Tilburg University

Which Words Bond? An Experiment on Signaling in a Public Good Game (replaced by CentER DP 2011-139)

Serra Garcia, M.; van Damme, E.E.C.; Potters, J.J.M.

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

No. 2010–33

WHICH WORDS BOND? AN EXPERIMENT ON SIGNALING IN A PUBLIC GOOD GAME

By Marta Serra-Garcia, Eric van Damme, Jan Potters

March 2010

ISSN 0924-7815

Which Words Bond?

An Experiment on Signaling in a Public Good Game*

Marta Serra-Garcia†, Eric van Damme‡ and Jan Potters§

March 26, 2010

Abstract

We compare signaling by words and actions in a one-shot 2-person public good game with private information. The informed player, who knows the exact return from contributing, can signal by contributing first (actions) or by sending a costless message (words). Words can be about the return or about her contribution decision. Theoretically, actions lead to fully efficient contributions. Words can be as influential as actions, and thus elicit the uninformed player's contribution, but allow the informed player to free-ride. The exact language used is not expected to matter. Experimentally, we find that words can be as influential as actions. Free-riding, however, does depend on the language: the informed player free-rides less when she talks about her contribution than when she talks about the returns.

JEL classification codes: C72; D82; D83.

Keywords: Information transmission; costly signaling; communication; experiment.

*We would like to thank Miguel Carvalho, Douglas DeJong, Eline van der Heijden, Andrew Schotter and seminar participants at the CESS experimental economics seminar at NYU, 2009 International ESA meetings, 2009 EEA-ESEM Meeting and Economics workshop at Tilburg University for their helpful comments and suggestions. We gratefully acknowledge funding from the Dutch Science Foundation (NWO), in the framework of the ESF-EUROCORES programme TECT.

†CentER, Tilburg University. Address: PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail:m.serragarcia@uvt.nl.

‡CentER and TILEC, Tilburg University. E-mail: eric.vandamme@uvt.nl.

§CentER, Tiber, TILEC and Netspar, Tilburg University. E-mail: j.j.m.potters@uvt.nl

1 Introduction

Popular proverbs about words and actions are abundant. While some say 'an Englishman's word is his bond', others say that 'actions speak louder than words' (Knowles, 2006). Indeed words can be just cheap talk (Farrell and Rabin, 1996). But can words speak as loud as actions? Furthermore, does the effectiveness of words depend on what words are spoken? Our aim is to compare words and actions in a public good game with private information, and vary the set of words (i.e., the language) that can be used.

In public good games, the influence of actions, or more precisely, of it being common knowledge that some actions are observed, has been widely studied. Theoretically, Hermalin (1998) and Vesterlund (2003), show that, if informed players contribute first to a team project or charity, they can 'lead by example': their contribution can elicit the contribution of uninformed players and enhance efficiency. Experimentally, Potters et al (2007) find support for these results[1]. The role of being allowed to talk about the return to a contribution, or about the size of the own contribution, however, has remained unexplored in contexts like these[2]. We examine the potential influence of words theoretically, and test the resulting hypotheses experimentally.

Our analysis proceeds in the context of a two-player one-shot public good game. The game is symmetric with respect to the players' contributions. The return to a contribution can take three different values, which are equally likely. If the return is low, it is individually rational and (Pareto) efficient not to contribute. If it is intermediate, the game is a prisoners' dilemma: it is efficient to contribute, but each player has an incentive to free ride. Finally, if the return is high, contributing is both individually rational and efficient. The exact state of nature, however, is only known to one of the players. The parameters are set such that, in case no signaling is possible, the uninformed player will not contribute. On the other hand, if the uninformed player knows that the return is either intermediate or high, and considers both possibilities to be equally likely, he will contribute. If no signaling is possible, the informed player only contributes when the return is high and the uninformed player never contributes, hence, contributions are inefficiently low.

[1] Several studies have investigated the effect of observing another player's contribution before deciding one's own (sequential moves) in complete information settings (e.g. Güth et al, 2007, Moxnes and van der Heijden, 2003). We consider a situation in which there is private information.

[2] The effect of communication in social dilemmas has been frequently studied, but in most cases, again assuming complete information (see Balliet, 2010, for a recent meta-analysis).

We compare two different kinds of signaling by the informed player: actions and words. In the first case, as in Potters et al (2007), the informed player moves first and her contribution is revealed before the uninformed player makes his contribution decision. The informed player now has an incentive to contribute if (and only if) the return is high or intermediate. Her contribution then signals to the uninformed player that he should contribute as well. Consequently, the actions of the informed player are influential: they determine the uninformed player's contribution. As both players contribute unless the returns are low, the game with signaling by actions produces a fully efficient outcome.

To study the effect of words, we allow for two different languages. The first language allows the informed player to talk about the return to a contribution. She can say 'the return is low', 'the return is intermediate', or 'the return is high'. The second language allows her to talk about her contribution decision. The informed player can say 'I do not contribute' or 'I contribute'. In both of these cases, talk is cheap, that is, the messages do not directly influence the payoffs.

The traditional cheap talk literature has focused on two disjoint classes of games (Farrell and Rabin, 1996): sender-receiver games with incomplete information, in which only the uninformed player takes payoff-relevant actions, and complete information games, where pre-play communication is used to foster coordination or cooperation. In the first case, the informed player is allowed to talk about her type (the private information); in the second case, she can talk about the action she intends to take. In our public good game, there is private information and both players take payoff-relevant actions. We allow the informed player to either talk about the return to a contribution (her type), or about the action she intends to take. The existing literature has shown that each type of communication can be effective in the respective class of games, and has investigated under which circumstances such communication is most effective. The game we employ allows us to investigate the effectiveness of these types of communication within one framework.

From a standard theoretical perspective, the exact language is irrelevant: for any language that allows at least two different messages, there are two pure equilibrium outcomes[3]. In the first equilibrium, words are ignored - considered as just cheap talk - and contribution levels are as in the game without signaling. In the second equilibrium, the informed player sends the same message (say G) when the state is intermediate and when it is high, and a

[3]The baseline game and the game with signaling through actions each have a unique equilibrium, and this is in pure strategies; the game with words also has mixed strategy equilibria.

different message (say B) when the return is low. The uninformed player contributes only after having heard G, hence, words can as be influential as actions.

Note that, for the two languages considered in this paper, all messages have a natural (or focal) meaning: although messages need not be believed, they will always be understood. Our work, hence, is in the tradition of Farrell (1985, 1993), who was the first to argue that messages having a literal meaning may destabilize certain equilibrium outcomes[4]. We show that, in our context, only the influential equilibrium outcome, is neologism-proof (Farrell, 1993), hence, we focus on this outcome. For the uninformed player, we thus predict the same behavior under words as under actions. In contrast, words allow the informed player to free ride when the return is intermediate. In the equilibrium with actions, this player is forced to contribute when the return is intermediate, but, since her contribution cannot be observed by the receiver in the case of words, theory predicts that she will contribute less in that case.

Existing theory thus predicts that (1) words can be as influential as actions (the informed player communicates the same information about the returns in both situations, to which the uninformed player responds in the same way); (2) the informed player will contribute less under words than under actions (as, under words, this player will free ride in the intermediate state); and (3) that it does not matter which words can be used. We test these hypotheses experimentally.

Our experiment reveals that words indeed can be as influential as actions. Informed players most frequently use the message 'the state is high' (resp. 'I contribute'), both when the state is intermediate and high, to which uninformed players react by contributing, as they do after observing a contribution of the informed player. Moreover, as predicted, when the state is intermediate, the rate of free riding by the informed player is much lower in case signaling is by actions (19% of the time) than in case signaling is by words (81% of the time, averaged across both languages). Still, in contrast to what theory predicts, it does matter what language is available. There are two key differences. First, while existing theory remains silent about which messages will be used, actual behavior displays important regularity: informed players strongly make use of the natural meaning of the words that are available. Secondly, and perhaps more striking, while free riding by the informed player

[4]There is a separate literature that builds on the presumption that messages, whilst not having an inherent meaning, may acquire meaning through an evolutionary learning approach; see Blume et al (2001) for a comparison of the two approaches.

is almost universal (94%) when talk is about the return, it falls significantly when she talks about her contribution (68%). In the specific case that the informed player says 'I contribute', she in fact contributes 41% of the time, revealing that for some players a word can be a bond.

We address both discrepancies in this paper. The first is rather easily dealt with by a theoretical extension of the ideas underlying Farrell's neologism-proofness concept: if uninformed players are likely to interpret messages according to their literal meaning, informed players will use messages according to their literal meaning, whenever this is a credible statement.

We suggest two, potentially complementary, explanations for the fact that the extent of free riding depends on the language that is available. Both explanations build on the idea that players dislike lying to some degree. The first explanation is in line with previous experimental studies, which find that lying depends on the associated consequences, that is, on the costs and benefits that follow from the lie (Gneezy, 2005, Hurkens and Kartik, 2009)[5]. In our game, not lying is less costly when talk is about the contribution than when talking about the returns. When talking about the return, if the informed player reveals the intermediate state truthfully, the uninformed player no longer contributes, which decreases the informed player's payoff substantially. In contrast, in talking about her contribution, the informed player can avoid lying at a low cost by indeed contributing if she says 'I contribute'. In this case, the uninformed player still contributes and the informed player does not forgo as much monetary payoff.

