Not Just Babble: A Voluntary Contribution Experiment with Iterative Numerical Messages

by Olivier Bochet and Louis Putterman*

Abstract

When subjects can make non-binding announcements of possible contributions to a public good numerically, there is no effect on average level of contributions in a public goods experiment relative to play without announcements. But a detailed analysis of this experiment shows that pre-play announcements increased the variance of achieved cooperation among groups, leading both to more highly cooperative groups and to more thoroughly uncooperative groups than in a treatment without announcements. We also add a treatment in which subjects can select a statement of (non-binding) "promise" to contribute a certain amount and we find that even though subjects were instructed that promise statements were not binding, the ability to issue them significantly increased both contributions and earnings in a treatment that includes costly punishment opportunities, although not in a treatment without punishment.

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

---

* University of Namur and Brown University. Send correspondence to olivier.bochet@fundp.ac.be 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 thank again the many students who assisted in carrying out the experiments reported in Bochet, Page and Putterman (forthcoming), as well as Dennis Zachary Schubert and Dmitri Lemmerman, who programmed the new treatments.

Not Just Babble: A Voluntary Contribution Experiment with Iterative Numerical Messages

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 to provision of effort in partnerships and work 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 which 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 sanction one another (Fehr and Gächter, 2000a), and excluding free riders from playing with more cooperative subjects (Gunnthorsdottir, Houser, McCabe and Ameden, 2002; Ones and Putterman, 2004). 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, 2004) found not only that pre-play 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.

In BPP, we also reported VCM experiments with two other kinds of communication. First, we performed 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 they might contribute to the group account. The chat room treatments yielded levels of cooperation only slightly lower and statistically indistinguishable from the treatments with face-to-face communication, thus casting doubt on suggestions that the efficacy of such communication is attributable to information conveyed by facial expression, vocal intonation, and body language. Unlike the treatments that allowed the exchange of verbal 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 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 signalling 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 signalling, achieved even less cooperation than their least successful no communication counterparts. Both the coordination successes and the effects of false signalling indicate that subjects took numerical announcements as something other than cheap talk in the sense 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 even though the efficacy of the promise option was compromised by our explicit instruction that it was nonbinding, the ability to make promises enhanced cooperation in our public goods treatment without sanction opportunities. The option of sending promise messages still more significantly improved outcomes in those groups whose members were also given the opportunity to sanction one another with costly punishment.

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 and with and without sanction options. Section 3 discusses numerical signalling theoretically and differing assumptions about preferences and beliefs. Section 4 analyzes and compares the behaviors with and without cheap talk. Section 5 presents the promise option experiments and analyzes the results. Section 6 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, say 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

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

where C ($0 \leq C \leq E$) is an individual's contribution to the group account and the summation is taken over all group members, i included. (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 returns 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 $C_i = 0$, $\forall\ i$. While an outcome having $C_i = E$ for all players dominates it, there is no credible way, within the payoffs of the game, to punish deviations from an agreement to contribute E, so communication cannot in principle alter the outcome.[1] Communication can change the outcome only if (a) payoffs can be altered in a manner external to the game, for example if a contract can be supported by the threat of penalties imposed by a third party, for example the state, or if (b) we drop the common knowledge assumption, allowing that some players' objective functions don't perfectly coincide with their material payoffs and/or that some players entertain the belief that players with such preferences might be present.[2]

In the first systematic investigation of the matter by economists, 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 that report 130 different experimental treatments whose results Sally (1995) entered in multi-variate regressions to study which treatment variables best account for differing levels of cooperation and free riding. Sally concluded that face-to-face communication was the single strongest of the treatment variables studied, 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.

---

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

[2] A variant of (a) might be to have players themselves punish defectors after the finitely repeated game ends, but a threat to do so can be credible, assuming self-interested players with common knowledge, only if the finitely repeated game is part of an infinitely repeated game that they are playing (e.g.) outside of the laboratory. Alternatively, players could effectively threaten post-play penalties *if* they have or are believed to have other preferences—for example, if it is believed that some may be sufficiently *angry* to penalize others in post-play interaction—which amounts to a variant of (b). Our experiments attempt to rule out post-experiment penalties by explicitly ruling out threats related to events after a session, by recruiting subjects from a student body numbering around six thousand in a random fashion so that participants are unlikely to know one another, and by preventing subjects from knowing who is in their group and which individual is responsible for a given decision. Social punishments after the experiment would in any case fall more into the category (b), involving preferences beyond payoffs, and thus support our main explanation for observed cooperation.

To understand better what lies behind the effects of face-to-face communication, Brosig, Ockenfels and Weimann (hereafter BOW, 2003) and BPP, who strongly confirmed the result noted by Sally, 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-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 also conducted two kinds of public goods game—a standard VCM experiment, and one like Fehr and Gächter's having a stage following each contribution round in which group members could engage in costly reductions of one anothers' earnings after learning of their contributions. BPP's chat room communication treatment without punishment resembled that of Frohlich and Oppenheimer (1998), except that those authors used e-mail messages, which do not provide the same 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.

BOW and BPP each found one treatment that achieved almost equally large efficiency gains as did their face-to-face communication treatments. For BOW, this was the audio-video conference treatment; for BPP, the chat room treatment. BPP's numerical cheap talk treatment performed no differently on average than did our no communication baseline. The tentative conclusion offered by BOW was that electronic communication can fully replace face-to-face interaction so long as group members can see one another's body language and facial expression and hear one another's words and intonation. However, even written communication was as effective as face-to-face communication in BPP's experiment. Because the numerical cheap talk treatment in BPP performed no differently on average than did our no communication baseline, we conjectured there that the difference was due to the ability in face-to-face, audio-visual, and chat room treatments, but the inability in the numerical announcement treatment, to frame the problem with language.[3]

There are many reasons why the use of language (in face-to-face and audio-visual conferences and in a chat room) might help individuals to achieve cooperation more effectively than do announcements of numbers. Oral and written communication allows group members who more quickly grasp the nature of the dilemma problem to educate others and move toward a common framing of the task that faces them. Such communication also makes it possible for joint strategies to be proposed, for individuals to verbally commit themselves to an agreement, and for subjects to size up one another's trustworthiness. They can frame the problem they are facing to one another in moral language and they can engage in efforts to build team spirit. All of this is impossible when subjects can transmit only the number of dollars that they are considering contributing to the public good. But it also should make no difference if subjects' preferences are strictly limited to maximizing their own payoffs, without social or psychological components.

In this paper, we analyze BPP's numerical communication treatments, previously described in the aggregate only, at the level of individual subject behaviors. We demonstrate that numerical announcements did affect binding play, and that their content was anything but white noise. Even though groups failed, on average, to achieve greater cooperation when allowed to signal possible choices numerically, our comparison of numerical communication treatments with the counterpart treatments without

[3] This leaves the further question of why BOW's audio-only treatment was relatively unsuccessful. In BPP, we conjectured that this had to do with the complete physical isolation of the subjects from one another by BOW but not BPP.

communication shows that some groups were more (less) successful at cooperating than were even relatively cooperative (uncooperative) counterparts in the baseline and punishment-only treatments. Furthermore, an analysis of numerical announcements shows them to be something quite other than random noise or babble. In actuality, the pattern of numerical announcements closely resembled patterns in binding play. For example, in the treatment with numerical communication and punishment stages, subjects announced larger punishments of those who announced smaller contributions, and those at whom such announced punishments were targeted responded by raising their announced contributions. Actual contributions are significantly correlated with both own and others' announced contributions, and contributing less than the amount announced tended to be punished. A careful study of announcements thus provides evidence of the fact that subjects did *not* understand their messages to be "cheap talk" in the full sense required by a theory of rational, self-interested agents with common knowledge of one anothers' type.

Our paper also reports on a set of new experiments designed to test a hypothesis offered by BPP. 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 verbal component, we conjectured that the ability to make 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 conducted sessions identical to those of BPP's "numerical cheap talk" treatments with and without punishment option, 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 these promises, which were also not binding other than by force of conscience, led to high contributions in the treatment without sanctioning opportunities, and to both higher contributions and higher earnings in the treatment with such opportunities.

2. Experimental Design

We discuss the decisions of eighteen experimental sessions, in each of which sixteen (in four of the sessions, twelve) undergraduate subjects made a series of contribution decisions in randomly assigned and anonymous groups of four which stayed together for a total of ten periods. 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 uniformly free rode.[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, of which the first four are discussed in this section.[5] 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 $\gamma =$ \$0.25 to the punisher per $\delta =$ \$1 of earnings loss to the person punished. Individuals were informed of the reductions they themselves received only, without knowing which and what combination of others were responsible. Subjects in all four treatments were identified to 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

[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.

