$\rho = 0.469$ [0.001]	$\rho = 0.485$ [0.001]	$\rho = 0.422$ [0.004]	$\rho = 0.612$ [0.001]	$\rho = 0.437$ [0.003]	$\rho = 0.237$ [0.120]

Note: Non-parametric Spearman correlation test. N = 44, all tests.

<u>Table 9a. The negative impact of "lying" on group performance in NCT and NCTwR</u>

Dependent variable: Average contribution in group, period *t*

	NCT			NCTwR		
	(1)	(2)	(3)	(4)	(5)	(6)
Average "lie" in *t*-1	-0.617 (0.078) [0.001]	-0.318 (0.088) [0.001]	-0.580 (0.084) [0.001]	-0.587 (0.114) [0.001]	-0.033 (0.096) [0.727]	-0.599 (0.116) [0.001]
Last average announcement in *t*	0.726 (0.060) [0.001]	0.381 (0.118) [0.001]	0.725 (0.056) [0.001]	0.984 (0.070) [0.001]	0.366 (0.103) [0.001]	0.984 (0.070) [0.001]
Constant	0.604 (0.434) [0.164]	3.286 (1.084) [0.002]	0.620 (1.612) [0.312]	-0.470 (0.618) [0.447]	2.142 (0.987) [0.030]	-0.754 (0.764) [0.324]
Group fixed effects	No	Yes [0.001]	No	No	Yes [0.001]	No
Period fixed effects	No	No	Yes [0.291]	No	No	Yes [0.971]
R^2	0.631	0.796	0.664	0.736	0.852	0.743

Note: Panel corrected standard errors regression. N = 88, all regressions.

<u>Table 9b.</u> Correlation between the average contribution of group *i* in period *t* and the average lie of group *i* in period *t*-1

Period	2	3	4	5	6	7	8	9
NCT	$\rho = -0.473$ [0.141]	$\rho = -0.447$ [0.168]	$\rho = -0.698$ [0.016]	$\rho = -0.391$ [0.233]	$\rho = -0.490$ [0.125]	$\rho = -0.146$ [0.668]	$\rho = -0.433$ [0.182]	$\rho = -0.255$ [0.449]
NCTwR	$\rho = -0.137$ [0.685]	$\rho = -0.346$ [0.296]	$\rho = -0.182$ [0.590]	$\rho = -0.278$ [0.407]	$\rho = -0.746$ [0.008]	$\rho = -0.127$ [0.708]	$\rho = -0.178$ [0.598]	$\rho = -0.664$ [0.025]

Note: Non-parametric Spearman correlation test. N = 11, all tests.

<u>Table 10a. The negative impact of "lying" on group performance, NCTwP and NCTwP&R treatments</u>

Dependent variable: Average contribution in group, period *t*

	NCTwP			NCTwP&R		
	(1)	(2)	(3)	(4)	(5)	(6)
Avg. "lie" on A, *t*-1	-0.295 (0.122) [0.016]	-0.259 (0.113) [0.023]	-0.265 (0.125) [0.035]	0.292 (0.210) [0.165]	0.390 (0.190) [0.040]	0.415 (0.230) [0.070]
Avg. last announced contribution	0.682 (0.126) [0.001]	0.439 (0.136) [0.001]	0.696 (0.122) [0.001]	0.651 (0.133) [0.001]	0.469 (0.122) [0.001]	0.644 (0.133) [0.001]
Avg. "lie" on P, *t*-1	-0.255 (0.112) [0.023]	-0.024 (0.112) [0.826]	-0.259 (0.110) [0.018]	-0.679 (0.265) [0.011]	-0.298 (0.267) [0.263]	-0.535 (0.292) [0.067]
Avg. last promised contribution	0.151 (0.145) [0.299]	0.165 (0.145) [0.253]	0.195 (0.139) [0.163]	0.284 (0.096) [0.003]	0.252 (0.075) [0.001]	0.267 (0.102) [0.009]
Constant	-0.360 (1.066) [0.735]	3.951 (1.510) [0.009]	-0.201 (1.161) [0.863]	0.236 (1.252) [0.850]	2.081 (1.427) [0.863]	0.445 (1.335) [0.739]
Group fixed effects	No	Yes [0.001]	No	No	Yes [0.001]	No
Period fixed effects	No	No	Yes [0.270]]	No	No	Yes [0.691]
R^2	0.583	0.758	0.623	0.569	0.757	0.576

Note: Panel corrected standard errors regression. N = 88, all regressions.