The second explanation elaborates on a similar idea by arguing that there may be different types of lies, and that some lies may be perceived as being more costly than others. In this respect, we note that the message 'I contribute' is similar to a promise, as it refers to an action of the speaker. In contrast, the message 'the returns are high' does not resemble a promise. The norm that promises should be kept may be stronger than the norm that one should not lie, and, therefore, players may be less likely to not contribute when they have announced a contribution. The similarity of the message 'I contribute' to 'I promise to contribute' could thus be a driving force behind the decrease in free-riding. In social

[5] See Kartik et al (2007) and Kartik (2009), among others, for models of sender receiver games in which such costs of lying are allowed. Demichelis and Weibull (2008) follow a similar approach, assuming that players have a lexicographic preference, after payoffs, for choosing an action which is in line with the meaning of the message they send.

dilemmas and trust games, with symmetric information, promises are often made and kept, especially when communication is free-form (Balliet, 2010, Charness and Dufwenberg, 2006, Ellingsen and Johannesson, 2004, Vanberg, 2008). Our experiment reveals a similar effect in a game of private information. It is noteworthy, however, and somewhat in contrast to these complete information studies, that we observe a relatively strong effect, even though we allow only a very restricted set of messages.

The contribution of our study, hence, is three-fold. First, we compare words and actions in a game with incomplete information and show that words can be as influential as actions. Previous studies comparing words and actions have only considered games of complete information (Bracht and Feltovich, 2009, Duffy and Feltovich, 2002 and 2006, and Wilson and Sell, 1997)[6]. Second, we slightly extend the reasoning underlying Farrell's neologism-proofness concept, show that it allows us to predict both messages and actions, and demonstrate that the prediction on which messages will be used is reasonably accurate. Third, we consider two different languages. In one case, the informed player can talk about her private information (returns), in the second case she can talk about her actions. We show that the language that is available matters for the informed player's own contribution. To the best of our knowledge, especially this latter aspect has remained unexplored in the literature on private information games[7].

The structure of the paper is as follows. In Section 2, we develop the theoretical framework, outlining the equilibria under actions and words. We then describe the experimental design in Section 3 and move to the results in Section 4. Section 5 concludes.

2 Theoretical Framework

We study a one-shot public good game with two players, one informed and one uninformed. The informed player has private information regarding the return of a contribution to the public good. The contribution's return, also called the state, s, can take three different values with equal probability, $S = \{a, b, c\}$, where $a \leq 0, 0 < b < 1$ and $c > 1$. Both the informed and the uninformed player decide whether to contribute or not to the project, where $x_i = 1$

[6] Also Brandts and Cooper (2007) compare words to financial incentives used by a 'manager' in a weak-link coordination game. Çelen et al (2009) compare advice to observation of other's actions in a social learning environment.

[7] Some previous studies have focused on the evolution of the strategic meaning of different sets of messages (Blume et al, 1998 and 2001, and Agranov and Schotter, 2009).

indicates a contribution and x_i= 0 none, with $i = \{I, U\}$. Whenever convenient, we will also denote the action of I by x and the action of U by y. The payoff function of the game is given by:

$$u_i = 1 - x_i + s(x_i + vx_j),\;\; j \neq i,\, j = \{I, U\}$$

where $v > 0$. Throughout we assume that $a+b+c < 3$, $b+c > 2$ and $b > 1/(1+v)$. These parameter restrictions imply: (i) against the prior distribution, the uninformed player's best response is not to contribute; (ii) if the uninformed player knows that the state is either b or c, and considers these to be equally likely, his best response is to contribute; (iii) if $s = a$, it is individually optimal and Pareto efficient not to contribute, while when $s = c$, the opposite is true; and (iv) it is socially optimal to contribute when the state is b.

Within this context, the baseline game does not allow any information transfer. In addition, we consider various games that allow signaling by the informed player. Under 'Actions', the informed player can signal through her contribution decision. In the case of 'Words', she can send a message, either about the state, or about her contribution decision. We, hence, consider four different games. In the subsections below we describe the equilibria of these games. Technical proofs are presented in the Appendix.

2.1 The Baseline Game

Let us first consider the Nash Equilibrium (NE) of the game when the uninformed player receives no signal. The strategy of the informed player is denoted as $\sigma = (x_a, x_b, x_c)$, where x_s denotes the probability of contributing in state s. The strategy of the uninformed player is specified as τ, the probability that he contributes.

Proposition 1 *The baseline game has a unique Nash Equilibrium, given by* $(\sigma^*, \tau^*) = \{(0, 0, 1); 0\}$.

In the unique NE of the game, only the informed player contributes, and then only if $s = c$. Since she cannot signal her private information to the uninformed player, the latter never contributes. However, if he would know that $s = c$, the uninformed player would prefer to contribute. Also, when $s = b$, neither player contributes while total payoffs would be maximized if both players did. Signaling the state with either words or actions can improve upon this outcome.

2.2 Actions

In the 'Actions' game, the informed player chooses her contribution x first; the uninformed player observes x and then chooses his contribution y. A strategy σ of the informed player is defined as above. Since the uninformed player can condition his decision on the observed choice of the other, his strategy space expands. A strategy τ of the uninformed player now is denoted as $\tau = (y_0, y_1)$, where y_z denotes the probability that the uninformed player contributes given $x = z$. The next Proposition states that, if the informed player can signal the return by revealing her contribution, both her contribution and that of the uninformed player increase. In particular, a contribution by the informed player is influential, as it leads to a contribution of the uninformed player as well.

Proposition 2 *The game with Actions has a unique Nash Equilibrium,* $(\sigma^*, \tau^*) = \{(0, 1, 1); (0, 1)\}$.

Note that signaling with the contribution decision ('leading by example') leads to a fully efficient NE. Players choose $x = y = 1$ when $s = b$ or $s = c$, while they choose $x = y = 0$ if $s = a$. This maximizes the sum of payoffs for each value of s.

2.3 Words

We introduce 'Words' by allowing the informed player to send a message m, from a given set of messages M, to the uninformed player. To allow that some information can indeed be transmitted, we assume that M contains at least two elements. The informed player first selects m, which is observed by the uninformed player before he decides about y. The uninformed player does not, however, observe x. The payoff function remains the same, hence, communication is costless.

Since the informed player observes the realization of s before sending a message, she can condition both her message and her contribution on the state of nature. We denote the strategy of the informed player as $\sigma = (\sigma_a, \sigma_b, \sigma_c)$ where $\sigma_s = (m_s, x_s)$. m_s is a probability distribution over M, and x_s is the probability of contributing in state s. Similarly τ specifies, for each $m \in M$, the probability $y(m)$ that the uninformed player contributes after the message m. We write $M_s(\sigma)$ for the set of messages in M that occur with positive probability when the state is s and σ is played. Similarly $X_s(\sigma)$ denotes the set of contributions that the informed player makes with positive probability when the state is s and σ is played.

Note that, since messages are costless, standard analysis leaves undetermined the messages that will be used, hence, there will always be multiple Nash equilibria. In Proposition 3, we, therefore, focus on the equilibrium outcomes: the contribution levels $x(s)$ and $y(s)$ in each state s.

There are two pure strategy equilibrium outcomes. In the equilibria of the first type, communication is uninformative, viewed as pure cheap talk, so that contribution levels are the same as in the baseline game. In the equilibria of the second type, the informed player's messages are influential, i.e. they induce the uninformed player to contribute when the state is b or c, but not when the state is a. In these equilibria, the informed player only contributes when $s = c$, hence, she free rides when $s = b$. We call these 'influential' equilibria.[8]

Proposition 3 *There are two pure strategy equilibrium outcomes in the game with Words, given by, respectively:*

(1) $X(\sigma) = (X_a(\sigma), X_b(\sigma), X_c(\sigma)) = (0, 0, 1)$ and
$\tau(m) = 0$ for all $m \in M_s(\sigma)$, where $s = \{a, b, c\}$

(2) $X(\sigma) = (X_a(\sigma), X_b(\sigma), X_c(\sigma)) = (0, 0, 1)$ and
$\tau(m) = 0$ for all $m \in M_a(\sigma)$,while $\tau(m) = 1$ for all $m \in M_b(\sigma) \cup M_c(\sigma)$

Introducing words can, hence, have two effects on contribution levels: a positive one, which increases the uninformed player's contribution levels, but not those of the informed player, or a null-effect, which leaves contribution levels as in the baseline case.

2.3.1 Words with a focal meaning: neologism-proof equilibrium

In this subsection, we show that only an influential equilibrium is neologism-proof, as defined in Farrell (1993). We also discuss why we consider this concept to be relevant in our context.