[5] These four are among the eight treatments discussed also by BPP.

amended announcements. During these periods in the **NCT** treatment, subjects simply entered a possible contribution amount in a screen identical to the binding decision interface 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). 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). The full instructions given to the subjects, including practice problems, are provided in the working paper version of BPP.

The designs of the modified versions of the **NCT** and **NCTwR** treatments in which subjects could also choose statements of promise following each numerical communication stage, denoted **NCTwP** and **NCTwP&R**, respectively, are discussed in Section 5. For now, it is helpful to note that they contained precisely the same elements as the **NCT** and **NCTwR** treatments, with only the addition of the promise stage to each period. Table 1 summarizes the six treatments.

Table 1 about here

Because each subject participated in a session of one treatment only, the analysis of treatment effects follows a between-subject design, which assumes that the subjects in each treatment are essentially the same, a reasonable assumption given their numbers and the common population from which they were drawn.[6] Much of the analysis is conducted at the level of the group of four, since group members had no knowledge of what was occurring in any of the other groups in their session, and group behaviors are accordingly statistically independent, whereas the behaviors of individuals within a given

[6] None of the 204 subjects had previously participated in an economics experiment.

group could begin to affect one another once the first set of decisions or announcements had been revealed.

3. Standard and Bayesian Predictions

As mentioned above, 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 argued 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 accordingly theory still predicts to 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 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. Unlike the common sense notion of cheap talk, which admits of the possibility of unscrupulous individuals using the communication opportunity to take advantage of less sophisticated or more trusting types[7], there is no reason to type one number rather than another if fellow players are known to be rational, to care only about their own payoffs, and to know that you are also of that type. If such subjects are required by the experiment to enter some number into a field and if no number is any more difficult to enter than another, then the predicted stream of messages will be a smear of random numbers, a pure numerical babble or gibberish.[8]

---

[7] Crawford (2002), in a context different than ours, shows that rational players exploit boundedly rational agents by misrepresenting their intentions: payoff-maximizing players set-up fellow group members who have non-standard type or who are simply more credulous as to the possible existence of agents with nonstandard preferences.

[8] Our game is a problem of coordination under conflict. With the common knowledge assumption that players are standard, the unique subgame perfect equilibrium outcome is Pareto dominated by the outcome in which agents fully contribute to the group account. As shown in Farrel-Rabin (1996), however, conflict of interest among agents means that

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 possibility. 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 two possible "nonstandard" preferences: reciprocity, and truth-telling. A reciprocating agent is one who prefers to cooperate if he believes that others are cooperating and who is willing to incur a cost to punish someone who exploits him by free riding while he contributes.[9] 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. 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).

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*., 2002). 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 may begin to reduce their contributions to the group account, leading to the gradual downward slide that is usually

---

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.

[9] 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. Ones and Putterman (2004) consider that the relative degrees of positive and of negative reciprocity may differ from one reciprocator to another.

seen in finitely repeated VCM experiments. Let the reciprocators punish the free riders while maintaining their own high contributions, however, as in our **R** treatment, and contributions may stabilize or rise rather than falling, as is found by Fehr and Gächter (2000a), Masclet, Noussair, Tucker and Villeval (2003), BPP, and Sefton, Shupp and Walker (2002).

Consider now what communication might add to the Bayesian story. With face-to-face communication, it is conceivable that members of a group can go so far as to signal their types by way of gesture, intonation, etc.[10] A group of reciprocators might thus assure one another of their types and initiate a run of self-sustaining high contributions. A group of truth-tellers might likewise pledge themselves to contributing their endowments, and, if confident of one another's commitments, proceed to fulfill their promises. In our **NCT** and **NCTwR** treatments, however, subjects were unable to communicate feelings or intentions, but could simply type into the relevant fields numbers described by the experimenter as possible contributions and possible reductions. While much less favorable to cooperation than verbal communication, this opportunity to send numerical signals wouldn't necessarily be viewed as useless by subjects who believe reciprocators or truth-tellers are common.[11] Unlike agents in a world of common

---

[10] Ordinarily, cheap talk is viewed as a device that has no impact on the payoff-structure of the underlying game. However, when there is uncertainty about types and thus about subjective payoffs (which can differ from material ones), the discussion that takes place can have a direct impact on the structure of payoffs since people respond to each other (a fact underlined by Farrel-Rabin (1996)). Even pay-off maximizing agents may be convinced that it is in their interest to behavior cooperatively until last period.

[11] As Farrel and Rabin (1996) underlined, in our game cooperation cannot be achieved if there are only standard agents in the population. They do not emphasize what happens if there are non-standard types present in some groups. In the numerical messages treatment, even if agents share a common language (a necessary condition for cheap talk to convey information), numbers may not be informative enough to convey information on the type of an agents, nor to build a sense of trust. The fact that some groups did poorly while others performed well can be related to the presence of opportunists in a group. In the language of Crawford-Sobel (1982), (notice that they do not consider the case of non-standard preferences), if their preferences are not "too far apart" (in the present case, if the proportion of reciprocators or truth-telling agents is high enough), then agents can achieve cooperation.

knowledge of payoff maximizing type, such subjects might attempt to signal intentions and to read one another's messages as possible signals of intention.[12]

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 typing 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 indeed follow. If the game includes punishment opportunities, the reciprocators might signal intentions to punish low contributions, and some might follow through with actual punishment not only when others free ride *per se*, but also when they see evidence of attempts to mislead by announcing high contributions but contributing little. Opportunistic subjects whose only goal is to maximize their payoffs might also signal and act cooperatively in some periods, but later they might attempt to exploit the signalling opportunity and the anticipated credulity of fellow players, "setting up" team members by suggesting 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 with respect to our main focus, communication. The standard model implies that "numerical cheap talk" will amount to a meaningless 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. Further analysis is required in

[12] The fact that some groups did poorly may also be interpreted (for instance, like in Crawford (2002), even though the context is different) on the ground that opportunistic players set-up fellow group members by pretending they are going to contribute. Real contributors got quickly discouraged by misrepresentations of intentions.

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 matter.

4. Analysis of the **NCT** *and no communication treatments*

Result 1 restates the main findings regarding numerical cheap talk in BPP, where the two treatments are analyzed in terms of averages only.

*Result 1: On average, contribution and earning trends did not significantly differ in a treatment with numerical communication (***NCT***, or* **NCTwR***) from those in the counterpart treatment without such communication (***B***, or* **R***). Contributions show declining trends with repetition in the* **B** *and* **NCT** *treatments, and no such trend in the* **R** *and* **NCTwR** *treatments. Average contributions are thus higher in the* **R** *and* **NCTwR** *treatments. Due to the costliness and in some cases misdirection of punishment, average earnings do not statistically differ among the four treatments.*[13]

Figures 1 and 2 show the average number of dollars contributed to the public good, and the average earnings after deduction of punishment costs, in the B, R, NCT, and NCTwR treatments, by period. 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, although with evidence of attempts to boost contributions in periods 2 and 6.[14] In the R treatment, as in the similar treatment in Fehr and Gacther (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

[13] Cinyabuguma, Page and Putterman (in process) show that about 20% of earnings reductions were aimed at high rather than low contributors, a phenomenon they label "perverse punishment." Ertan, Page and Putterman (2004) 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).

[14] Past results are surveyed in Davis and Holt (1993) and in Ledyard (1995). Not too much should be made of the timing of these increases, since the result averages patterns in twelve separate groups.

of many subjects to impose costly punishment on low contributors.[15] 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.

Figure 1 about here

Figure 2 about here

The trends of *average* contributions in the **NCT** and **NCTwR** treatments resemble closely those of their no communication counterparts, the **B** and **R** treatments. Mann-Whitney tests confirm that average contributions over the ten periods as a whole do not differ significantly as between the **NCT** and **B** treatments, or as between the **NCTwR** and **R** treatments.[16] From averaged behavior, therefore, it appears that giving subjects the opportunity to announce possible decisions before each set of binding decisions made no difference to outcomes.

