Strategic information transmission and social preferences

Antonio Cabrales, Francesco Feri, Piero Gottardi
and Miguel A. Meléndez-Jiménez*

May 2025

Abstract: This paper reports on experiments regarding cheap talk games where senders engage in deception also when their interests are not in conflict with those of the receiver. The amount of miscommunication we observe is higher than in previous experimental findings on cheap talk games, even though, as in previous work, some participants appear to feature a cost of lying. A novel feature of our framework is that sometimes senders' and receivers' interests are in conflict and some other times they are aligned. We show that our findings can be attributed to distributional preferences of senders, which may be sufficiently high to induce them to lie, even when they face a cost of lying, to avoid the receiver getting a higher payoff than the sender.

Keywords: Experiments, Cheap talk, Deception, Conflicts of interest, Social preferences

JEL Classification: D83, C72, G14

---

* Antonio Cabrales, Dept. of Economics, Universidad Carlos III de Madrid; Francesco Feri, Dept. of Economics, Royal Holloway University of London; Piero Gottardi, Dept. of Economics, University of Essex and Venice; Miguel A. Meléndez-Jiménez, Dept. Teoría e Historia Económica, Universidad de Málaga. We are grateful to Valeria Maggian for very useful comments and suggestions. Miguel A. Meléndez-Jiménez and Antonio Cabrales acknowledge financial support from the Agencia Estatal de Investigación through the María de Maeztu Programme CEX2021-001181-M and PID2021-126892NB-100 as well as RTI2018-097620-B-I00 and PID2021-127736NB-I00, and from the Regional Government of Andalusia-FEDER through Projects UMA18-FEDERJA-243 and P18-FR-3840. The research reported in this paper is declared to the Research Ethics Committee of Royal Holloway University of London. All errors are our own.

1. Introduction

Since the seminal work of Crawford and Sobel (1982), we learnt that under standard preferences there are sharp limits to the amount of information that can be credibly transmitted among parties. Those limits arise from the extent of the alignment of preferences between the players. Yet a recent literature, prompted by experimental findings, has argued that individuals may have moral concerns, like an aversion to lying, in which case the limits to information transmission are relaxed.[1]

In this paper we show that other factors, going in a different direction, are important in explaining behavior in situations of strategic information transmission. To be precise, we find that distributional concerns can create impediments to information communication which overcome the effects of the aversion to lying found by this] literature. We do this in a fully strategic context, where the receivers can suspect the motives of the senders.[2]

The typical situation we examine is one where the sender, by telling the truth and being believed, receives a benefit (or is not harmed) but the receiver receives an even bigger benefit. In that case, concerns for equality might tempt the sender to mislead the receiver.[3] Think of a corporate raider networking in a social occasion. She can pass on a piece of information to a colleague about an investment opportunity. However, the "sender" could herself be interested in the investment, or could not be interested, and this influences the receiver's beliefs. Something that complicates the situation is that relative status concerns may also be in operation. If the sender is competing for promotion with the receiver, she may not want to transmit the information honestly. This is a kind of realistic dilemma which firms and individuals encounter in reality and may hamper effective information transmission.[4]

[1] See, e.g. Kartik et al. 2007, Kartik 2009, Gneezy, Kajackaite, and Sobel, 2018, Abeler, Nosenzo and Raymond, 2019.

[2] This is a key difference between our approach and the related one of Gneezy (2005) as we discuss in more detail in the literature review section.

[3] Gneezy (2005) and Erat and Gneezy (2012) also explore the interaction of social preferences with a norm of truth-telling. In the next section we discuss how our research differs from this work.

[4] We believe this phenomenon happens in many relevant applications, where these impediments to information transmission could have significant implications. For example, in Cabrales et al. (2020) we experimentally tested the viability of markets for cheap-talk information. We found that sellers provide low-quality information even when doing so did not increase their monetary payoff, which results in a low level of trade. Our new results provide a rationale for this observation in terms of social preferences.

The kind of situation we just described cannot happen in the cheap talk games that follow Crawford and Sobel (1982). In that framework, when preferences are aligned players have the same utility for the possible actions which can be chosen by the receiver and there is thus no room for distributional concerns and no reason for not telling the truth. When preferences are misaligned, on the other hand, distributional concerns may arise but there is also a reason for lying independently of whether such concerns are present or not.