Thus far, we left the message space M to be an abstract set, and just assumed it to be large enough for partial separation. The existing game theoretic literature on 'cheap talk' can be divided into two classes. Most papers have assumed that messages do not have an a priori meaning, but that they may acquire a meaning through their use in equilibrium. Starting from Farrell (1985, 1993), there is a smaller literature that assumes that players

[8] There are also equilibria in which the informed player randomizes over messages, but these still yield the same contribution levels. If a=0, then there are also other mixed strategy equilibria in which some messages are used in all states of nature by the informed player; see the proof (in the Appendix) for details. We will focus on the pure strategy equilibria.

share a common language, in which messages have a natural, focal meaning. In this setting, although messages do not need to be believed, they will be understood. The idea is that, in such a context, players cannot (or will not) fully neglect the meaning that a message has outside of the specific game under consideration. In his seminal papers, Farrell has shown that, under this assumption, some equilibria are no longer plausible, since they can be destabilized by reference to the focal meaning of the messages; formally they are not neologism-proof. In the experiments that we conducted, see the next section, we used messages that have a literal meaning; hence, our work is in this second tradition. We will show that only an influential equilibrium is neologism-proof[9].

Strictly speaking, however, there are two reasons why the neologism-proofness concept is not directly applicable to our context. First, our public goods game with 'Words' is not of the type that has been considered in the traditional cheap talk literature, as it is a game with private information in which both players take payoff-relevant actions. Nevertheless, the informed player, I, has a strictly dominant contribution level $x_I(s)$, in each state of nature s. If we assume that I will always choose this contribution, we are back in the standard setting, to which Farrell's ideas can be applied[10]. Second, and perhaps more important, the interpretation of Farrell's concept relies on the players having a *rich* language at their disposal. In our experiments, we used a restricted language. We return to this aspect after having given the formal definition and having formulated the result.

For a subset T of S write $b_U(T)$ for the best response of player U, given the prior, but conditional on the state s being in T. Let $e = (\sigma, \tau)$ be an equilibrium and denote by $u_I^e(s)$ the equilibrium payoff of player I, given that the state is s. Farrell (1993) defines the set T to be self-signaling with respect to e if

$$T = \{s \in S : u_I(s, b_U(T)) > u_I^e(s)\}$$

[9] Rabin (1990) has argued that Farrell's definition rules out too many equilibrium outcomes. For further discussion, see also Farrell and Rabin (1996). It can, however, be shown that, if $a<0$, only the influential equilibrium satisfies Rabin's condition of Credible Message Rationalizability. If $a=0$, player I is indifferent between all responses of player U and also the uninfluential equilibrium satisfies CMR. If I would have social preferences and attach some positive weight to the utility of U, then I strictly prefers U to choose $y=0$ if $a=0$, and in this case again only the influential equilibrium is CMR. Details are available from the authors upon request.

[10] It is innocuous to make this assumption as also the best reply of the uninformed player only depends on the posterior distribution over the states and not on the contributions that player I makes.

and he defines the equilibrium e to be *neologism-proof* if there is no set of types T that is self-signaling with respect to it. The interpretation is as follows. Suppose e is the equilibrium under consideration, and suppose that player I says "the state belongs to the set T". If player U interprets the message literally, he will be inclined to choose $b_U(T)$. On the other hand, player U should not be credulous, but rather ask himself the question: when does player I have an incentive to use this message, assuming that it would be believed? If T is self-signaling, player I strictly benefits from using the message 'the state is in T' exactly when this statement is true. When T is self-signaling, there are good arguments to believe this message as the literal meaning of the message 'the state is in T' is consistent with the incentives that the game provides. Consequently, if an equilibrium e is not neologism-proof, and the language that is available to the players is rich enough to allow a self-signaling set to identify itself, e can be upset by the corresponding self-signaling message. We have

Proposition 4 *Only an influential equilibrium is neologism-proof.*

The proof relies on the fact that the set $T = \{b, c\}$ is self-signaling. If the informed player uses the message "the state is b or c", the uninformed player should thus believe her. Farrell (1993) assumes that players have a *rich* natural language at their disposal, so that this message is available. In our experiments, although we used messages with a natural meaning, we did *not* use a rich language. In particular, in none of the two games that we experimented with was the message "the state is b or c" available. Nevertheless, in each of these games, there were messages (such as "the state is c" or "I contribute") available, that could naturally be interpreted like this. In other words, the self-signaling set $\{b, c\}$ might be able to signal through a different message than "the state is in $\{b, c\}$". Furthermore, although the *interpretation* of Farrell's concept relies on this richness assumption, the formal definition only refers to the mathematical structure of the game under consideration. For both of these reasons, we believe that the concept is relevant to our game.

It should be noted that, although the concept of neologism-proofness limits the number of equilibrium *outcomes* to one, it does not lead to restrictions on the messages that will be used. As already mentioned in the context of Proposition 3, there are multiple *equilibria*. For example, consider the case discussed in the Introduction, where there are (at least) the messages B (Bad) and G (Good). In this case, in one neologism-proof equilibrium, player I sends the message B when $s = b$, or $s = c$, to which player U responds with $y = 1$, while player I uses the message G when $s = a$, which is then followed by the response $y = 0$.

In another equilibrium, player I sends the message G if $s = b$, or $s = c$, (with response $y = 1$), while the message is B if $s = a$ (with response $y = 0$). Formally, according to the logic of the concept, both of these equilibria are neologism-proof. Nevertheless, the latter equilibrium seems more natural than the first. After all, in this latter equilibrium, player I communicates that the state is Bad exactly when this is the case, while she communicates that the state is G, when it is not bad. In other words, the latter equilibrium is closer to the truth than the former.

2.3.2 Talking about the state or talking about the contributions

To further develop the above idea, let us now focus on the two specific message sets that will be discussed in the remainder of this paper. In the first case, $M= M^s = \{a, b, c\}$, so that messages correspond to the state of nature[11]. In the second case, $M = M^x =\{x = 1, x = 0\}$, the messages correspond to the contribution decision of the informed player. To select among the equilibria, hence, to also pin down the messages that will be used, we make two assumptions, each of them corroborated by extensive experimental evidence. The first assumption is that players (or at least some of them) have at least a minimal aversion to lying. Several experiments (e.g. Gneezy, 2005, Sánchez-Pagés and Vorsatz, 2007, and Hurkens and Kartik, 2009) have shown that players dislike lying. As in Demichelis and Weibull (2009), we adopt a very minimal version of this idea, namely that, when the material payoffs are the same, players prefer not to lie[12].

This assumption is sufficient to obtain a unique, focal, equilibrium in the case where messages are about the contribution of the informed player, $M^x =\{x = 1, x = 0\}$. In this case, there are two pure equilibria that produce the influential equilibrium outcome. In the first, I sends the message $x = 0$ when $s = a$ and the message $x = 1$ when $s = b$, c. In the second, messages are reversed: I says $x = 1$ when $s = a$ and says $x = 0$ when $s = b$, c. In the first equilibrium, I tells the truth when $s=a$ and c; in the second, she always lies. We

[11] We chose this set of messages because it is precise and corresponds directly to the informed player's private information. In Serra-Garcia et al (2008) we consider a richer set of messages allowing for two or more states to be stated in one message and a blank message. In that paper, the action of the informed player is observed by the uninformed player as well as the informed player's message. We find that players' contribution behavior is not significantly affected by the richer message space, but that informed players are often vague.

[12] We note that Farrell (1993, p. 519) also explicitly refers to players having a slight preference for telling the truth in order to justify his refinement of neologism-proofness.

consider the first equilibrium to be focal.

Now consider the case in which player I can talk about the state, but is required to provide full (precise) information, $M^s = \{a, b, c\}$. Table 1 describes the 6 message combinations that are possible in the various influential pure equilibria.

	Message sent if state			
Equilibrium nr.	a	b	c	# states lie
1	a	b	b	1
2	a	c	c	1
3	b	a	a	3
4	b	c	c	2
5	c	a	a	3
6	c	b	b	2

Table 1: Message use in influential equilibria and lies

An argument as above points in the direction of the first or the second equilibrium, but it does not discriminate between those. Nevertheless, we argue that only the second equilibrium is focal. The additional assumption leading to this conclusion is that a small but positive portion of uninformed players is naïve and interprets messages literally and naïvely. Such an assumption is also used in Crawford (2003), Kartik et al (2007) and Ellingsen and Östling (2009). Experiments have indeed shown that some receivers are credulous and interpret messages literally and naïvely (e.g. Cai and Wang, 2006). Under this additional assumption, only the second equilibrium is focal. Since player I wants to induce U to contribute when the state is b or c, and U might interpret messages literally, I uses message c. He assumes that U will react to the unused message b by interpreting it literally and, hence, by not contributing. Note that the natural language reinforces the equilibrium. For this reason we call this equilibrium focal.