Similarity also exists between the **B** and **NCT** treatments, and between the **R** and **NCTwR** treatments, with respect to the patterns of average earnings. The similarities are closer for the **B** and **NCT** treatments, where patterns of contributions directly map into patterns of earnings and the patterns of declining average contributions is mirrored in patterns of declining earnings. The **R** and **NCTwR** treatments contain extra degrees of freedom because reductions (punishments) and contributions may be related in a multiplicity of different ways. For both treatments, nonetheless, earnings (which are net of reduction costs) lie below those in the treatments without punishment during the early

---

[15] Cinyabuguma, Page and Putterman (in process) show that the amount of punishment received was significantly increasing in the negative deviation of a subject's contribution from the average of fellow group members, in the BPP **R** treatment and in Fehr and Gacther's punishment condition. They also show that low contributors responded to punishment by raising their contributions, in both experiments. 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).

[16] These tests are discussed further in Section 3 when we turn to comparisons of our new treatments with the option to select a promise statement. A summary of the tests' *p*-values appears in Table 7.

periods and, because earnings in the no punishment treatments fall while those in the treatments with punishment show no consistent pattern, the punishment treatment earnings lie above those of the no punishment treatments in the later periods. Mann-Whitney tests find no significant difference of overall average earnings either between the **B** and **NCT** treatments or between the **R** and **NCTwR** treatments, except that earnings in **B** differ from (exceed) those in **R** significant at the 10% level in a two-tailed test. Apart from this, the tests find no difference in earnings among any of the four treatments.

As remarked earlier, the similarity of averaged outcomes across numerical communication treatments and their no-communication counterparts does not prove one way or the other that numerical communications were babble. The next three results will show that numerical messages were not random noise or babble: the messages were taken quite seriously by most subjects and had real effects on their binding, payoff determining decisions. The first step is to show that subjects adjusted their announced plans to one anothers' announced plans in much the same way that subjects' binding decisions react to one anothers' binding decisions in treatments *without* communication.

Result 2. During communication periods, subjects in the **NCT** *treatment adjusted their announced contributions in the direction of the average announced contributions of other group members.*

Table 2 reports results of OLS regressions in which the dependent variables are the initial change in a subject's announced ("possible") contribution during a communication period, and the independent variables are the difference between that subject's initial announcement and the average initial announcement of the other members of his/her group.[17] The middle column shows the results for **NCT** treatment data, the right column those for **NCTwR** treatment data. In both cases, there is a highly significant negative coefficient on the difference in announced contribution, which means that the further below (above) the other's average was *i*'s original announced

[17] All OLS regressions are reported with robust (Huber-White) standard errors calculated using the robust command in Stata.

contribution, the more did i increase (decrease) the announced amount in his or her first adjustment.[18]

Table 2 about here

Next, we verify that the influence of announcements is not limited only to other announcements (which are costless), but extends as well to costly decisions.

Result 3. In the **NCT** and **NCTwR** *treatments, actual contributions in a period were positively related to the average of others' announced contributions as well as to one's own announced contribution.*

Table 3 shows 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.[19] The regressions include individual fixed effect terms, not shown.

Table 3 about here

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 as a potentially profitable investment in reputation).

In both treatments, actual last period contributions by others significantly and positively affect own contribution. Others' most recent announcement are also positively correlated

[18] The regressions include all cases in which a subject changed his or her announced contribution during a communication period. With 48 subjects and 10 periods, sample size is potentially 480 for each regression, but the smaller actual sample sizes result from the fact that announcement changes were actually made in only about 20% of periods—in part because many groups settled into repeating patterns after the early periods of play.

[19] 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.

with own actual contribution, although the effect is significant at the 10% level only in the **NCTwR** treatment.[20]

Next, we look at the effects of announced punishments, beginning with their effects on other announcements.

Result 4. "Possible" behaviors announced in communication stages of the **NCTwR** *treatment display the same pattern of interaction as do actual behaviors in the* **R** *treatment, in that subjects targeted their announced reductions primarily at announced low contributors, and the latter reacted by raising their announced contributions.*

The first column of Table 4 reports a 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 from the average of other group members, and that average itself, are independent variables.[21] Apart from the fact that the variables are announced rather than actual contributions and reductions, the specification is identical to that used by Fehr and Gächter to demonstrate that punishment was mainly aimed at low contributors in their experiment, a specification replicated for the **R** treatment in BPP (2003).[22] 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 second column shows a parallel regression, also using data from the **NCTwR** treatment, but this time data on actual, as opposed to

[20] Note that the combination of truthful announcements and persistence of binding contributions causes last actual and current announced contributions to be correlated, so that the estimates may be biased towards zero.

[21] A tobit 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.

[22] 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.

announced, contributions and reductions. The similarity of the coefficients on the absolute deviation terms shows 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 prediction of the theory based on common knowledge of uniformly payoff maximizing preferences—namely, that numbers communicated would amount to random noise.

Table 4 about here

The second part of Result 4, subjects' adjustments of their announced contributions in response to one anothers' announced reductions, is demonstrated by the OLS regression shown in Table 5. The dependent variable is each subject'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 the subject's earnings, interacted with 0,1 dummy variables the first of which takes the value 1 only if the would-be target of punishment announced a contribution less than the maximum one announced in her group, the second of which takes the value 1 only if the would-be target communicated the highest announcement in the group. The first coefficient indicates that those who announced a less-than-maximum contribution increased their announcement by an average of 36 cents per dollar of announced punishment "received," a reaction qualitatively identical and quantitatively similar to that found by BPP for actual contributions following actual punishment. The second suggests that a targeted high contributor might slightly reduce her announcement, but this coefficient is not statistically significant.[23]

Table 5 about here

Result 5. A significant number of subjects in the **NCTwR** *treatment treated the announcements of others as implicit obligations in the sense that those who contributed less than they announced tended to receive actual costly punishment over and above what*

[23] Cinyabuguma, Page and Putterman (in process) find that when a group's highest contributor receives a unit of costly punishment, she tends to reduce her contribution in the next period (whereas a low contributor tends to increase his contribution when he receives punishment).

is predicted by considering the level of their actual contribution alone, with the amount of punishment being significantly increasing in the difference between announced and actual contribution.

The third column of Table 4 reports a tobit regression with the same specification as column 2 except that the difference between subject *j*'s last announced contribution of the period and his/her actual contribution in the same period is added as an independent variable. The new variable has a positive coefficient which is significant at the 1% level. The estimate implies that for every one dollar of difference between announced and actual contribution, a subject received about 19 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 subjects themselves didn't treat one anothers' announcements as noise, but instead predicated costly decisions upon them.

Results 2 – 5 provide evidence 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 false signals to improve their individual returns from free riding. One way to test this conjecture is to see 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.

Result 6. The larger the average gap between announced and actual contributions, in early periods, the smaller were average contributions in a group in later periods.

Let "lying" denote a situation in which a subject's contribution to the group account is less than her last announced contribution during the communication stage of the same period, and let the extent of lying in a group be measured by the average difference between the last announcement and the actual contribution.[24] In Table 6, we present OLS regressions with one observation per group, including both **NCT** treatment and **NCTwR** treatment groups, in which the average extent of lying in periods 1 – 4 is an explanatory variable, and the average contribution to the group account in periods 5 – 9 is the dependent variable.[25] The result suggests that a one dollar increase in the average gap between announced and actual contributions in periods 1 – 4 reduced the average contribution in periods 5 through 9 by 95 experimental cents, with this effect being significant at the 1% level despite the small sample size.

Table 6 about here

The finding in Result 6 indicates that there was variation among groups in the extent to which numerical communication aided cooperation. Since *average* outcomes in each communication treatment were statistically indistinguishable from those in its no communication counterpart, the possibility arises that the outcomes of groups in the two kinds treatments might differ in their degrees of *dispersion*. The next two results show that this is indeed the case.

Result 7. Although **NCT** *behaviors were like* **B** *behaviors on average, there was more dispersion among groups in the* **NCT** *treatment, meaning its "successful" groups achieved higher contribution levels than the more successful groups in* **B***, and its "unsuccessful" groups achieved lower contribution levels than did the less successful groups in* **B**.

We demonstrate this result with two figures. First, Figure 3 plots average contribution by period for the three (of eleven) groups in the **NCT** treatment and the three (of twelve) groups in the **B** treatment that had the highest average contribution levels, and

---

[24] Although the extent of an individual's "lying" might be defined as being equal to zero whenever his actual contribution exceeded his last announced contribution in the communication round, we let lying (in the few cases of this type) take negative values.
[25] It's necessary to pool observations from the **NCT** and **NCTwR** treatments because with only one observation per group, the number of observations proves too small to yield significant results for either treatment's observations tested separately.

for the three **NCT** and three **B** groups with the lowest contribution levels.[26] For both highest and lowest groups, performance of **NCT** groups is more extreme than that of **B** treatment groups in all but one period (period 2, for the high groups; period 9, for the low groups).