<u>Table 10b.</u> NCTwP. Correlation between the average contribution of group *i* in period *t* and the average lie (announcements and promises) of group *i* in period *t*-1

Period	2	3	4	5	6	7	8	9
Average lie on announcement, period t-1	$\rho = -0.305$ [0.361]	$\rho = -0.321$ [0.335]	$\rho = -0.150$ [0.658]	$\rho = -0.200$ [0.553]	$\rho = -0.220$ [0.514]	$\rho = -0.461$ [0.153]	$\rho = -0.247$ [0.462]	$\rho = -0.375$ [0.254]
Average lie on promises, period t-1	$\rho = -0.418$ [0.199]	$\rho = -0.746$ [0.008]	$\rho = -0.478$ [0.136]	$\rho = -0.482$ [0.132]	$\rho = -0.336$ [0.311]	$\rho = -0.393$ [0.231]	$\rho = -0.694$ [0.017]	$\rho = -0.479$ [0.135]

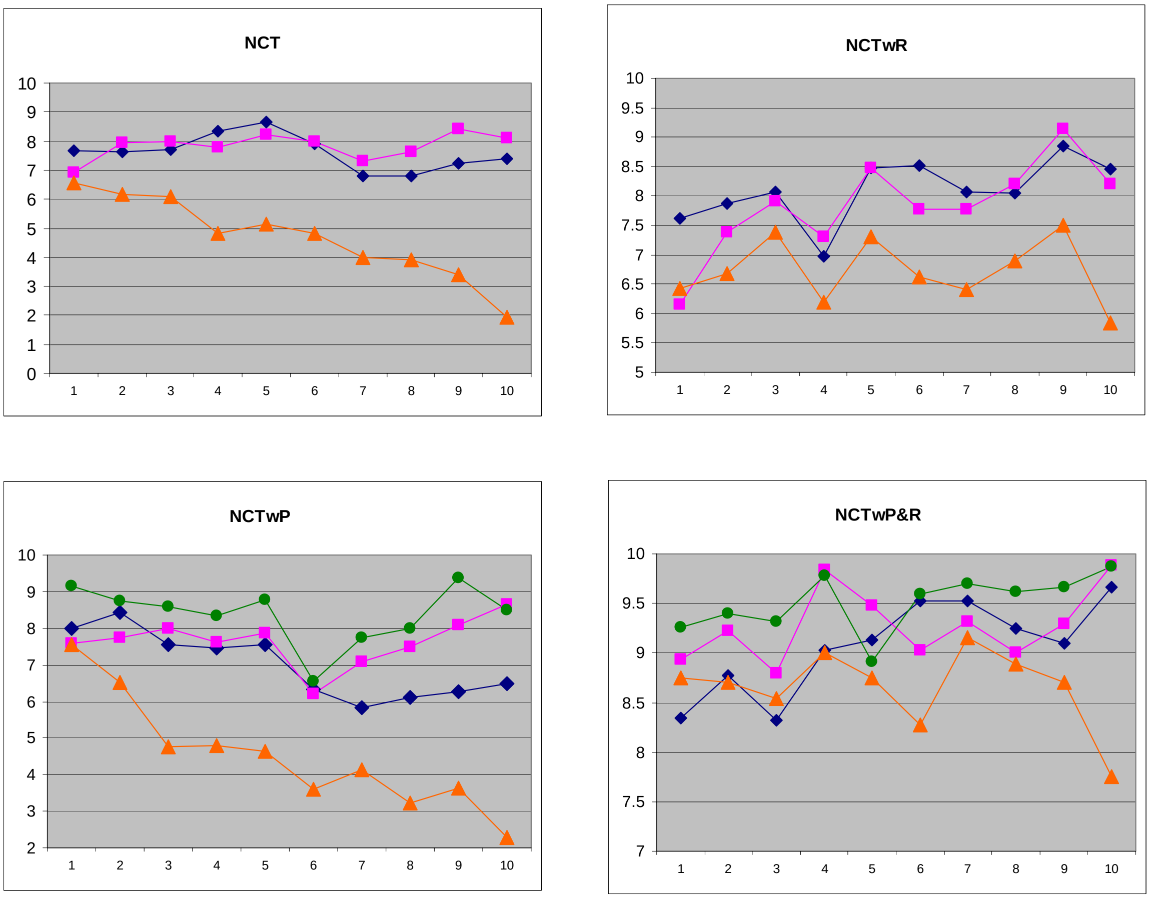
Note: Non-parametric Spearman correlation test. N = 11, all tests.