We have proved:

Proposition 5 *The games with words M^s and M^x each have a unique focal equilibrium. The focal equilibrium is influential. If player I talks about the state, she will reveal it when it is a, whilst she will say c when the state is b or c. Alternatively, when player I talks about her contribution, she will honestly reveal her contribution when the state is a or c, but she will lie and say that she contributes in state b. Player I only contributes when the state is c, and player U only does so after message c or after a message stating that I contributes.*

3 Experimental Design and Hypotheses

3.1 Parametrization and Treatments

In the experiment, the payoff function of our game is the following, $u_i = 40[1 - x_i + s(x_i + vx_j)]$, where $s = \{0, 0.75, 1.5\}$ and $v = 2$. Subjects are asked to choose between A (equivalent to $x_i = 0$) and B (equivalent to $x_i = 1$) in each round. The payoffs of a player depend on her choice, the choice of the other player and the earnings table selected. The earnings table number (1,2 or 3) corresponds to the value of s (s = 0, 0.75 or 1.5, respectively). Payoffs (in points) are shown in Table 2 for each earnings table number. These tables were shown to subjects both in the instructions (reproduced in the supplementary material[13]) as well as on the computer screens.

		Earnings Table 1			Earnings Table 2			Earnings Table 3	
		Other person's choice			Other person's choice			Other person's choice	
		A	B		A	B		A	B
Your choice	A	40	40	A	40	100	A	40	160
	B	0	0	B	30	90	B	60	180

Table 2: Payoff Matrices

In all treatments, at the beginning of each round, the informed player, named first mover in the experiment, is informed about the earnings table selected, and next decides whether to contribute or not. In the Baseline, the uninformed player, named second mover, receives no information and is simply asked to make a decision. In Words and Actions, the uninformed player first receives the signal from the informed player and is then asked to make a decision. In Actions, the signal is the decision of the informed player (A or B). In Words, the informed player is explicitly asked to also select a message to send to the uninformed player. In Words(s), the three possible messages are 'The earnings table selected by the computer is s', where s is either 1, 2 or 3. In this game, the informed player thus talks about the state. In Words(x), two messages are possible: 'I choose A' or 'I choose B'. In this game, the informed player thus talks about (her) contributions. The roles of informed and uninformed player are randomly determined within each pair in each round. The information available in each treatment is detailed in Table 3 below.

[13] The supplementary material can be downloaded at http://center.uvt.nl/phd_stud/garcia/WWB.rar.

	Informed player	Uninformed player
Baseline	Observes s	No information
Words(s)	Observes s	Observes $m \in M^s$
Words(x)	Observes s	Observes $m \in M^x$
Actions	Observes s	Observes x

Table 3: Experimental Design - Information Structure by Treatment

In each period, both players have a history table at the bottom of their screens, displaying the following information for each previous period: the earnings table selected, the role of the player, the own decision and that of the other player, including the message sent if applicable, and the earnings of both players. From this information, players could not identify the players with whom they had previously played.

3.2 Hypotheses

We take the results from Propositions 1 to 5 and summarize the equilibrium contributions of the different treatments in Table 4, below. The informed player never contributes when s=0, and always does when s=1.5. When s=0.75, she only does in Actions, that is, if her contribution is observed. The reactions of the uninformed player range from never contributing (as in Base) to imitating the informed player (in Actions).

	Choices[a]		
Treatment	s=0	s=0.75	s=1.5
Baseline	$(0, 0)$	$(0, 0)$	$(1, 0)$
Words	$(0, 0)$	$(0, 1)$	$(1, 1)$
Actions	$(0, 0)$	$(1, 1)$	$(1, 1)$

Note: [a] (x, y)

Table 4: Expected Choices

The hypotheses 1 and 3 are derived from the contribution behavior of both players as described in this table. Hypothesis 2 focuses on the communication between the players and is derived from Proposition 5. Relatedly, the efficiency[14] (ξ) of each treatment can be ranked as follows: ξ_{Base} (61.3%) $\leq$ $\xi_{Words(s) \text{ and } (x)}$(91.9%) $<$ $\xi_{Actions}$ = (100%).These

[14] Efficiency is calculated throughout the paper as the sum of payoffs of the leader and the follower in each treatment, divided by the maximum sum of payoffs attainable.

inequalities lead to hypothesis 4[15].

Hypothesis 1 (informed player contribution behavior): when s=0.75, the informed player contributes:

(a) more frequently under Actions than in Words(s) or in Words(x)

(b) with equal frequency in Words(s) as in Words(x).

Hypothesis 2 (message use and information transmission):

(a) if s=0, the message 'the state is 0' is used in Words(s), whilst the message 'I do not contribute' is used in Words(x). If s=0.75 or s=1.5, the messages that are used are 'the state is 1.5' and 'I contribute', respectively.

(b) the same information is transmitted in Words(s), Words(x) and Actions.

Hypothesis 3 (uninformed player contribution behavior): the messages 'the state is 1.5' and 'I contribute', in Words(s) and Words(x), respectively, are as influential as a contribution is in Actions.

Hypothesis 4 (efficiency):

(a) efficiency is highest under Actions, compared to all other treatments.

(b) efficiency under Words(s) is equal to that under Words(x).

3.3 Experimental Procedures

Four matching groups (of 8 subjects each) participated in each treatment. Subjects were re-paired every period with another subject in their matching group and roles were randomly assigned. To have enough learning possibilities for each earnings table (value of s), subjects played the game for 21 periods. Further, since there were 8 subjects in each matching group, each subject met the same person at most 3 times, without coinciding two consecutive periods in the same role. Overall, 84 pairings were obtained per matching group (4 pairs x 21 periods): 25 faced Earnings Table 1, 30 Earnings Table 2 and 29 Earnings Table 3[16]. The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). It was conducted in CentERlab, at Tilburg University. Subjects received an invitation to

[15] We do not formulate a hypothesis about payoffs since the treatment effects are expected to be small for the informed player's payoffs. We briefly discuss predicted and actual payoffs in Section 4.4.

[16] The matching schemes, roles and states of nature for each period and pair were randomly drawn before the experiment. This allowed us to have the same exact patterns across different matching groups.

participate in the experiment via e-mail. They could enrol online to the session of the experiment, which was most convenient for them, subject to availability of places. Subjects were paid their accumulated earnings in cash and in private at the end of the experiment. Average earnings were 12.20 Euro (sd: 2.46) and sessions lasted approximately 60 minutes.

4 Results

We report results from the second half of our experiment (periods 11 to 21). This is motivated by the fact that, in the first 10 periods, informed players exhibit strong learning for s=0.75. Our unit of observation will be each matching group in the experiment; we thus have 4 independent observations per treatment[17].

4.1 Contributions by the informed player

The informed player's contribution decision is determined by two main factors. The first one is the state, s, and the second one is the treatment. In Figure 1, we observe the average frequency with which informed players contribute by state and treatment.

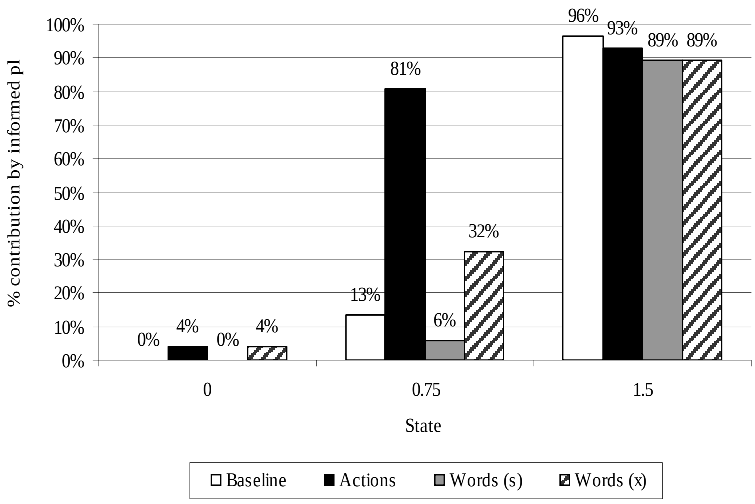

Figure 1: Contribution Frequency by Informed Player, by State and Treatment

The four leftmost columns of Figure 1 reveal that, when $s = 0$, the informed player contributes between 0 and 4% of the time. In contrast, when $s = 1.5$ (four rightmost

[17]The z-tree code, raw data and stata code are provided in the supplementary material.

columns), she contributes approximately 90% of the time. In neither of these cases is there a significant difference across treatments (Kruskall-Wallis test, p-value=0.1718 and 0.8152, respectively).

Treatment differences become significant when $s = 0.75$. First, the informed player contributes significantly more often (81% of the time) in the Actions treatment, when her contribution is observed, than in any other treatment (Mann-Whitney (MW) test, p-value=0.0194 comparing Actions to Baseline, or Actions and Words(s); p-value=0.0202 comparing Actions and Words(x)).

The informed player's contribution is also affected by the words she can use. When the informed player talks about her contribution decision, her contribution frequency increases to 32%, compared to 6%, when she talks about the state (MW test, p-value=0.0421).

Result 1 (contributions of the informed player):

(a) When $s = 0.75$, the informed player's contribution is higher in Actions than in Words(s) and in Words(x). Thus, we do not reject Hypothesis 1 (a).