Figure 3 about here

Second, a measure of the degree of dispersion among groups which takes into account all groups, not just the highest and lowest, is the coefficient of variation. We calculate the coefficient of variation of the average contribution among the NCT groups, and among the B groups, in each of the ten periods, and display the results in Figure 4. In eight of ten periods, the coefficient of variation is clearly larger for the groups in the NCT treatment; in only one period (period 9) is the coefficient clearly larger for B treatment groups.[27]

Figure 4 about here

Our final result of the section shows that a similar property of dispersion holds when comparing the no communication and numerical communication treatments that include reduction opportunities:

Result 8. **NCTwR** *treatment groups likewise had a greater dispersion of outcomes than did groups in the* **R** *treatment.*

Figures 5 and 6 parallel Figures 3 and 4, respectively, but for the **NCTwR** and **R** treatments. In Figure 5, the difference among highest and among lowest groups is not decisive for the highest three groups, but a pattern like that in Figure 3 does hold for the low-contributing groups, which almost always contributed less in the **NCTwR** than in the

[26] Each line follows the same three groups for the ten periods. That is, we identify the three groups which, on average, had the highest or lowest contributions over ten periods, rather than graphing values for which ever three groups were highest or lowest in each individual period.
[27] Formal tests of statistical significance can't be applied here because the coefficients of variation from different periods of a given treatment cannot be taken as statistically independent of one another.

R treatment. However, Figure 6 shows that when all groups are taken into account by calculating period by period coefficients of variation for each treatment, the greater dispersion of outcomes for the treatment with numerical communication is clear: the coefficient of variation of **NCTwR** treatment groups exceeds that of **R** groups in all but one period.[28]

Figure 5 about here

Figure 6 about here

5. **NCT** *with Promise*

a. Design

In Section 4, we've seen that while average outcomes in the treatments with numerical communication did not differ from their baseline and reduction counterparts, as reported in BPP, this was not because subjects disregarded the opportunity to announce possible choices or treated one anothers' announcements as mere noise. In fact, non-binding (announced) decisions displayed an iterated adjustment process closely resembling adjustments from period to period in the binding decision process of both the baseline and the reduction (punishment) treatments. These announcements also had real effects on binding decisions, and were treated as implicit promises by many subjects, who used costly reductions to punish deviations between announced and actual contributions to the group account. Some groups in the communication treatments succeeded in cooperating to a degree not matched in the counterpart treatments, while others failed more miserably than their worst no-communication counterparts, evidently because overt attempts to mislead produced an especially untrusting atmosphere.

In this section, we use a new treatment to explore the possibility that it was the inability of subjects to verbally *pledge* themselves to a strategy that accounts for the inferior average outcome in the **NCT** treatments as compared with the face-to-face,

[28] Lack of statistical independence across periods again makes formal significance tests impossible.

audio-video, and chat room treatments in BOW and BPP, where cooperation was close to 100%. Sally's (1995) meta-analysis found that experiments in which the experimenter explicitly suggested that subjects consider communicating promises achieved (statistically) significantly higher levels of cooperation. BPP cited that finding in support of their conjecture that the inability to bind themselves with promises is what accounts for the poorer average performance of **NCT** and **NCTwR** groups than of face-to-face and chat room communication subjects.

We designed a simple variation on the **NCT** and **NCTwR** treatments as a partial test of this conjecture. The new treatments are identical to the old ones, including a period of iteractive numerical communication of non-binding "possible" choices. However, at the end of each period's numerical iterative communication stage and before its binding contribution stage (a stage that follows immediately after iterative numerical communication in the **NCT** and **NCTwR** treatments), subjects were asked to choose between two statements. The first option read: "I promise to contribute __ to the group account this period." and it required the selection 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 __ to the group account." 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 call 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 experiments 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 doing so. To rule out the possibility that promises were effective because of such a misunderstanding, we included in the instructions about the binding decision stage 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 chuckles or raised eyebrows among the subjects at this point in every session. For this reason, we think that our decision to make clear to the subjects that promise statements were not automatically binding is likely to have reduced the prospects of raising cooperation levels—although not to have included such a statement might well have introduced an equally large bias in the opposite direction.

b. Results

Result 9. The inclusion of a promise option did not significantly alter the average performance of groups in the treatment without reductions. But in the treatment with reductions, both average contributions and average earnings were significantly higher when subjects could select promise statements.

Figures 1 and 2 include plots of average contributions and average earnings by period in the **NCTwP** and **NCTwP&R** treatments, respectively, along with the corresponding plots of the four treatments previously discussed. Table 7 summarizes the results of a set of Mann-Whitney tests comparing both contributions and earnings among all pairings of the six treatments by listing the *p*-values of the tests, where a low *p*-value indicates that we can reject the null hypothesis that the treatments in question differ in a random manner only.[29] In Figure 1, the graph of contributions in the **NCTwP** treatment follows roughly the same downward course as in the **B** and **NCT** treatments, and the lack of significant difference between **B** and **NCTwP** or between **NCT** and **NCTwP** is confirmed in Table 7. By contrast, Figures 1 and 2 show contributions and earnings to be distinctly higher than in all other treatments in the **NCTwP&R** treatment, and Table 7 confirms that both contributions and earnings were significantly different in this treatment in paired comparisons against each other treatment investigated. The

[29] In these non-parametric tests, we compare either average contributions or average earnings at the level of the groups of 4 subjects, averaged over the entire 10 periods of play. The number of observations for each treatment is therefore the same as the number of groups shown in Table 1.

opportunity to select a statement of promise led to higher contributions and higher earnings when subjects also had the opportunity to target costly punishment at one another.

Table 7 about here

We next provide some micro-analysis of the treatments with promise option to show that, as with the **NCT** treatment, announcements were not simply babble but rather had both internal rationality and effects on the costly actual decisions of the subjects.

Result 10: In the **NCTwP&R** *treatment as in the* **NCTwR** *treatment, subjects adjusted their announced contributions in response to announced punishments in the same manner as actual contributions react to actual punishment in conventional treatments.*

Table 8 reports the results of a regression that exactly parallels the one in Table 5, with the same qualitative result. If the group member announcing the highest of the indicated possible contributions is targeted for announced punishment, he or she tends to lower his/her announced contribution by *E*\$0.19 per dollar of announced punishment, whereas if any other group member is targeted for announced punishment, he or she tends to raise his or her announced contribution by an average of *E*\$0.17 per dollar of announced punishment.

Table 8 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** because other group members had the possibility of inflicting monetary damage on them were they to contribute less than the promised amount. That subjects who reneged on promises were often punished is confirmed in the next result.

Result 11. Contributing less than the amount specified in a promise statement drew actual costly punishment in even larger amount than does contributing less than indicated as a "possible" amount.

Table 9 reports a series of tobit regressions[30] resembling the last specification in Table 4, with a few changes made necessary by unusually high correlations among certain variables.[31] To reduce multicolinearity, we drop the average contribution variable and in place of the size of the deviations between actual and announced or promised contribution, we use categorical variables equalling 1 if the individual contributed less than announced or promised, and zero otherwise.

Table 9 about here

Column 1 contains the basic result paralleling Fehr and Gächter's: subjects received more punishment the further below other's average was their contribution. In column 2, we add the dummy variable for contributing less than promised, and find it to have a highly significant positive coefficient implying that a subject received an average of 4.12 experimental dollars of punishment if she broke a promise. In column 3, 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 a highly significant positive coefficient, implying that contributing less than announced elicited 3.18 experimental dollars of punishment. Finally, in column 4 both dummy variables are included, and both obtain significant positive coefficients. The larger absolute magnitude and greater statistical significance of the coefficient on the dummy for breaking a promise in the column 4 estimate, along with smaller magnitude of the coefficient on the promise than on the announcement breaking dummy when columns 2 and 3 are compared, leads to the conclusion, stated in Result 11, that breaking a promise leads to even more punishment than did contributing less than announced.[32]

---

[30] As with Table 4, 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.

[31] Specifically, the average contribution of others is highly correlated with the absolute positive deviation from others' average contribution (corr. = -.6238) and the absolute negative deviation from others' average contribution is highly correlated with both the deviation between last announced and actual contribution (corr. = .7644) and the deviation between promised and actual contribution (corr. = .7201).