We study a sender-receiver game that features the essential ingredients of our motivating story. There is a state of nature known to the sender, but not to the receiver. The sender sends a message about the state to the receiver, who then chooses an action. In addition, we assume that the sender has a private type, known only to herself, while the receiver's type is common knowledge. These types and the state of nature determine whether a conflict of interest is present. This introduces a reason for the receiver to be cautious about the sender. The sender should anticipate that and be cautious, which adds richness to the strategic aspects of our environment.

When the sender's type is equal to the state of nature she would like to gain the asset and her payoff is negatively affected if the receiver chooses to compete for it. In the complementary event, her payoff is invariant with respect to the receiver's action. But in that event, if the receiver chooses the action that is optimal for him in each state of nature, he gets a higher payoff than the sender. This is a crucial difference with earlier designs. There is no conflict of interest in that event, but there is a distributional asymmetry.

We first provide a complete characterization of all the pure strategy equilibria of this game, under standard preferences. There is an equilibrium in which the sender sends a truthful message, except in the event where interests are misaligned (both would like to gain the asset). As in all sender-receiver games, there is also a babbling equilibrium, where the sender's message is totally uninformative. Finally, there is another equilibrium in which the sender's message is only truthful when the types of sender and receiver differ.

The experimental results we obtain show that, as in the previous experimental literature, subjects acting as senders sometimes tell the truth when there is a conflict of interest between sender and receiver. The crucial novelty is that we also find a remarkable amount of deception when the sender's payoff is independent of the action of the receiver. To gain some explanation for these different findings, recall that by telling the truth when the sender's type differs from the state of nature, and being believed, the sender does not harm her payoff, but clearly benefits the receiver.

As mentioned earlier, a candidate explanation for the observed behavior is that participants display sufficient concerns for equity to overcome the tendency to tell the truth that we also observe in the experiment. To test whether these aspects (social preference and cost of lying) affect behavior, we elicited the distributional preferences of the participants in our experiments using the method of Bartling et al. (2009). We also elicited the cost of lying for a large fraction of the sample.

We show that a model accounting for social preferences, lying aversion and for heterogeneity in the role of these traits across individuals can explain around three quarters of the observed behavior by participants as the outcome of best-response behavior.

To better understand the behavior of experimental participants, we then estimate a mixture model (à la Costa-Gomes, Crawford and Broseta 2001). In such a model each sender is assumed to best respond, on the basis of some appropriately identified preferences, to the receivers' behavior in her group, and departures from it arise from random mistakes. Our results show that around 40% of subjects best respond in line with a model specification where both social preferences and lying aversion are taken into account, 30% of subjects best respond according to standard preferences, and 30% lie in the remaining categories. Hence, this analysis shows that both social preferences and lying aversion play a relevant role to explain behavior at the individual level.[5]

Finally, a key contribution in our paper is that, in contrast with earlier literature, we examine a setting with true strategic interaction between sender and receiver. To show the importance of this element, we compare the results of our main treatment with the ones we obtain in a setting where there is no conflict of interest between sender and receiver. We find significantly less deception in the latter one. Thus, the (realistic) awareness about potential misaligned interest by the two parties, which is inherent to many real-world situations, matters, and our study sheds light on the determinants of behavior in this kind of situation.

The paper is organized as follows. We first briefly describe the related literature. Then we present the design of the experiment and the analysis of the equilibria of the

[5] While a full characterization of equilibria with heterogeneous social preferences and lying aversion traits is difficult, our overall findings support the claim that individuals act according to equilibrium behavior.

game we consider. The experimental results and their discussion follow. We conclude with some reflections for future research.

2. Literature

First, we should mention the seminal work of Crawford and Sobel (1982) on strategic information transmission, which studies how the alignment of preferences between sender and receiver affects information transmission.[6] We consider a novel game structure where distributional preferences have distinct implications for agents' behavior from attitudes towards truth-telling.[7]

The experimental literature on information transmission has primarily analyzed sender-receiver games à la Crawford and Sobel (1982). A first series of papers (e.g., Dickhaut, McCabe and Mukherji (1995), Blume et al. (1998, 2001), and Kawagoe and Takizawa (1999)) demonstrates that, when the interests of the sender and the receiver are well aligned, play tends to converge to informative/separating equilibria. A more recent strand of the literature (see Sánchez-Pagés and Vorsatz (2007), Kawagoe and Takizawa (2005), Cai and Wang (2006), and Wang, Spezio and Camerer (2010)) finds evidence of a higher level of truthful communication than in the most informative equilibrium, and attributes this to a truth-telling norm.[8] We also find some evidence of aversion to lying, but also observe a substantial amount of deception/misinformation, not explained by a conflict of interest, for subjects who display (independently measured) non-pro-social or envious preferences.