<u>Table 10c.</u> NCTwP&R. Correlation between the average contribution of group *i* in period *t* and the average lie (announcements and promises) of group *i* in period *t*-1

Period	2	3	4	5	6	7	8	9
Average lie on announcement, period t-1	$\rho = -0.220$ [0.514]	$\rho = -0.341$ [0.304]	$\rho = -0.674$ [0.022]	$\rho = -0.552$ [0.078]	$\rho = -0.538$ [0.087]	$\rho = -0.649$ [0.030]	$\rho = -0.207$ [0.540]	$\rho = -0.091$ [0.790]
Average lie on promises, period t-1	$\rho = -0.100$ [0.769]	$\rho = -0.583$ [0.059]	$\rho = -0.674$ [0.022]	$\rho = -0.743$ [0.008]	$\rho = -0.283$ [0.398]	$\rho = -0.828$ [0.001]	$\rho = -0.426$ [0.190]	$\rho = -0.686$ [0.019]

Note: Non-parametric Spearman correlation test. N = 11, all tests.

Figure 4: Average announcements, promises and binding contributions



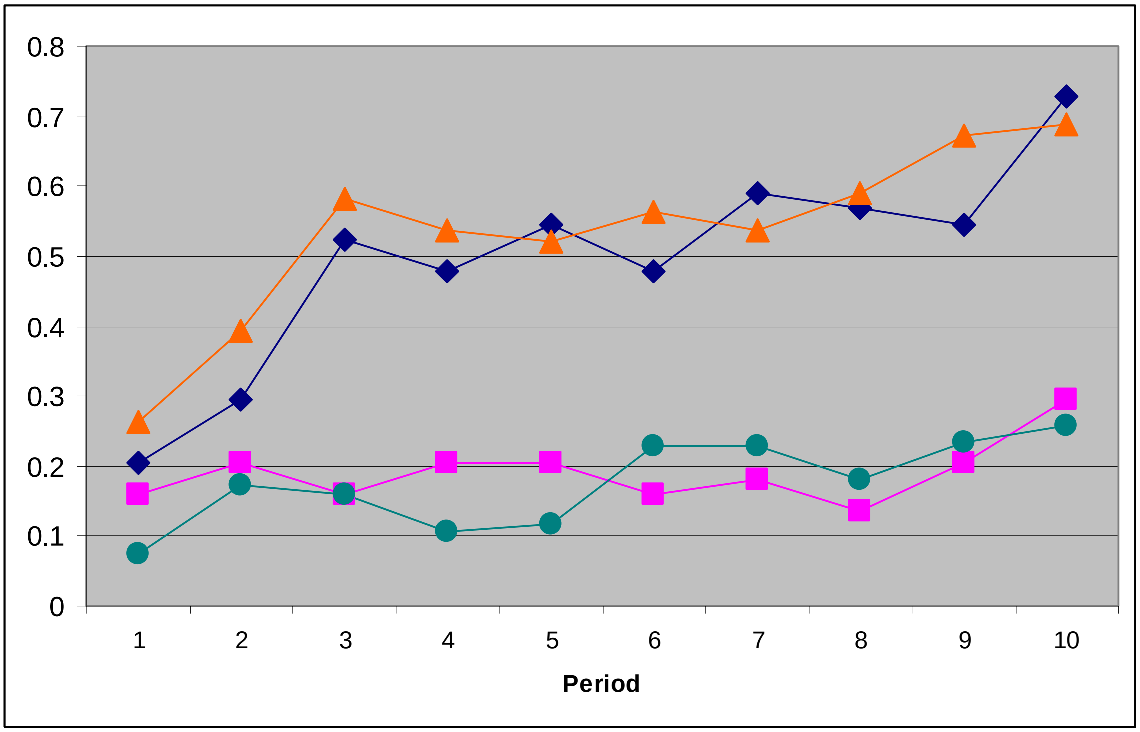
Note: Triangles = Average contribution, Diamonds = Average first announcement, Squares = Average last announcement, Circles = Average promise. Notice that vertical scales differ as appropriate to each treatment's observations.

Table 11. *p*-values of two-tailed Mann-Whitney tests of differences in group average gaps between average last announced contribution and average binding contribution; and differences in group average gaps between average promised contributions and average binding contributions

	NCTwR	NCTwP	NCTwP&R
NCT	**0.007** **NCT>NCTwR**	1	**0.001** **NCT>NCTwP&R**
NCTwR	.	**0.013** **NCTwP>NCTwR**	0.300
NCTwP	.	.	**0.001 0.001** **NCTwP>NCTwP&R**

Note: The number in each cell displays the *p*-value of the test. The first number in the **NCTwP-NCTwP&R** cell refers to the test with average announced contribution, while the second number refers to the test with average promised contribution. The number of observations for each treatment equals the number of groups (see Table 1).