[32] Of course, there is a fairly high correlation between the two dummy variables, because subjects often entered the same contribution number in their promise statement as they

As in the **NCT** and **NCTwR** treatments, we again suspect that a major factor explaining differences in outcomes among groups are differences in the prevalence of the use of numerical signals to mislead other group members. We confirm this with the following result.

Result 12. At the group level, the more subjects misrepresented their intentions in early periods both in their announcements of "possible" contributions and in their promise statements, the lower were group members' average contributions in later periods.

Table 10 reports OLS regressions at group level which parallel those in Table 6 but this time combining the data for the groups in the **NCTwP** and **NCTwP&R** treatments. In the first column, both average undercontribution ("lying") with respect to announced "possible" contribution and average undercontribution ("lying") with respect to announced promise (if any) are entered as explanatory variables, and both attract highly significant negative coefficients, with a somewhat larger magnitude for the promise term. The coefficients imply that for every experimental dollar of average underprovision relative to announced possible contribution in periods 1-4, average contributions per period were approximately one experimental dollar lower in periods 5-9, while for each experimental dollar of underprovision relative to the amount typed in the promise statement, average contributions were almost *E*$1.77 less during periods 5-9—rather large effects given the potential contribution range of 0 to 10 only. In the second column, the analysis is repeated using breaking of promise statements only. The result is that both the significance and the absolute value of the coefficient on this term rise somewhat, but its general affect is reconfirmed.

Table 10 about here

6. Discussion and Conclusions

---

had put in their last announcement of the "numerical cheap talk" stage, and this weakens the reliability of the Column 4 estimates. Nonetheless, the correlation is less than complete—the correlation coefficient is .7001—and our conclusion also rests on the independent evidence from the column 2 and 3 specifications.

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 would be expected to be meaningless babble.

Our 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. However, by disaggregating outcomes to the level of individual groups and to within-group interactions, we discover 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 in the early periods of play. Both opportunistic and truthful 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 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 common sense of that phrase, but not in the more demanding sense of standard economic theory which assumes common knowledge and payoff maximization; that theory would have talk be uniformly ignored!

In BPP (2004), we had speculated that one reason why numerical cheap talk was less effective overall than was verbal communication is that the former treatment prevented subjects from framing their announcements in the moral language of explicit promises. As a partial test of that conjecture, we describe here new experiments in which, in addition to typing "possible" decisions into the message space used in our "numerical cheap talk" treatment, 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 interpretation that promises were 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 manner. 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 the promises came to be taken more seriously than the ordinary announcements, permitting many groups to achieve high levels of cooperation.

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 own payoffs alone, and that (b) the overwhelming majority of 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 verbal 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 anothers' 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, 2004). Finally, the willingness of many to incur monetary costs in order to penalize free riders and those who deliberately mislead in their announcements and promises goes a long way toward explaining why the one-shot and finitely-repeated game predictions of standard theory are so frequently violated (Fehr and Gächter, 2000b).

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 their cooperation and the penalties and other mechanisms they will put in place to police their agreements. 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 form of their utility functions. Giving one's word alters subsequent play both 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

Andreoni, James and John H. Miller, 1993, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence," *Economic Journal* 103: 570-85.

Bochet, Olivier, Talbot Page and Louis Putterman, forthcoming, "Communication and Punishment in Voluntary Contribution Experiments," *Journal of Economic Behavior and Organization* (in press).

Brosig, Jeannette, Axel Ockenfels and Joachim Weimann, 2003, "The Effect of Communication Media on Cooperation" *German Economic Review* 4 (2): 217-42.

Cinyabuguma, Matthias, Talbot Page and Louis Putterman, "On Perverse Punishment in Voluntary Contribution and Punishment Games," in process, Brown University Dept. of Economics.

Crawford Vincent, 2003, "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentations of Intentions," *American Economic Review* 93: 133-149.

Crawford, Vincent and Joel Sobel, 1982, "Strategic Information Transmission," *Econometrica* 50 (6): 1431-1451.

Davis, Douglas D. and Charles A. Holt, 1993, *Experimental Economics.* Princeton: Princeton University Press.

Ertan, Arhan, Talbot Page and Louis Putterman, 2004, "Can Endogenous Institutions Mitigate the Free Rider Problem," unpublished paper, Brown University Dept. of Economics.

Falkinger, Josef, Ernst Fehr, Simon Gächter, and Rudolf Winter-Ebmer, 2000, "A Simple Mechanism for the Efficient Provision of Public Goods: Experimental Evidence," *American Economic Review* 90: 247-264.

Farrel, Joseph and Matthew Rabin, 1996, "Cheap talk", *Journal of Economic Perspectives* 10 (3): 103-118.

Fehr, Ernst and Simon Gächter, 2000a, "Cooperation and Punishment," American Economic Review 90: 980-94.

Fehr, Ernst and Simon Gächter, 2000b, "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14 (3): 159-81.

Frohlich, Norman and Joe Oppenheimer, 1998, "Some Consequences of e-mail vs. Face-to-Face Communication in Experiment," *Journal of Economic Behavior and Organization* 35 (3): 389-403.

Gunnthorsdottir, Anna, Daniel Houser, Kevin McCabe, and Holly Ameden, 2002, "Disposition, History and Contributions in a Public Goods Experiment," unpublished manuscript, Department of Economics and Economic Science Laboratory, University of Arizona.

Guttman, Joel, 2003, "Repeated Interaction and the Evolution of Preferences for Reciprocity," *Economic Journal* 113, no. 489: 631-656.

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

Isaac, R. Mark and James M. Walker, 1988, "Communication and Free-Riding Behavior: The Voluntary Contributions Mechanism," *Economic Inquiry* 26: 585-608.

Kreps, David, Paul Milgrom, John Roberts and Robert Wilson, 1982, "Rational Cooperation in Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* 27: 245-52.

Ledyard, John, 1995, "Public Goods: A Survey of Experimental Research," pp. 111-94 in John Kagel and Alvin Roth, eds., *Handbook of Experimental Economics*. Princeton: Princeton University Press.

Masclet, David, Charles Noussair, Steven Tucker and Marie-Claire Villeval, 2003, "Monetary and Non-Monetary Punishment in the VCM," *American Economic Review* 93(1): 366-80.

Offerman, Theo, Joep Sonnemans and Arthur Schram, 1996, "Value Orientations, Expectations, and Voluntary Contributions in Public Goods," *Economic Journal* 106: 817-45.

Ones, Umut and Louis Putterman, 2004, "The Ecology of Collective Action: A Public Goods and Sanctions Experiment with Controlled Group Formation," Working Paper No. 2004-01, Brown University Department of Economics.

Ostrom, Elinor, James Walker and Roy Gardner, 1992, "Covenants with and without a Sword: Self Governance is Possible." *American Political Science Review.* 86 (2): 404-416.

Page, Talbot, Louis Putterman and Bulent Unel, 2004, "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry, and Efficiency," Brown University Department of Economics Working Paper No. 2002-19, revised.

Palfrey, Thomas and Jeffrey Prisbrey, 1997, "Anomalous Behavior in Public Goods Games: How Much and Why?," *American Economic Review* 5: 829-846.

Sally, David, 1995, “Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992,” *Rationality and Society* 7 (1): 58-92.

Sefton, Martin, Robert Shupp and James Walker, 2002, “The Effect of Rewards and Sanctions in Provision of Public Goods,” Working Paper, University of Nottingham and Indiana University.

Wilson, Rick and Jane Sell, 1997, “”Liar, Liar...”: Cheap Talk and Reputation in Repeated Public Goods Settings,” *Journal of Conflict Resolution* 41 (5): 695-717.

Table 1. Summary of Treatments.

Reduction Option / Communication	None (Contribution Stage Only)	Contribution and Reduction Stages
No communication	Baseline (**B**) – 12 groups	Reduction (**R**) – 12 groups
Numerical announcements	Numerical Cheap Talk (**NCT**) – 11 groups	Numerical Cheap Talk with Reduction Option (**NCTwR**) – 11 groups
Numerical announcements and possible promise statements	Numerical Cheap Talk with Promise Option (**NCTwP**) – 11 groups	Numerical Cheap Talk with Promise and Reduction Options (**NCTwP&R**) – 11 groups

Note: Each treatment was carried out in 3 sessions of 4 or 3 groups of four subjects. Hence, a total of 280 subjects participated.