(b) The contribution frequency of the informed player is also affected by the language that is available. The informed player contributes more often when sending messages about her contribution (Words(x)), than when she sends messages about the state (Words(s)). We, thus, reject Hypothesis 1 (b).

In contrast to what standard theory predicts, it, hence, matters what the informed player can talk about. We will examine this result in more detail at the end of this section, after having studied the use of messages by the informed player, the information transmitted through these messages, and the reaction of the uninformed player.

4.2 Message use and information transmission

In Table 5, we display the informed player's message use in Words(s) and Words(x). The rows display the possible messages and the columns the frequencies with which they are used in the various states. For example, in treatment Words(s), the message 'the state is 0' is used 71.1% of the time when $s = 0$.

		Message use[a]		
Treatment	Message (m)	s=0	s=0.75	s=1.5
Words(s)				
	'The state is 0'	71.1%	8.8%	1.8%
	'The state is 0.75'	11.6%	16.2%	3.6%
	'The state is 1.5'	17.3%	75.0%	94.7%
Words(x)				
a)	*Matching groups 13,15 and 16*			
	'I do not contribute'	94.9%	23.5%	9.5%
	'I contribute'	5.1%	76.5%	90.5%
b)	*Matching group 14*			
	'I do not contribute'	61.5%	17.6%	28.6%
	'I contribute'	38.5%	82.4%	71.4%

Note: [a] Number of times m is sent over total number of times that s is drawn

Table 5: Message use in Words(s) and Words(x), by treatment and state

Let us first focus on the Words(s) treatment. When $s = 0$, informed players most frequently use the message 'the state is 0' (71.1%). Instead when $s = 0.75$ or $s = 1.5$, the informed player most frequently uses the message 'the state is 1.5' (75% and 94.7%, respectively). The frequency with which this message is used in these states is not significantly different (Wilcoxon signed-rank (WSR) test, p-value=0.1441). Note that, when $s = 0$ or $s = 1.5$, the informed player most frequently tells the truth, but that, when $s = 0.75$, lies are very frequent. In any case, the natural meaning of the words plays a role.

Let us now turn to Words(x). In this treatment, we observe differences in message use across matching groups. Three matching groups (the groups 13, 15 and 16), use messages as expected in the focal equilibrium, while one matching group (group 14) does not. In this matching group, when $s = 0$, the message 'I contribute' is sent much more frequently than in any other matching group (38.5%, versus 0% in matching group 13, or 7.7% in groups 15 and 16). Furthermore, in this group 14, the message 'I contribute' also is used more often when $s = 0.75$ than when $s = 1.5$. We find that this difference in message use in matching group 14 has important consequences in terms of the information transmitted by the informed player. In the tables that follow, we therefore report separate statistics for this group[18].

[18] In treatment Words(s) we find no substantial differences across matching groups and, therefore, report averages across all matching groups throughout.

In matching groups 13, 15 and 16, when $s = 0$, the informed player most frequently says 'I do not contribute' (94.9%). When $s = 0.75$ or $s = 1.5$, she most frequently sends the message 'I contribute' (76.5% and 90.5%). Again, the frequency with which she sends this message does not differ significantly between these two states (WSR test, p-value= 0.2850). We also here see that the natural meaning of the message plays a role.

To consider the information transmitted in Actions, Words(s) and Words(x), we now take the behavior of the informed player during periods 11 to 21 and calculate (using Bayes' rule) the posterior probability that the state is s, given the signal received. Table 6 displays the results. The rows represent the different signals (distinguished also by matching group in the case of Words(x)), while the final three columns give the posterior probability of each state.

		Probability that		
Treatment	Signal	s=0	s=0.75	s=1.5
Actions	Informed player's decision			
	x=0	0.75	0.18	0.06
	x=1	0.02	0.5	0.48
Words(s)	Message about the state			
	'The state is 0'	0.85	0.13	0.02
	'The state is 0.75'	0.18	0.54	0.28
	'The state is 1.5'	0.07	0.44	0.48
Words(x)	Message about the contribution			
a)	*Matching groups 13,15 and 16*			
	'I do not contribute'	0.70	0.23	0.07
	'I contribute'	0.03	0.49	0.48
b)	*Matching group 14*			
	'I do not contribute'	0.53	0.20	0.27
	'I contribute'	0.17	0.48	0.35

Table 6: Posterior probability of each state conditional on signal by informed player

In Actions, after a contribution ($x = 1$), the probability that $s = 0.75$ is 0.5, while the probability that $s = 1.5$ is 0.48. Instead, if the informed player does not contribute, the probability that $s = 0$ is 0.75.

In Words(s), after the message 'the state is 1.5' the probability that $s = 0.75$ is 0.44. This probability is not significantly different from the corresponding probability, 0.5, after a contribution in Actions (MW test, p-value=0.1489). The probability that $s = 1.5$ is 0.48, which again is not significantly different from that after a contribution in Actions (MW test, p-value=1.000). This message therefore did not transmit significantly different information than a contribution decision of the informed player, in Actions. Furthermore, the probability that $s = 0$ after the message 'the state is 0' (0.85) is not significantly different from that (0.75) after no contribution by the informed player in Actions (MW test, p-value=0.2482).

In the treatment Words(x), for matching groups 13, 15 and 16, after a message 'I contribute', the probability that $s = 0.75$ is 0.49, and that of $s = 1.5$ is 0.48. These are not significantly different to those after a contribution in the Actions treatment (MW test, p-value=0.5637 for state 0.75 and 0.4678 for state 1.5). Furthermore, again excluding matching group 14, the probability that $s = 0$ after the message 'I do not contribute' (0.70,) is not significantly different from that (0.75) after no contribution in Actions (MW test, p-value 0.1102). Instead, for matching group 14, the probability that $s = 1.5$, after the message 'I contribute' is 0.35.

Result 2 (message use and information transmission):

(a) In Words(s), the message 'the state is 0' is most frequently used when s=0, while the message 'the state is 1.5' is most frequently used when s=0.75 or 1.5. In Words(x), 'I do not contribute' is most frequently used when s=0, and 'I contribute' is used most often when s=0.75 or 1.5 (especially in matching groups 13, 15 and 16). We therefore do not reject Hypothesis 2a.

(b) Compared to a contribution decision in Actions, the message 'the state is 1.5' in Words(s), or the message 'I contribute' in Words(x) (except in one matching group) does not convey significantly different information. Compared to no contribution in Actions, the messages 'the state is 0' and 'I do not contribute' also do not convey significantly different information. Thus, we do not reject Hypothesis 2b.

4.3 Contributions by the uninformed player

The uninformed player reacts to the information transmitted by the informed player. In Table 7, rows again display the different possible signals. Column (1) gives the average contribution frequency of the uninformed player. Columns (2) and (3) give the expected payoff in points from not contributing, or contributing, calculated using the posterior probabilities displayed in Table 6, as well as (for Words(s) and Words(x)), the frequency with which the informed player contributes conditional on each message sent. The last column of Table 7, (4), displays the empirical best reply, based on the expected payoff calculation. The choice with the highest expected payoff is then displayed for each signal.

In the baseline treatment, the first row in Table 7, the uninformed player receives no signal but contributes 39.2% of the time. This is an unexpectedly high level of contributions, since the empirical best reply is not to contribute. This contribution rate is, however, similar to that in Ellingsen and Johannesson (2004), who find that 35% of sellers invest when there is no communication, despite the prediction of no investment. One possible explanation in our game is that individuals try to 'guess' when the state will be high and that they fall prey of the 'gambler's fallacy' (Kahneman and Tversky, 1974). For example, the likelihood of a contribution decreases in the period after the state was 1.5, despite the fact that players are informed that in every period the state is 0, 0.75 or 1.5 with equal probability. Another possible explanation is that social preferences play a role. After all, with an expected value of s of 0.75 it is socially efficient to contribute.

		(1)	(2)	(3)	(4)
		Uninformed Player's	Expected Payoffs		Empirical
Treatment	Signal	Contribution Frequency	$\pi(y{=}0)$	$\pi(y{=}1)$	best reply
Baseline	-	39.2%	81.22	71.22	$y{=}0$
Actions	$x{=}0$	4.4%	40.00	9.27	$y{=}0$
	$x{=}1$	88.0%	127.77	131.65	$y{=}1$
Words(s)	'The state is 0'	2.3%	43.67	8.60	$y{=}0$
	'The state is 0.75'	42.0%	71.67	64.69	$y{=}0$
	'The state is 1.5'	69.7%	93.30	95.51	$y{=}1$
Words(x)					
a)	*Matching groups 13,15,16*				
	'I do not contribute'	7.6%	53.40	24.67	$y{=}0$
	'I contribute'	62.3%	109.69	113.35	$y{=}1$
b)	*Matching group 14*				
	'I do not contribute'	13.3%	52.00	34.00	$y{=}0$
	'I contribute'	13.8%	71.03	66.21	$y{=}0$

Table 7: Uninformed player's contribution frequency, expected payoffs and best reply, by treatment

In the treatments where signals are received, the uninformed player responds optimally to signals in most cases. In Actions, after observing a contribution by the informed player, the uninformed player contributes 88% of the time. This is the choice that yields the highest expected payoff (131.65>127.77), and thus it is also the empirical best reply. In Words(s), after a message 'the state is 1.5', the uninformed player contributes 69.7% of the time, which again is also his best reply.