Figure 5a: Proportion of lies on promises and announcement, NCTwP and NCTwP&R



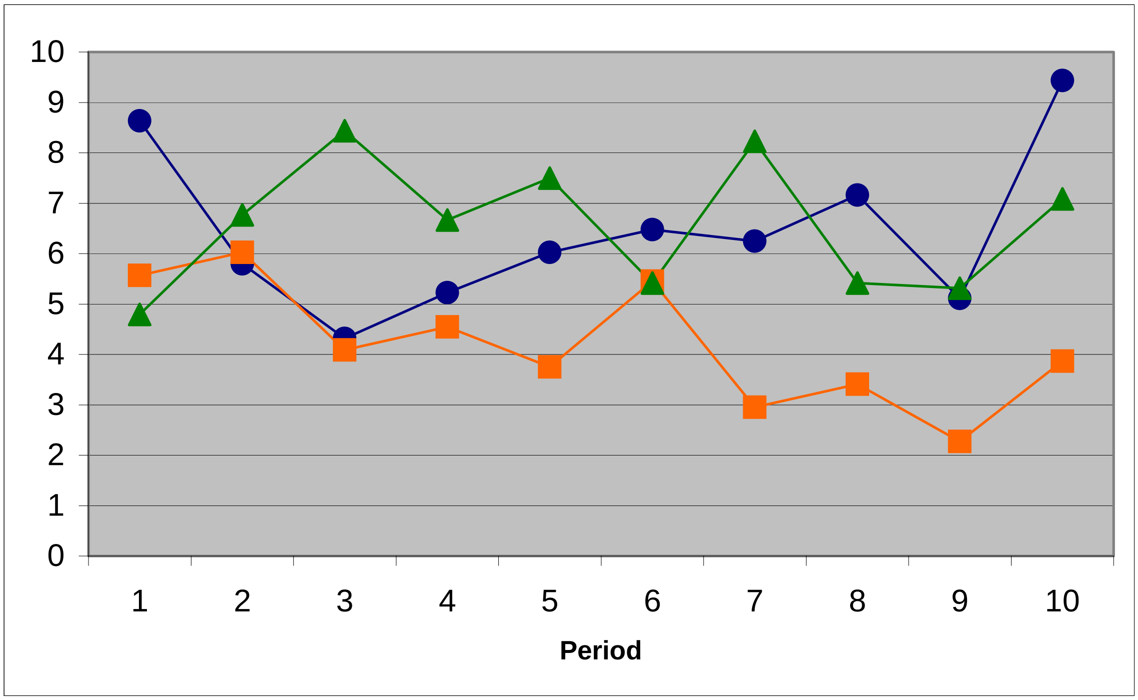
Note: Diamonds = Proportion of lies on announcements in **NCTwP**, Squares = Proportion of lies on announcements in **NCTwP&R**, Triangles = Proportion of lies on promises in **NCTwP**, Circles = Proportion of lies on promises in **NCTwP&R**.

Table 12. *p*-values of two-tailed Mann-Whitney tests of differences in group average proportion of lies on announcements of contribution and group average proportion of lies on promises.

	NCTwR	NCTwP	NCTwP&R
NCT	**0.010** NCT>NCTwR	0.898	**0.001** NCT>NCTwP&R
NCTwR	.	**0.013** NCTwP>NCTwR	0.332
NCTwP	.	.	**0.001 0.001** NCTwP>NCTwP&R

Note: The number in each cell is the *p*-value of the test. The first number in the **NCTwP-NCTwP&R** cell refers to the test with average proportion of lies on announced contributions, while the second number refers to the test with average proportion of lies on promised contribution. The number of observations for each treatment equals the number of groups (see Table 1).

Figure 5b. Average earnings loss from punishment, NCTwR and NCTwP&R



Note: Triangles = Average loss from punishment, R treatment, Circles = Average loss from punishment, NCTwR treatment, Squares = Average loss from punishment, NCTwP&R treatment. Average loss is at group level; thus, loss per subject is one quarter of the amount shown.

Table 13. *p*-values of two-tailed Mann-Whitney tests of differences in group average earnings loss due to punishment

	NCTwR	**NCTwP&R**
R	0.608	0.288
NCTwR	.	**0.088** **NCTwR>NCTwP&R**

Note: In each row, the first number displays the *p*-value of the test. Number of observations per treatment is same as number of groups (see Table 1).