Table 2: 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**
Constant	-0.876** (0.383)	-1.112** (0.501)
(1st announced contribution by *i*) –(average 1st announced contribution by others in *i*'s group)	-1.010*** (0.085)	-0.955*** (0.092)
	N = 95, R^2 = 0.567	N = 82, R^2 = 0.467

Note: numbers in parentheses are Huber-White robust standard errors.

Table 3: 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**
Constant	-2.588* (1.427)	4.440*** (1.443)
Average binding contribution of group members other than *i* in period *t*-1	0.288*** (0.870)	0.140* (0.091)
Average last announced contribution of group members other than *i* in period *t*	0.124 (0.120)	0.146* (0.079)
I's last announced contribution in period *t*	0.298*** (0.089)	0.255*** (0.073)
N = 352	R^2 = 0.589	R^2 = 0.684

Note: numbers in parentheses are Huber-White robust standard errors. Regressions include individual fixed effects, not shown.

Table 4. Announced and actual reductions as a function of announced and actual contribution deviations and the deviation of actual from announced contribution.

Dependent variable:

	announced pun. "received" by *j*	actual pun. received by *j*	actual pun. received by *j*
Constant	-0.327 (2.829)	-2.576*** (0.720)	-3.00*** (0.745)
Abs. neg. dev.	1.839*** (0.197)	1.118*** (0.105)	1.105*** (0.104)
Abs. pos. dev.	-0.108* (0.395)	-0.533*** (0.179)	-0.505*** (0.178)
Avg. contrib. (*j* excluded)	-0.327 (2.829)	0.016 (0.083)	0.034 (0.083)
difference of annc'd. and actual contrib.			0.187*** (0.078)
N = 440	R^2 = 0.058	R^2 = 0.110	R^2 = 0.114
Log Likelihood	-801.610	-579.564	-576.720

Table 5: Responses to announced reductions in **NCTwR** communication periods

Independent variables	Dependent variable: first change in *j*'s announced possible contribution
Constant	-0.772^{***} (0.167)
Initial announced reductions, if *j* is not the maximum announced contributor (otherwise 0)	0.362^{***} (0.061)
Initial announced reductions, if *j* is the maximum announced contributor (otherwise 0)	-0.079 (0.059)
N = 440	$R^2 = 0.1943$

Note: numbers in parentheses are Huber-White robust standard errors.

Table 6: The negative impact of "lying" on the group performance in **NCT** and **NCTwR**

Independent variables	Dependent variable: Average contribution in group, rounds 5-9
Constant	6.666^{***} (0.731)
Average "lying" in group, rounds 1-4	-0.951^{***} (0.215)
N = 22	$R^2 = 0.266$

Note: numbers in parentheses are Huber-White robust standard errors.

Table 7: *p*-values of two-tailed Mann-Whitney tests of differences in group average contributions and group average earnings

	B	**R**	**NCT**	**NCTwR**	**NCTwP**	**NCTwP&R**
B	.	0.002	0.74	0.096	0.235	0.001
R	0.097	.	0.016	1	0.016	0.023
NCT	0.74	0.786	.	0.065	0.65	0.001
NCTwR	0.525	0.786	0.699	.	0.087	0.087
NCTwP	0.235	0.832	0.699	0.796	.	0.001
NCTwP&R	0.059	0.068	0.033	0.0233	0.013	.

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.

Table 8: Responses to announced reductions in **NCTwP&R** communication periods

Independent variables	Dependent variable: first change in *j*'s announced possible contribution
Constant	-0.610*** (0.138)
Initial announced reductions, if *j* is not the maximum announced contributor (otherwise 0)	0.169*** (0.058)
Initial announced reductions, if *j* is the maximum announced contributor (otherwise 0)	-0.190 (0.258)
N = 440	R^2 = 0.041

Note: numbers in parentheses are Huber-White robust standard errors.

(Table 9 is on the next page.)

Table 10: Impact of "lying" on the group performance, **NCTwP** and **NCTwP&R** treatments

	Dependent variable: Average contribution in group j. Round 5-9	
Constant	9.779*** (0.387)	9.324*** (0.367)
Average "lying" on contribution in group, periods 1-4	-1.053** (0.411)	
Average "lying" on promises in group, periods 1-4	-1.767*** (0.382)	-2.552*** (0.306)
N = 22	R^2 = 0.821	R^2 = 0.767

Note: numbers in parentheses are Huber-White robust standard errors.

Table 9: Punishment received as a function of broken promises and other variables, **NCtwP&R** treatment

	actual pun. received by *j*	actual pun. received by *j*	actual pun. received by *j*	actual pun. received by *j*
Constant	-4.006*** (0.961)	-3.228*** (0.924)	-4.611*** (0.974)	-3.741*** (1.250)
Abs. neg. dev.	1.401*** (0.143)	1.017*** (0.158)	1.093*** (0.154)	0.974*** (0.159)
Abs. pos. dev.	-0.087 (0.273)	-0.069 (0.272)	-0.101 (0.269)	-0.082*** (0.270)
Dummy = 1 if promise was made, 0 otherwise	-0.943 (0.903)	-2.312*** (0.944)	-0.727 (0.881)	-1.828* (0.992)
Dummy = 1 if promise was broken, 0 otherwise.		4.120*** (0.995)		2.923*** (1.252)
Dummy =1 if contributed less than last announced, 0 otherwise.			3.178*** (0.840)	1.588* (1.063)
N = 440	R^2 = 0.110	R^2 = 0.128	R^2 =0.125	R^2 = 0.131
Log Likelihood	-429.268	-420.266	-421.966	-419.167