In Words(x) and for matching groups 13, 15 and 16, the uninformed player contributes 62.3% of the time after message 'I contribute', which is also his best reply. Interestingly, for matching group 14, the uninformed player rarely contributes after a message 'I contribute' (only 13.8%). This is his empirical best reply, as can be seen by comparing 71.03 to 66.21. This is mainly driven by the informed player's use of message 'I contribute' when the state is 0 in 38.5% of the cases (as shown in Table 5).

Uninformed player contributions in Actions are very similar to those in the treatments Words(s) and Words(x). If we compare the reaction to a contribution of the informed player in Actions to the reaction to the message 'the state is 1.5', we find that these are not significantly different (MW test, p-value=0.1489). If we compare that reaction to a contribution (88%) to the reaction to the message 'I contribute' (62.3%), we find that the difference is only marginally significant (MW test, p-value=0.0771). Finally, comparing the reaction to the message 'the state is 1.5' to the message 'I contribute', we find no significant differences (MW test, p-value=0.7237). This leads to Result 3.

Result 3 (contributions of the uninformed player):

The uninformed player frequently contributes (more than 60% of the time) after observing the contribution of the informed player, or after hearing the message 'the state is 1.5', or after the message 'I contribute'. Furthermore, the reaction to 'the state is 1.5' is not significantly different from the reaction after observing a contribution, while the reaction to the message 'I contribute' is only marginally different from that after observing a contribution (except for one matching group). Thus, the messages 'the state is 1.5' and 'I contribute' are as influential as actions, and we do not reject Hypothesis 3.

4.4 Payoffs and Efficiency

In Table 8 below we display average payoffs and efficiency by treatment. We also display the predicted average payoffs and efficiency in equilibrium.

Table 8 reveals that the informed player does remarkably well in the baseline treatment, compared to the theoretical prediction. This is due to the fact that the uninformed player contributes more frequently than predicted. In contrast, under Actions, Words(s) and (x), the informed player does worse as predicted, while the uninformed player's payoff comes close to the theoretical prediction in most cases. Interestingly, the uninformed player's payoff is significantly higher in matching groups 13, 15 and 16 in Words(x) compared to Words(s), while the informed player's payoff suffers a slight (non-significant) decrease (MW test, p-value=0.0339 and 0.4795, respectively). These changes reveal that the decrease in free-riding by the informed player in Words(x) has important effects, particularly for the uninformed player.

Taking both the informed and uninformed player's payoff, we can calculate efficiency. Table 8 shows that efficiency is highest in Actions (89.1%), and that it is significantly higher there than in Words(s) and Words(x), where it is 76.1% and 78.6% respectively (MW test, comparing Actions and Words(s), p-value=0.0209, comparing Actions and Words(x) in matching groups 13, 15 and 16, p-value=0.0497). Thus, we find that, as predicted, Actions leads to the most efficient outcome. If we compare efficiency between Words(s) and Words(x), we do not find a significant difference (MW test, p-value=0.4795).

	Informed player's average payoff		Uninformed player's average payoff		Efficiency	
Treatment	Observed	Predicted	Observed	Predicted	Observed	Predicted
Baseline	73.24	46.36	78.01	78.18	72.8%	61.0%
	(1.97)		(2.25)		(0.02)	
Actions	89.72	103.86	95.40	103.86	89.1%	100.0%
	(2.74)		(3.30)		(0.02)	
Words(s)	83.30	107.73	74.83	80.68	76.1%	91.9%
	(11.93)		(4.20)		(0.06)	
Words(x)						
a) *Matching groups 13,15,16*	76.06	107.73	87.12	80.68	78.6%	91.9%
	(14.45)		(3.29)		(0.05)	
b) *Matching group 14*	51.36	107.73	63.18	80.68	55.1%	91.9%

Note: standard deviations in parentheses.

Table 8: Average Payoffs and Efficiency, by treatment

Result 4 (efficiency):

(a) Efficiency is highest under Actions, as predicted. We therefore do not reject Hypothesis 4 (a).

(b) Efficiency is not significantly different in Words(s) and Words(x). We therefore do not reject Hypothesis 4 (b).

4.5 Discussion: messages and contributions by the informed player

All in all, the theoretical predictions from Section 2 organize the data very well. As we, however, have seen at the beginning of this section, hypothesis 1 (b) is rejected: when the informed player talks about her contribution, she contributes more often than when she talks about the returns to the contribution. Our objective here is to discuss this result in somewhat greater detail.

We display in Table 9, in the rows labeled Contribution Freq, the contribution frequencies by the informed player, conditional on the state and the message that she sends. For completeness, this table also displays, in the rows labeled Message Freq, the frequency with which each message is used. This latter information was already been displayed in Table 5.

			State		
Treatment	Message (m)		s=0	s=0.75	s=1.5
Words(s)	'The state is 0'	Contribution Freq[a].	0.0%	16.7%	100.0%
		Message Freq[b].	71.1%	8.8%	1.8%
	'The state is 0.75'	Contribution Freq.	0.0%	6.7%	50.0%
		Message Freq.	11.6%	16.2%	3.6%
	'The state is 1.5'	Contribution Freq.	0.0%	**4.1%**	90.7%
		Message Freq.	17.3%	**75.0%**	94.7%
Words(x)					
a)	*Matching groups 13,15 and 16*				
	'I do not contribute'	Contribution Freq.	2.8%	23.3%	100.0%
		Message Freq.	94.9%	23.5%	9.5%
	'I contribute'	Contribution Freq.	50.0%	**41.2%**	100.0%
		Message Freq.	5.1%	**76.5%**	90.5%
b)	*Matching group 14*				
	'I do not contribute'	Contribution Freq.	0.0%	33.3%	25.0%
		Message Freq.	61.5%	17.6%	28.6%
	'I contribute'	Contribution Freq.	0.0%	7.1%	70.0%
		Message Freq.	38.5%	82.4%	71.4%

Note:[a] Number of times the informed player contributes and sends m over total number of times m is sent, by state;[b] Number of times m is sent over total number of times that s is drawn.

Table 9: Contribution frequency by the informed player, conditional on the message sent, and message use

Let us focus on the case $s = 0.75$ and the focal equilibrium. In Words(s), the informed player sends the message 'the state is 1.5' in 75% of the cases, and, in such case, she very rarely contributes (only in 4.1% of the cases), as shown in bold. In particular, the informed player lies frequently. Let us contrast this with the behavior in the matching groups 13, 15 and 16 in the Words(x) treatment. First of all, when $s = 0.75$, the informed player frequently states that she contributes (76.5%). However, conditional on sending the message 'I contribute', she indeed contributes in 41.2% of the cases. Hence, when s=0.75, the informed player contributes more often conditional on saying 'I contribute' as compared to when saying 'the state is 1.5' (MW test, p-value=0.0745). In contrast, conditional on sending the message 'I do not contribute' or 'the state is 0', contributions are not significantly different (MW test, p-value=0.6374).

This difference in behavior across messages 'I contribute' and 'the state is 1.5' is not driven by differences in the informativeness of the messages, as we saw in section 4.2, or in the reactions of the uninformed player, as we saw in section 4.3.

We suggest two explanations for this result, both relying on the idea that players may dislike lying. Existing research has shown that often individuals indeed have an aversion to lying about private information (e.g., Gneezy, 2005) or about intended actions (e.g., Ellingsen and Johannesson, 2004), and that the extent of lying may depend on the costs and benefits involved.

Let us first assume that players dislike lying as such. Formally, assume that the informed player's utility not only depends on her own material payoff but that she also suffers a disutility of c, when sending a message which is not true. Kartik (2009) follows this approach, which we simplify greatly here[19]. We will argue that it is less costly to avoid lying when the informed player talks about her contribution than when she talks about the state. Again, suppose s=0.75 and that we are in the focal equilibrium. When words are about the state and the informed player says 'the state is 1.5', her utility is u_I(x=0,'the state is 1.5',y=1)=100-c. In contrast, if she deviates and tells the truth about the state, she can expect the uninformed player not to contribute, hence, her utility will only be 40. Consequently, a lie brings considerable benefits. Only when the cost of lying is high, if $c \geq 60$, will the informed player say 'the state is 0.75'.

Now consider the situation in which the informed player talks about her contribution. As in the previous case, if she says 'I contribute' but does not, her utility is 100-c. If, instead, she says 'I do not contribute', her payoff drops to 40. However, in contrast to the previous case, the informed player can protect herself against this drop in payoff by saying

[19] A similar approach is used in Kartik et al (2007), Chen et al (2008) and Chen (2009).