Maggian and Villeval (2016) study the connection between lying and social preferences in a dictator game. They find that subjects playing the dictator often lie to avoid unequal distributions. However, the game considered is not one of communication, as the other participant is passive, and the lie is directed to the experimenter. This is an important difference in our view. The real-life settings we are modelling involve receivers of the message who understand the senders' incentives and can react in different ways. A similar design is studied by Barron, Stüber, and van Veldhuizen (2019) who again consider a "lying dictator game."

[6] Sobel (2013) reviews the vast theoretical literature following that paper.

[7] There is also a relevant theoretical literature that studies information transmission when agents may have a preference for telling the truth (for either moral reasons, as in Kartik, Ottaviani and Squintani (2007), or reputational ones, as in Bolton, Freixas and Shapiro (2007)).

[8] Some papers also argue that level-k thinking can explain overcommunication outcomes in their game (see, for instance, Cai and Wang (2006) and Lafky et al. (2022)).

Gneezy (2005) also explores deception by senders in situations where senders always benefit materially from the deception, while the receiver is harmed. There are a couple of key differences between Gneezy (2005) and our experiment. One is that we show that envy can be a factor leading to lying, whereas he shows that lying aversion can be a deterrent to pursue self-interest, more in line with the findings of the experimental literature on cheap talk games recalled above (something we also observe as some players tell the truth when they are interested in the object). Perhaps more importantly, another significant difference is that in our case receivers are aware of the potential conflict of interest, and this is common knowledge, while in Gneezy (2005) there is always a conflict of interest, but receivers are not aware of it. Hence in our set-up the sender should take into account how the receiver interprets her messages. Strategic analysis is then crucial, and connected to our motivating real-life situations, and we study it carefully.

Brandts and Charness (2003) study a situation where a player sends a cheap talk message about her intended action in a 2×2 game she plays with a receiver. The receiver is then allowed to inflict a costly punishment to the sender. Like them, we also find that both truth-telling and distributional concerns contribute to determine subjects' behavior, though the way in which these factors manifest themselves is different and so is the environment considered.

We should also mention some interesting recent work in the field, though a bit more distant from our motivation. Agranov, Dasgupta and Schotter (2020) study a cheap talk game where sellers have private information about the quality of the good they sell and send messages to the buyers regarding the quality. Fréchette, Lizzeri and Perego (2021) study the implications of commitment and verifiability in information transmission. Farina et al. (2024) study a model where the agents can selectively disclose information.

Since we make a connection between deception and social preferences, there is a literature relating those two preference traits. Gneezy et al. (2020) show that experimental participants can deceive themselves into thinking that a self-serving recommendation to another player is done for the other players' benefit. The mental mechanism exploited in these papers may also be used by players in our experiment to alleviate their psychological cost of lying. Kerschbamer et al. (2019) show that altruists lie less, but only when lies hurt others. Amato et al. (2020) show that children who are inequality averse are less likely to lie. This is not universal, though. Vranceanu and Dubart (2019) use a modification of Gneezy (2005) with a multiple price list to measure the internal cost of deception. They find that this cost of deception is uncorrelated with social preferences. Erat and Gneezy (2012) show that lying aversion is strong enough that people sometimes do not want to lie even if the lie would benefit both the sender and the receiver. As far as we can see, none of their treatments induce pure spiteful lies (that do not benefit the sender), as the ones we consider here.[9]

3. The game

We now describe the game. In each round, each player is randomly assigned a color (either black or white with equal probability), defining his type. The color of the object is also drawn from the same distribution and defines the state of nature. The assigned colors of players and the object are drawn independently. Player 1 – the sender – is informed of the colors assigned to both agents and of the color of the object whereas player 2 – the receiver – is only informed of his assigned color but knows that the sender is fully informed and thus is aware of the potential conflict of interest. If the assigned color of a player coincides with the color of the object, we say the player is *interested* (otherwise, *uninterested*). One can think of the color of a player as the type of investment in which an asset manager specializes, or the type of worker a human resource manager needs. The color of the object is the type of investment, or the type of worker, the sender (trader or HR manager) is informed of.