FIGURE 1

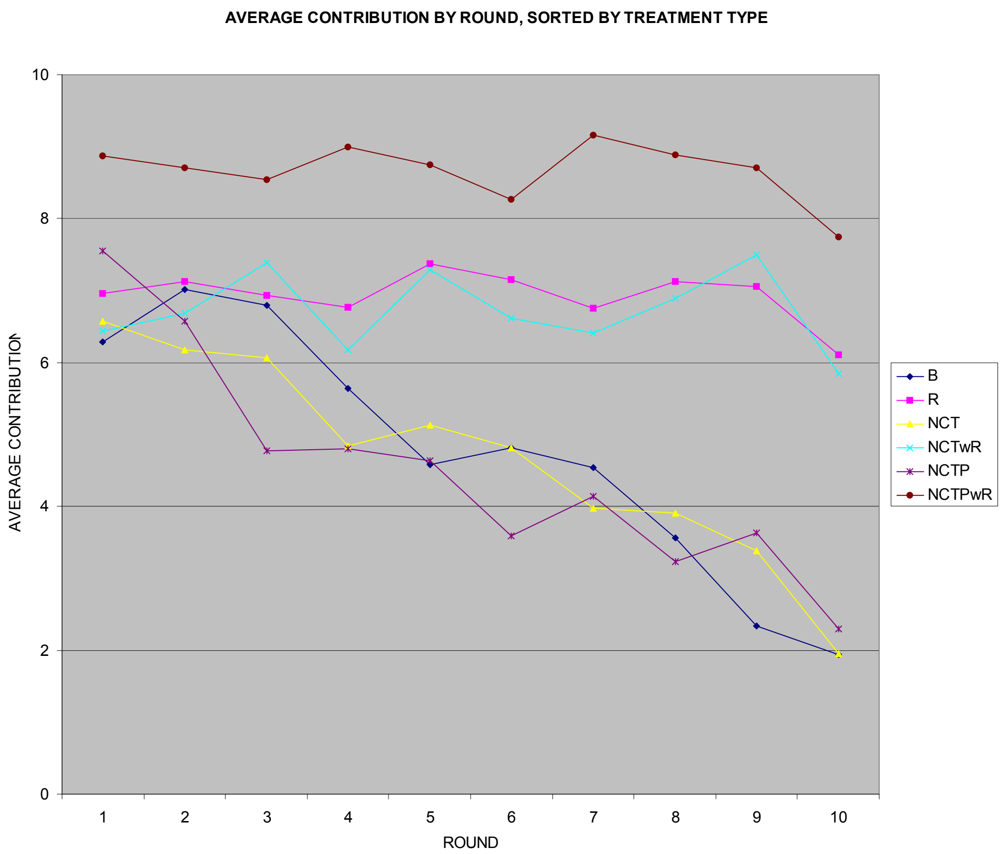
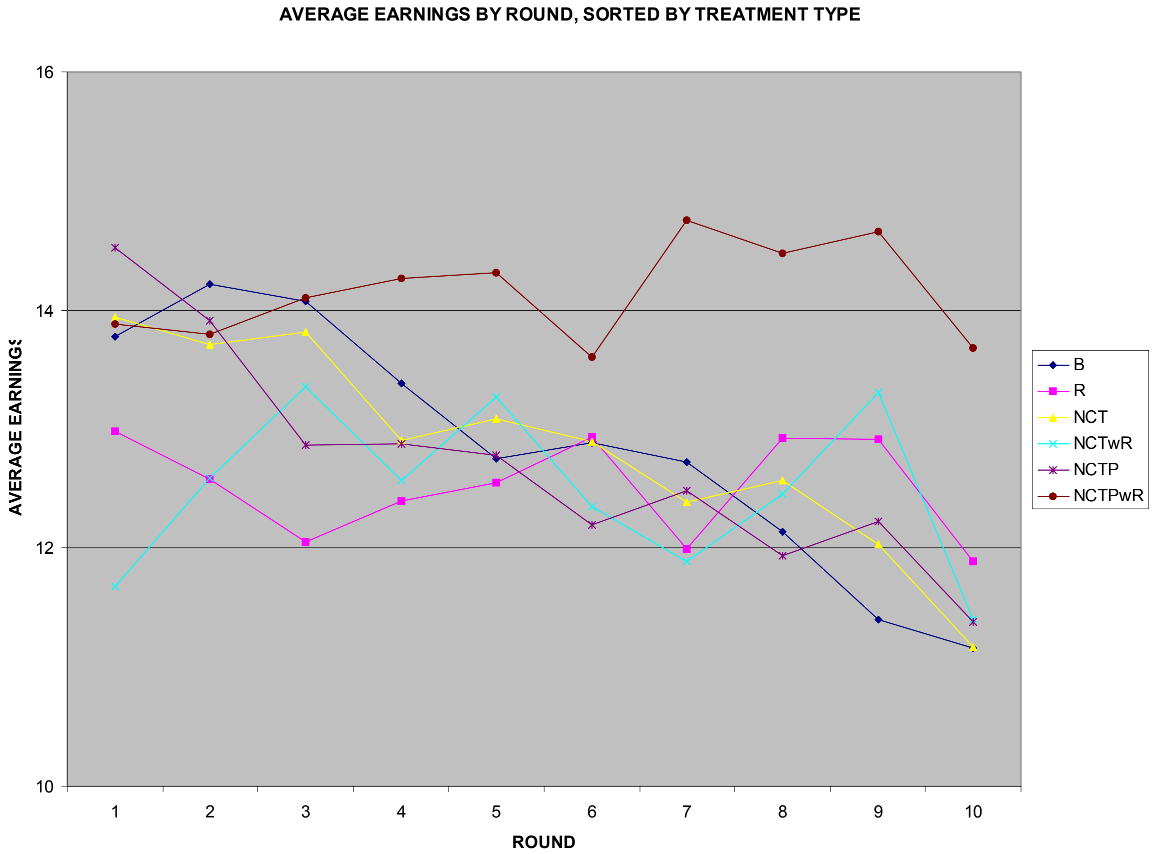
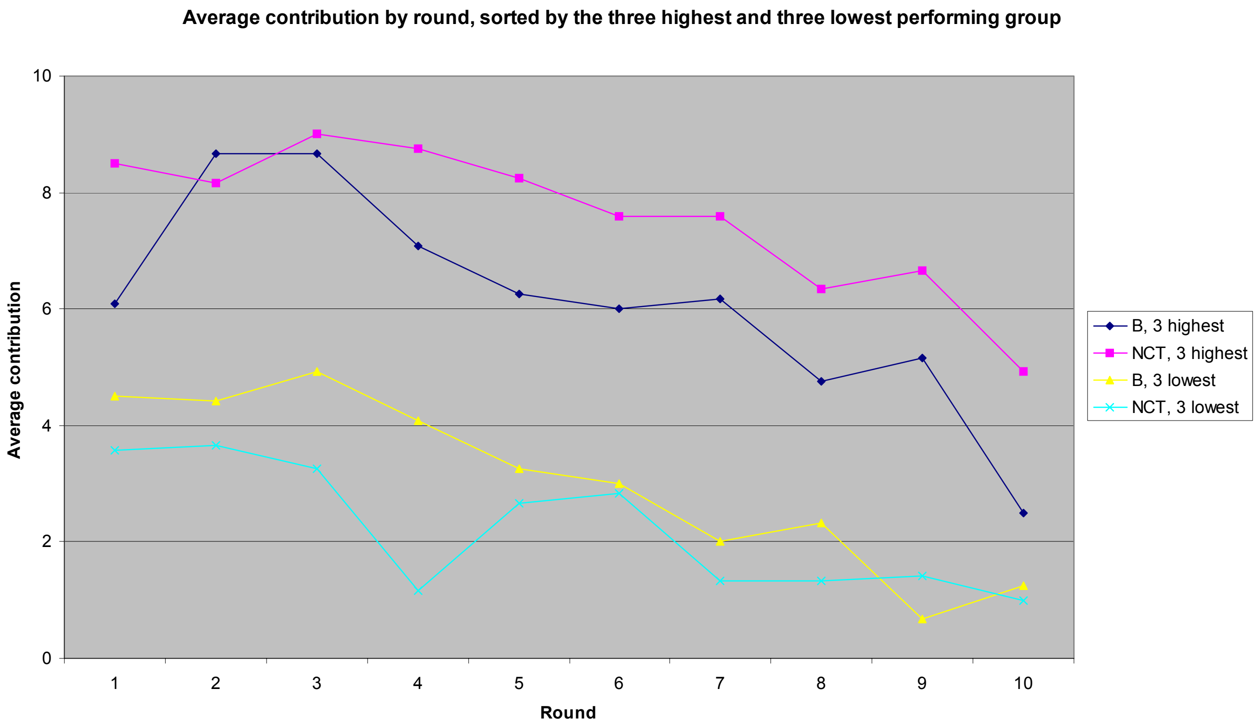

AVERAGE CONTRIBUTION BY ROUND, SORTED BY TREATMENT TYPE
AVERAGE CONTRIBUTION
ROUND
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10
B
R
NCT
NCTwR
NCTP
NCTPwR

FIGURE 2



AVERAGE EARNINGS BY ROUND, SORTED BY TREATMENT TYPE
16
14
12
10
AVERAGE EARNINGS
1
2
3
4
5
6
7
8
9
10
ROUND
B
R
NCT
NCTwR
NCTP
NCTPwR

FIGURE 3



Average contribution by round, sorted by the three highest and three lowest performing group
Average contribution
Round
0
2
4
6
8
10
1
2
3
4
5
6
7
8
9
10
B, 3 highest
NCT, 3 highest
B, 3 lowest
NCT, 3 lowest