‘I contribute’ and choosing to indeed contribute. In this case, her payoff drops to 90, but she avoids the lie. For players that dislike lying somewhat, but not too much ($10<c<60$), this combination is the preferred one. In other words, players who dislike lying somewhat, but not too much, will choose to contribute in state s=0.75 and announce to do so, while they will choose to report ‘s=1.5’ in that state and to not contribute. Note that, if the informed player talks about the state, there is no cheap way to avoid the lie: even if she would contribute, she would still lie. This may explain why the level of free-riding depends on the language available.

The second explanation is based on the assumption that the informed player may have a taste for keeping her word. Ellingsen and Johannesson (2004) and Miettinen (2008) proposed models in which players suffer a disutility if they do not act as they announced or promised to do, and Vanberg (2008) provided evidence that people have a preference for keeping promises per se[20]. Now, in our game, there are no explicit promises, but saying 'I contribute' is somewhat similar to making a promise, while, in contrast, saying 'the state is 1.5' clearly is not. If individuals dislike breaking promises, they might be willing to forgo monetary payoffs in order to avoid breaking a promise, but not when talking about the state. To a certain extent, this explanation thus relies on the assumption that lying about intentions is perceived as being more costly than lying about a more neutral aspect, such as the state of nature.

The reader might wonder whether also guilt aversion (Batigalli and Dufwenberg, 2007) could not explain the difference in behavior across the two languages. According to this theory, an individual suffers a disutility when she lets another player down. In order to avoid this disutility or guilt, an individual might act according to what he believes others expect him to do (see Charness and Dufwenberg, 2006, Vanberg, 2008, and Ellingsen et al, 2009, for experimental tests of this theory, with the latter two papers arguing that guilt aversion may be less prominent than previously thought). Thus, if the informed player makes a promise and others expect her to keep it, she might keep it to avoid guilt. This theory, however, does not predict a priori that messages 'I contribute' and 'the state is

[20] Interestingly, existing studies on games with complete information show mixed results when communication about intentions is restricted, as in our case, to pre-formulated messages. Such restricted communication does not increase cooperation or trust in some studies (e.g. Bochet et al, 2006, and Charness and Dufwenberg, 2010), while it does in others (e.g. Duffy and Feltovich, 2002). See Balliet (2010) for meta-analysis, as well as the reviews by Bicchieri and Lev-On (2007) and Koukoumelis et al (2009), and the references therein.

1.5' generate different beliefs regarding what the uninformed player expects, while also the realized equilibrium payoff is the same in both cases (and equal to 30 in case of earnings table 2), so that the extent of letting the other down also is the same. It follows that guilt aversion does not imply a different behavior of the informed player across languages.

5 Conclusion

In the context of a two-player, one-shot, public good game in which only one player is informed about the return from contributing, we have compared signaling by words and actions. Using actions, the informed player reveals her contribution decision to the uninformed before the latter decides on his contribution. Using words, the informed player sends (cheap talk) messages, either about the return or about her contribution decision, before the other player decides on his contribution. We compare these signaling devices using also a baseline game in which no signaling is available.

From a theoretical perspective, by using actions, fully efficient contribution levels can be achieved. In the experiment, we find that contribution levels are indeed most efficient using this kind of signal. This result is in line with that of Potters et al (2007).

According to standard theory, whether messages are about the return, or about the contribution, is irrelevant. By allowing cheap talk, two Nash equilibrium outcomes become possible, but only one of these is neologism-proof. In this equilibrium, by using the appropriate words, the informed player can elicit the uninformed player's contribution. Consequently, words can be as influential as actions. However, 'cheap talk' has a 'dark side': it allows the informed player to free-ride on the contribution of the uninformed one.

Our experiment shows that words can indeed be as influential as actions. In most matching groups, messages are informative, as much as contribution decisions. And uninformed players react to the messages 'the state is 1.5' or 'I contribute' in a similar way as to a contribution decision. Broadly, the messages used in the experiment are also in line with what (our slight refinement of) neologism-proofness predicts.

In sharp contrast to the standard theoretical prediction, however, we find that it matters whether messages are about the return to the public good, or about the contribution of the informed player. Informed players contribute more often when saying 'I contribute', than when saying 'the return is 1.5'. Two possible explanations for this 'anomaly' were advanced in this paper: aversion to lying (coupled with the fact that it is less costly to avoid lies about contribution levels) and an intrinsic desire to keep one's word.

It is not straightforward to come up with a design that could separate the two explanations. The key is to find messages that affect whether the decision to contribute or not would turn the focal meaning of the message into a lie, while not involving an explicit promise. An admittedly somewhat contrived example would be a treatment with the following two messages: "if you contribute your payoff will be the same as my payoff" and "if you do not contribute your payoff will be the same as my payoff". In equilibrium, the informed player could send the latter message when the state is low, and the former message when the state is intermediate or high. In both the intermediate and the high state, the informed player can only prevent the message "if you contribute your payoff will be the same as my payoff" from being a lie by contributing herself since this message will induce the uninformed player to contribute. At the same time, the message is not an explicit promise about the informed player's action. So, a treatment with these two messages would separate the cost of lying argument from the argument that people want to keep their promises.

Irrespective of the outcomes of such a treatment, though, the present paper shows that it is interesting and important to pay attention to the focal meaning of messages. This meaning affects the non-material, psychological costs of different combinations of messages and actions.

Appendix: Proofs

Proposition 1 *The baseline game has a unique Nash Equilibrium, given by* $(\sigma^*, \tau^*) = \{(0, 0, 1); 0\}$.

Proof. Since $\frac{a+b+c}{3} < 1$, it is a strictly dominant strategy for U to choose $x_U = 0$. Since a, $b < 1$, $x_I = 0$ is a strictly dominant action for I, when $s = a$ or $s = b$. On the contrary, since $c > 1$, when $s = c$, it is a strictly dominant strategy for I to choose $x_I = 1$.

Proposition 2 *The game with Actions has a unique Nash Equilibrium,* $(\sigma^*, \tau^*) = \{(0, 1, 1); (0, 1)\}$.

Proof.- As in Serra–Garcia, van Damme, Potters (2008)- We will prove the stronger result that strategy profile X^* is the only one that survives iterated elimination of strictly dominated strategies.

Since $a \leq 0$ I has $x_I^* = 0$ as a strictly dominant action for $s = a$. From $a + b + c < 3$, it follows that U will respond to $x_I = 0$ by not contributing either: seeing $x_I = 0$ makes him less optimistic that the state is intermediate or good. This in turn implies that I has $x_{Is} = 1$ as her dominant action when $s = c$. Since $b + c > 2$, this in turn implies that U will contribute after a contribution of I. Having established that, for U, only $\tau^* = (0, 1)$ survives the elimination of dominated strategies, it easily follows that $x_{Ib} = 1$, hence, that $\sigma^* = (0, 1, 1)$ is the unique surviving strategy for I

Proposition 3 *There are two pure strategy equilibrium outcomes in the game with Words, given by, respectively:*

(1) $X(\sigma) = (X_a(\sigma), X_b(\sigma), X_c(\sigma)) = (0, 0, 1)$ and
$\tau(m) = 0$ for all $m \in M_s(\sigma)$, where $s = \{a, b, c\}$

(2) $X(\sigma) = (X_a(\sigma), X_b(\sigma), X_c(\sigma)) = (0, 0, 1)$ and
$\tau(m) = 0$ for all $m \in M_a(\sigma)$,while $\tau(m) = 1$ for all $m \in M_b(\sigma) \cup M_c(\sigma)$

Proof. First of all, note that player i strictly prefers player j ($j \neq i$) to contribute when $s>0$, while she strictly prefers the other not to contribute when $s<0$. Write D_i for the difference in (expected) payoff for player i between contributing (x_i=1) and not (x_i=0). It is easily seen that $D_i = E(s) - 1$. It immediately follows that the informed player will not contribute when s=a,b and will contribute when s=c.

For now, assume $a<0$. If player U follows a constant strategy ($y(m) = y^*$ for all $m \in M$), then equilibrium requires $y^* = 0$, as there is at least one message where $y^* = 0$ is a best response. There can be many equilibria of this type, and there are at least $|M|$ pure equilibria of this type, one for each $m \in M$. All these equilibria are uninformative; talk is considered pure cheap talk.

Next, assume that player U's strategy is not constant. Let $\overline{m}$ be a message with the highest probability that player U contributes, while $\underline{m}$ denotes one with the lowest. If these messages are unique, then types b and c will choose $\overline{m}$, while type a will choose $\underline{m}$. Equilibrium requires that $y(\overline{m}) = 1$ and $y(\underline{m}) = 0$, and this is indeed an equilibrium. We see that there are multiple pure semi-separating equilibria, but that these all yield the same outcome. Of course, $\overline{m}$ and $\underline{m}$ need not be unique. Non-uniqueness of $\underline{m}$ does not create specific problems. Suppose $\overline{m}$ is not unique. As type a will not choose any such $\overline{m}$, there must exist at least one $\overline{m}$ where player U attaches beliefs of at least $\frac{1}{2}$ to facing type c and, hence chooses $y(\overline{m}) = 1$. In the equilibrium, only such $\overline{m}$ will be chosen by both b and c. These types can differ a bit in their strategies, but not too much. This still generates the same pure semi-separating outcome. Consequently, if a<0, there are only two equilibrium outcomes: one in which player U never contributes, and another one in which he contributes for sure after some messages, there is no contribution after other messages, and there is randomization after a third set of messages. In this second equilibrium, types b and c randomize among messages in the first set, a randomizes among messages in the second set, and messages in the third set appear with zero probability.