Next, player 1 sends a message (or report) regarding the color of the object to player 2. The message can be either "the color is white" or "the color is black", i.e., player 1 can either send a truthful message or a false message. Player 2 observes the message sent by player 1 and chooses an action that can be either *left*, *center*, or *right*. The action is the "bid" of the receiver (salary offer, purchase price) for the asset, with *left* being "high", *center* being "medium" and *right* being "low". The players' payoffs depend on whether they are interested in the object and on the choice of player 2, as indicated in Table 1.

The payoff structure resembles the conflicts of interest that would arise in a hypothetical competition for the object where, as mentioned before, player 1 (the

[9] Abeler et al. (2019) show that lying aversion is a very strong motivator for behavior using a theoretical model they then validate both with past experiments and some new ones specifically designed to test the theory. Jacobsen et al. (2018) is an insightful survey on the recent experimental literature on deception.

informed player) makes a high offer when she is interested in the object and a low offer otherwise. One can interpret Player 2's action *left* as a high offer in the same situation, action *center* as a medium offer, and action *right* as a low offer. When player 2 is interested, action *left* is preferred by this player to *center*, which in turn is better than *right*. If he is not interested, the order is reversed.[10]

Player 1's payoffs, when she is not interested, are then independent of the action of player 2. In contrast, when player 1 is interested, she prefers that player 2 chooses *right* rather than *center* and *center* rather than *left*. In this case the preferences of player 1 may be strictly aligned with those of player 2 (when player 2 is not interested), or strictly misaligned (when player 2 is also interested).

Table 1. Payoffs in the game *Sender Receiver*

	Player 2 not interested			*Player 2 interested*		
Player 2's choice →	Left	Center	Right	Left	Center	Right
Player 1 not interested	20, 20	20, 60	20, 100	20, 160	20, 120	20, 50
Player 1 Interested	20, 30	70, 90	120, 120	20, 60	70, 50	120, 40

In each cell: player 1's payoff, player 2's payoff

We want to stress the richness of our design. The sender has four information sets: in two of them (when she is interested) she has a strict interest in the receiver believing her messages, in the other two (when she is uninterested) she is indifferent. However, she could use her behavior in these last two information sets to shape the receiver's beliefs over the truthfulness of the messages she sends. We believe this characteristic is uniquely novel in our design.

These features characterize situations where each agent can act on the basis of his or her information, information has a partly rival nature, and its value varies with the state. These features are somewhat different from the ones present in the cheap talk games considered in most of the literature following Crawford and Sobel (1982), where the conflict of interest is known, independent of the state of nature.[11]

---

[10] Moreover, player 2's payoff is strictly higher when her assigned color differs from the assigned color of player 1 than when they coincide.

[11] For an exception, see Thordal-Le Quement (2016). One important difference with our setup is that he considers the case where the senders' bias is private information and is fixed across states of nature, whereas in our case the bias depends on the state. A more important difference is that there are many possible senders the receiver can consult (at a cost). Since some of the senders are unbiased, while others are biased in possibly different directions, the actions of senders are affected by competition among them and their

The analysis of the equilibria of the game described is developed in Appendix A, where we characterize all the pure strategy (weak) perfect Bayesian equilibria. There are three equilibria: the first one is the *informative equilibrium*, in which player 1 sends a message that informs the receiver of the true color of the object except in the event where she is interested in the object and player 2 is also interested. The optimal response of player 2 in this equilibrium is then to choose action *left* if the report says that he is interested (i.e., if it contains his assigned color) and action *right* otherwise. Thus player 2 trusts sufficiently the truthfulness of the message sent by player 1 to choose *left* (an "aggressive" action) when told he is interested and to choose *right* (an "accommodating" one) when told he is not.

The second equilibrium is a *babbling* one in which player 1's message is uninformative and player 2 then chooses action *center* regardless of the content of the message received.[12]

In the third equilibrium player 1's message reports truthfully the color of the object when the assigned colors of players 1 and 2 differ, but reports instead the opposite of the true color of the object when the assigned colors of players 1 and 2 coincide.[13] Player 2 chooses action *left* or action *center* if the report says that he is interested (both actions provide the same expected payoff to player 2) and action *right* otherwise. We refer to this as the *opposite equilibrium*.