FIGURE 4

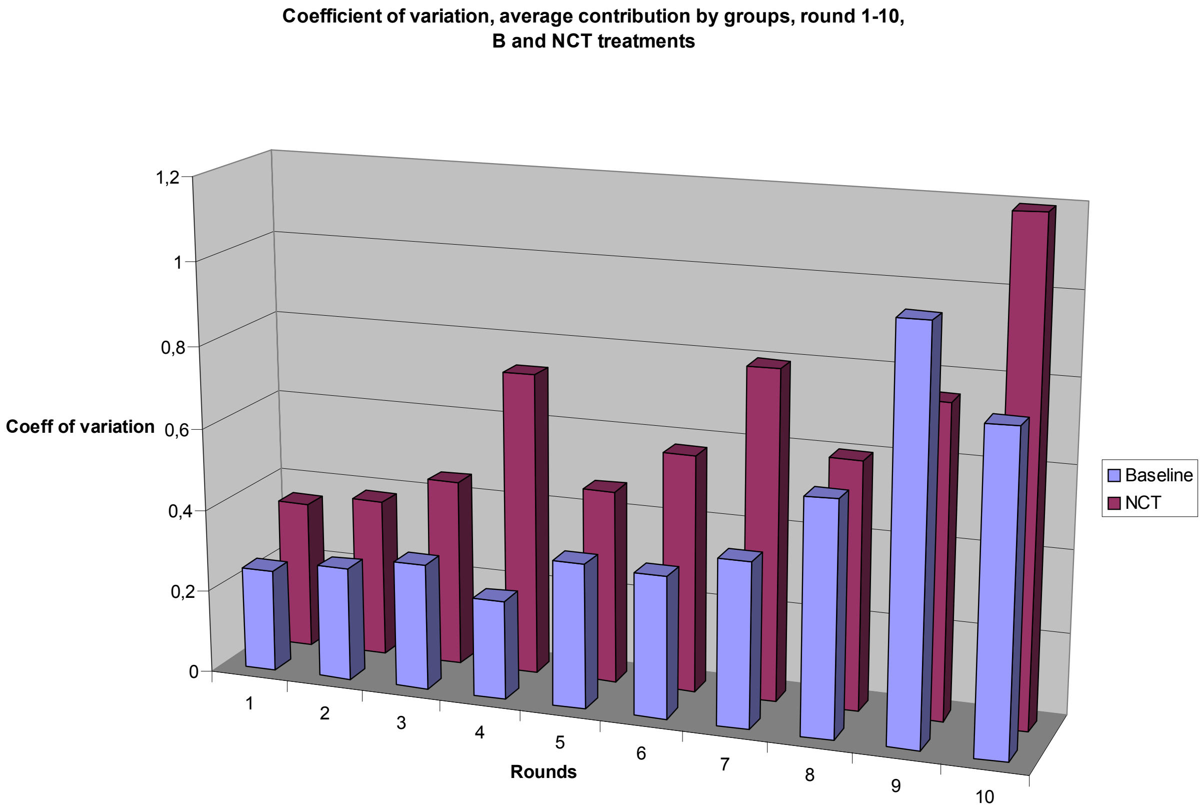

Coefficient of variation, average contribution by groups, round 1-10,
B and NCT treatments
Coeff of variation
1,2
1
0,8
0,6
0,4
0,2
0
1
2
3
4
5
6
7
8
9
10
Rounds
Baseline
NCT

FIGURE 5

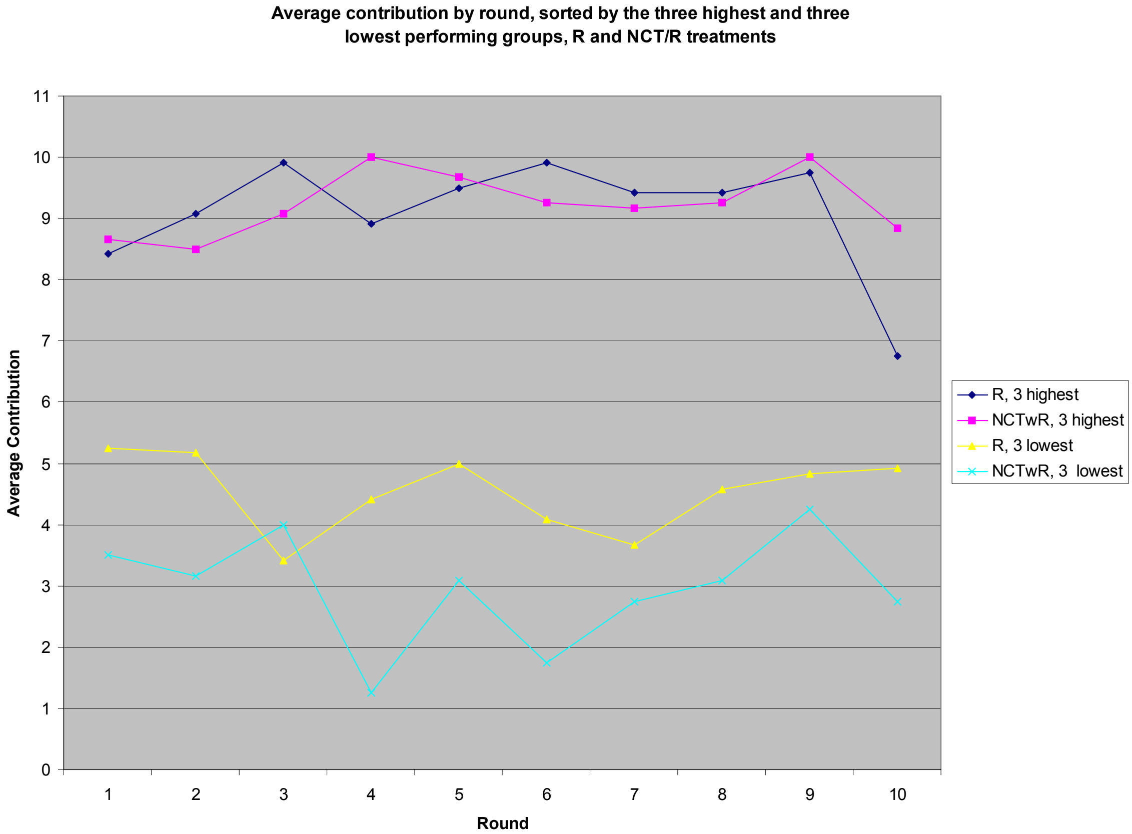
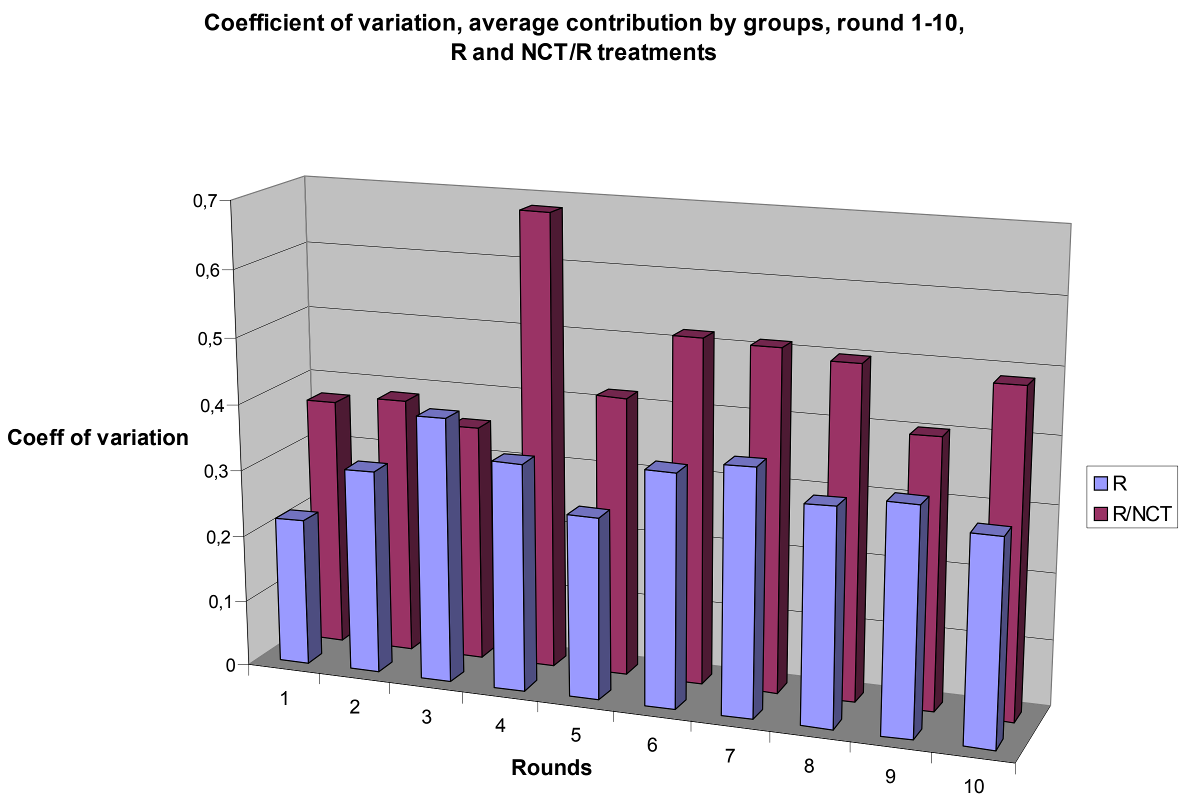

Average contribution by round, sorted by the three highest and three lowest performing groups, R and NCT/R treatments
Average Contribution
Round
R, 3 highest
NCTwR, 3 highest
R, 3 lowest
NCTwR, 3 lowest
0
1
2
3
4
5
6
7
8
9
10
11

FIGURE 6



Coefficient of variation, average contribution by groups, round 1-10,
R and NCT/R treatments
Coeff of variation
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
1
2
3
4
5
6
7
8
9
10
Rounds
R
R/NCT