Let us finally consider $a = 0$. It will be clear from the above argument that, if we restrict ourselves to pure strategies, (only) the two equilibrium outcomes exist that were identified above. If $a = 0$, however, type a is indifferent between what U does, hence, he could randomize between $\overline{m}$ and $\underline{m}$. If he randomizes, the result can be such that $D_U = E(s|m) = 0$, so that U can randomize as well. This then gives rise to various mixed equilibria. We do not specify further details here, as these can be filled in by the reader.

Proposition 4 *Only an influential equilibrium is neologism-proof.*

Proof. A pooling equilibrium is not neologism-proof as the set $T = \{b, c\}$ is a self-signaling set, relative to this equilibrium. A similar remark applies for any mixed strategy equilibrium in which types b and c do not receive their best payoff. On the other hand, a partially separating equilibrium is trivially neologism-proof, as, in this case, the informed player receives the best possible payoff in each state of nature. (Formally: $u_I(s) = \max_{x,y} u_I(s, x, y)$for all s, so that there cannot be a self-signaling set with respect to this equilibrium.)

References

[1] Agranov, Marina and Andrew Schotter (2009). "An Experimental Study of Ambiguity and Vagueness in the Announcement Game." Working paper, New York University.

[2] Balliet, Daniel (2010). "Communication and cooperation in social dilemmas: A meta-analytic review." *Journal of Conflict Resolution,* 59, 34-57.

[3] Batigalli, Pierpaolo, and Martin Dufwenberg (2007). "Guilt in Games." *AEA Papers and Proceedings*, 97, 170-176.

[4] Bicchieri, Christina, and Azi Lev-On (2007). "Computer-mediated communication and cooperation in social dilemmas: An experimental analysis." *Politics, Philosophy, and Economics,* 6,139-168.

[5] Blume, Andreas, DeJong, Douglas V., Kim, Yong-Gwan, and Geoffrey B. Sprinkle (1998). "Experimental Evidence on the Evolution of the Meaning of Messages in Sender-Receiver Games." *American Economic Review,* 88, 1323–1340.

[6] Blume, Andreas, DeJong, Douglas V., Kim, Yong-Gwan, and Geoffrey B. Sprinkle (2001). "Evolution of Communication with Partial Common Interest." *Games and Economic Behavior,* 37, 79-101.

[7] Bochet, Olivier, Page, Talbot, and Louis Putterman (2006). "Communication and punishment in voluntary contribution experiments." *Journal of Economic Behavior & Organization,* 60, 11–26.

[8] Bracht, Juergen and Nick Feltovich (2009). "Whatever you say, your reputation preceeds you: Observation and cheap talk in the trust game." *Journal of Public Economics,* 93, 1036-1044.

[9] Brandts, Jordi and David Cooper (2007). "It's What You Say Not What You Pay: An Experimental Study of Manager-Employee Relationships in Overcoming Coordination Failure". *Journal of the European Economic Association*, 5, 1223-1268.

[10] Cai, Hongbin and Joseph Tao-Yi Wang (2006). "Overcommunication in strategic information transmission games." *Games and Economic Behavior,* 56, 7-36.

[11] Charness, Gary and Martin Dufwenberg (2006). "Promises and Partnership." *Econometrica,* 74, 1579-1601.

[12] Charness, Gary and Martin Dufwenberg (2010). "Bare Promises: An Experiment" *Economics Letters*, forthcoming.

[13] Chen, Ying (2009). "Perturbed Communication Games with Honest Senders and Naive Receivers." Working paper, Arizona State University.

[14] Chen, Ying, Kartik, Navin, and Joel Sobel (2008). "Selecting Cheap-Talk Equilibria". *Econometrica*, 76, 117–136.

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

[16] Crawford, Vincent (2003). "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review* 93, 133-149.

[17] Çelen, Boğaçhan, Kariv, Shachar, and Andrew Schotter (2009). "An Experimental Test of Advice and Social Learning." Working paper, New York University.

[18] Demichelis, Stefano and Jörgen Weibull (2008). "Language, Meaning, and Games: A Model of Communication, Coordination, and Evolution." *American Economic Review,* 98, 1292–1311.

[19] Duffy, John and Nick Feltovich (2002). "Do Actions Speak Louder Than Words? An Experimental Comparison of Observation and Cheap Talk." *Games and Economic Behavior,* 39, 1-27.

[20] Duffy, John and Nick Feltovich (2006). "Words, Deeds and Lies: Strategic Behaviour in Games wtih Multiple Signals". *Review of Economic Studies,* 73, 669-688.

[21] Ellingsen, Tore and Magnus Johannesson (2004). "Promises, Threats and Fairness." *Economic Journal*, 114, 397–420

[22] Ellingsen, Tore, Johannesson, Magnus, Tjotta, Sigve, and Gaute Torsvik (2009). "Testing Guilt Aversion." *Games and Economic Behavior, 68*, 95-107.

[23] Ellingsen, Tore, and Robert Östling (2009). "When Does Communication Improve Coordination?" *American Economic Review*, forthcoming.

[24] Farrell, Joseph (1985). "Credible Neologisms in Games of Communication." MIT Working Paper 386.

[25] Farrell, Joseph (1993). "Meaning and Credibility in Cheap-Talk Games." *Games and Economic Behavior,* 5, 514-531.

[26] Farrell, Joseph and Matthew Rabin (1996). "Cheap Talk." *Journal of Economic Perspectives* 10, 103-118.

[27] Fischbacher, Urs (2007). "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics,* 10, 171-178.

[28] Gneezy, Uri (2005). "Deception: The role of consequences." *American Economic Review,* 95, 384–294.

[29] Güth, Werner, Levati, M. Vittoria, Sutter, Matthias and Eline van der Heijden (2007). "Leading-by-example with and without exclusion power in voluntary contribution experiments." *Journal of Public Economics,* 91, 1023-1042.

[30] Hermalin, Benjamin (1998). "Toward an economic theory of leadership: Leading-by-example." *American Economic Review,* 88, 1188-1206.

[31] Hurkens, Sjaak, and Navin Kartik (2009). "Would I lie to you? On social preferences and lying aversion." *Experimental Economics,* 12, 180-192.

[32] Kahneman, Daniel and Amos Tversky (1974). "Judgment under Uncertainty: Heuristics and Biases." *Science,* 185, 1124-1131.

[33] Kartik, Navin (2009). "Strategic Communication with Lying Costs." *Review of Economic Studies*, forthcoming.

[34] Kartik, Navin, Ottaviani, Marco, and Francesco Squintani (2007). "Credulity, Lies, and Costly Talk." *Journal of Economic Theory*, 134, 93–116.

[35] Knowles, Elizabeth (2006). "Actions speak louder than words." *The Oxford Dictionary of Phrase and Fable.* Encyclopedia.com (March, 2010).

[36] Knowles, Elizabeth (2006). "Englishman." *The Oxford Dictionary of Phrase and Fable.* Encyclopedia.com (March, 2010).

[37] Koukoumelis, Anastasios, Levati, M. Vittoria, and Johannes Weisser (2009). "Leading by Words: A Voluntary Contribution Experiment With One-Way Communication." Jena Economic Research Papers 106.

[38] Miettinen, Topi (2008). "Contracts and Promises - An Approach to Pre-play Agreements." SSE/EFI Working Paper Series in Economics and Finance No 707.

[39] Moxnes, Erling, and Eline van der Heijden (2003). "The effect of leadership in a public bad experiment." *Journal of Conflict Resolution,* 47, 773-795.

[40] Potters, Jan, Sefton, Martin and Lise Vesterlund (2007). "Leading-by-example and signaling in voluntary contribution games: An experimental study." *Economic Theory,* 33, 169-182.

[41] Rabin, Matthew (1990). "Communication between rational agents." *Journal of Economic Theory*, 51, 144-170.

[42] Sánchez-Pagés, Santiago and Marc Vorsatz (2007). "An experimental study of truth-telling in sender-receiver games." *Games and Economic Behavior,* 61, 86-112.

[43] Serra-Garcia, Marta, van Damme, Eric and Jan Potters (2008). "Truth or Efficiency? Communication in a Sequential Public Good Game." TILEC Discussion Paper No. 2008-047; CentER Discussion Paper Series No. 2008-107.

[44] Vanberg, Christoph (2008). "Why do people keep their promises? An experiment test of two explanations." *Econometrica,* 76, 1467–1480.

[45] Vesterlund, Lise (2003). "The Informational Value of Sequential Fundraising." *Journal of Public Economics,* 87, 627–657.

[46] Wilson, Rick K. and Jane Sell (1997). "Liar, Liar...: Cheap Talk and Reputation in Repeated Public Goods Settings." *Journal of Conflict Resolution,* 41, 695-717.