The equilibrium (expected) payoffs to players 1 and 2 are, respectively, 70 and 105 (in the *informative* equilibrium), 45 and 80 (in the *babbling* equilibrium), and 70 and 85 (in the *opposite* equilibrium).

In real life, individuals often have interdependent preferences, in which envy and/or pro-sociality play an important role, and there are also social/moral norms that induce them to avoid lying, at least if it is not too costly. Moreover, we know, from both past and our experimental experience, that there is a lot of heterogeneity in those preferences/norms. So, a better equilibrium benchmark would be one where individuals have social preferences and aversion to lying. To illustrate these preferences, we consider

private information. The equilibrium structure is then much more complicated than ours and the research objective, focused on the disciplining effect of competition on truthfulness, is quite different.

[12] The case in which player 2 responds using action *left* on the out-of-equilibrium path (and action *center* on the equilibrium path) also constitutes a babbling (weak) perfect Bayesian equilibrium that is outcome equivalent to the babbling equilibrium described in the text. See Appendix A for details.

[13] Of course, associated with each one of the three equilibria there is another equilibrium, in which the colors reported by player 1 are exactly the opposite in every state. See Appendix A for details.

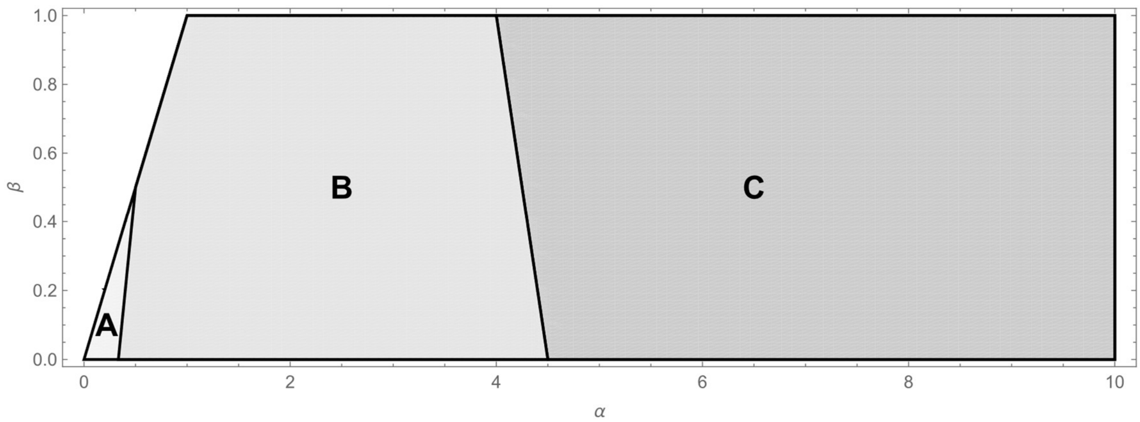
utilities that follow the Fehr and Schmidt (1999) specification, also including a loss of utility equal to $L_i \geq 0$ when the individual sends a false message. Thus, the utility of an individual i is:

$$U_i(\pi_i, \pi_j) = \pi_i - \alpha_i \cdot max\{\pi_j - \pi_i, 0\} - \beta_i \cdot max\{\pi_i - \pi_j, 0\} - I_i \cdot L_i \quad (1)$$

where $0 \leq \beta_i < 1$, $\beta_i \leq \alpha_i$, π_i and π_j are the payoffs of, respectively, individuals i and j (as given in Table 1), and I_i is an indicator function taking the value 1 when the individual sends a false message (0 otherwise).

In Appendix G we report the full characterization of equilibria with homogeneous preferences ($\alpha_i = \alpha$, $\beta_i = \beta$ and $L_i = L$ for all i). In Figure 1 we illustrate the ranges of (α, β) in which the informative, the babbling and the opposite equilibria exist without lying costs.[14] The babbling equilibria exists in all areas A, B and C, i.e., for all $\beta \leq \alpha$. The opposite equilibria exist in areas B and C, i.e., for $\beta \leq \min\{\alpha, 3\alpha - 1\}$. Finally, the informative equilibria exist in area C, i.e., for $\beta \geq 3 - 3\alpha$. All three equilibria exist in the baseline case $(\alpha, \beta) = (0,0)$

Figure 1. Informative, babbling and opposite equilibria without lying costs



Although Figure 1 suggests that the babbling and the opposite equilibria are more robust than the informative one, this is due to the absence of lying costs. Indeed, when $L > 0$, the babbling and the opposite equilibria no longer exist (except for the non-generic case $L = 40\alpha$), whereas the informative equilibrium exists for an open set of parameter values (which may be even larger than when $L = 0$). We report this set in Figure 2 in

[14] As reported in Appendix G, with social preferences there are various informative, babbling and opposite equilibria, which share the sender's strategy but differ in the receiver's response.

Appendix G for various values of L, illustrating the robustness of informative equilibria to the introduction of lying costs.

New equilibria also arise with lying costs, as described in Appendix G. For L sufficiently high, there is an equilibrium where senders follow a *full revelation* strategy. Also, when L takes intermediate values, and the parameters α and β are sufficiently high, the following also exist:

(i) Equilibria in which the senders always send the true message except for the case in which the sender is not interested, and the receiver is interested.

(ii) Equilibria in which the message of the sender only tells the truth either when she is interested, or when she is not interested.

4. Experimental design

Regarding the implementation of the experiment, participants were assigned to fixed groups of four players and played 40 rounds of the game.[15] In each round, subjects were randomly matched within their group, and in each pair one subject was randomly assigned the role of player 1 (sender) with the other subject acting as player 2 (receiver). Hence, players' roles were randomly assigned each round.[16]

After the 40 rounds of play, we elicited the subjects' attitudes towards risk and social preferences. We used the risk test proposed by Charness and Gneezy (2010). Subjects had to decide how much of their endowment (5 euros) to invest in a risky asset and how much to keep. They earned 2.5 times the amount invested if the asset has a high yield (which occurred with prob. 0.5) and otherwise lost the entire amount invested.

Regarding social preferences, we used the approach proposed by Bartling et al. (2009) to identify pro-social and envious attitudes. This information then allows us to investigate possible rationales for the behavior observed in the sender-receiver game in terms of these attitudes. Each subject was asked to make four decisions corresponding to

[15] The game is not simple, and the participants need time to learn how to play well. It is rather difficult to make the right inferences about the behavior of responders, for example, without playing enough periods. In many cases we refer to behavior in the last 20 periods, which we believe is a more reliable indicator of subjects' behavior. As we show in Table 6, the proportion of best responses (given the participants' elicited preferences) is above 75%, which suggests the general comprehension of the game was quite good.

[16] The fact that we allow for changing roles over time is a choice that, if anything, leads us to underestimate the importance of distributional concerns for subjects' behavior. In this way, players might disregard inequality thinking that one period they earn less, but the next period they will earn more. The small size of groups is then necessary to have enough independent observations.

four dictator games. Each decision consists of a choice between *distribution 1* and *distribution 2*. The choice of a distribution determines a payoff for the player and a payoff for another player.[17] These payoffs are shown in columns 2 and 3 of Table 2.

Using the choices made in these games, we can classify subjects with respect to their pro-sociality and envy attitudes. Regarding pro-sociality (games I and II), subjects choosing distribution 1 in game I and distribution 2 in game II are classified as *weakly pro-social* and those choosing distribution 1 in both games are classified as *strongly pro-social*. Those who choose instead distribution 2 in both games are classified as *non-pro-social*. Regarding envy (games III and IV), subjects choosing distribution 1 in game III and distribution 2 in game IV are classified as *weakly envious*, while those choosing distribution 1 in both games are classified as *strongly envious*. In contrast, those choosing distribution 2 in both games are classified as *non-envious*.[18] To explain the terms, a prosocial person prefers (perhaps weakly) to avoid having higher payoffs than others. An envious person prefers (perhaps weakly) to avoid having lower payoffs than others. It is indeed possible to be both pro-social and envious. The last two columns of Table 2 report the range of values of the parameters of the utility function in equation (1) that are compatible with the choices described above.

Table 2. Dictator games for the elicitation of social preferences

Game (All payoffs in euros)	*Distribution 1 self: other*	*Distribution 2 self: other*	α_i	β_i
(I) Pro-sociality	2: 2	2: 1	-	$[0, 0.5]$
(II) Costly pro-sociality	2: 2	3: 1	-	$[0.5, 1)$
(III) Envy	2: 2	2: 4	$[0, 0.5]$	-
(IV) Costly envy	2: 2	3: 5	$[0.5, \infty)$	-

Finally, for a large subset of our sample we also carried out, again at the end of the session, a lying aversion test, following Gibson et al. (2013). Subjects are asked to answer five questions, in which to lie provides a payoff of 1.50 Euro, whereas to say the truth provides a (weakly lower) payoff of *X*. The questions differ in the value of $X \in$

[17] Every subject acted as the decision maker in each of the four dictator games, with another subject randomly chosen to be the recipient, and one (randomly selected) choice of a decision maker in the four games was paid.
[18] Note that a subject choosing distribution 2 in game I and distribution 1 in game II would be hard to rationalize in terms of pro-social attitudes. Similarly, a subject choosing distribution 2 in game III and distribution 1 in game IV would be hard to rationalize in terms of envy attitudes. These subjects are classified as inconsistent.

$\{0.30, 0.60, 0.90, 1.20, 1.50\}$. This test allows us to identify, for each individual, a range of values for her lying cost L_i. This is identified as the largest difference $1.50 - X$ for which the subject chooses not to lie, converted into ECU.[19] The instructions of this test are reported in Appendix B.

4.1. Procedural details

We ran fourteen sessions at the laboratory of experimental economics of the University of Siena (LabSi). A total of 204 subjects participated in these sessions, recruited from the LabSi pool of human subjects, primarily consisting of undergraduate students from the University of Siena.[20] The experimental sessions were run in December 2014, December 2021 and June 2023.[21] The average duration of the sessions was 70 minutes (including the reading of instructions but excluding payment procedures). The experiment was computerized and conducted using the experimental software z-Tree (Fischbacher (2007)). Eight randomly selected rounds (out of 40) were paid (4 rounds selected from the rounds in which the participant was assigned the role of player 1 and 4 rounds selected from those in which she was assigned the role of player 2). The conversion rate was 40 ECUs = 1 euro. The average payment was around 20 euro.

At the end of each round, all the players are informed about what happened in that round: assigned role, color of the object, color assigned to each player in the group, the message of player 1, the action chosen by player 2 and the payoff of each player. However, participants were not assigned unique identifiers to minimize the impact of reputational concerns.

The experimental instructions, as well as the instructions of the risk, social preferences, and lying aversion elicitation (translated into English) are reported in Appendix B.

[19] A subject who lies for a value of $X = x_i$ and tells the truth when $X = x_j$, with $x_j < x_i$, is classified as inconsistent.

[20] This yields 51 independent observations (matching groups of 4 players). In our sample, when we compute the frequency of (senders') true messages by information set at the matching group level, we get standard deviations always lower than 0.22. Given this, if we fix a significance level of 0.05 and a power of 0.8, with our sample size we can identify effects of size 0.09 or higher in frequencies of true messages (one sample t-tests).

[21] The lying aversion test was carried out in the sessions run in 2021 and 2023, with a total of 164 subjects. In Appendix F we compare the choices made by subjects in 2014 and 2021-2023 and find there are very few significant differences, none in the last 20 periods.

5. Results

In Table 3 we report the behavior of senders. More specifically, we report the frequency of true messages sent, distinguishing the cases in which the senders and the receivers are, respectively, interested or not interested in the object. According to the informative equilibrium, the frequency should be 1 except for the cell in which both players 1 and 2 are interested, in which case it should be 0. We observe that the modal play coincides with the prediction of that equilibrium (the frequency of true messages is above 0.5 in the top cells and in the bottom-left cell of Table 3, and it is below 0.5 in the bottom-right cell),[22] but there are significant deviations. These deviations are particularly evident in the top cells, in which approximately one-third of the uninterested senders (who are expected to report the truth since they are not interested) lie, and in the bottom-right cell, in which approximately one-third of the interested senders (who are expected to lie since they are interested and so is the receiver) report the truth. These deviations from the informative equilibrium are significant (all the p-values are 0.000, two-tailed t-tests).

Table 3. Player 1's behavior. Frequency of true messages

	Player 2 not interested		*Player 2 interested*	
	All Rounds	Rounds 21-40	All Rounds	Rounds 21-40
Player 1 not interested	67.58 [1027]	68.69 [511]	63.56 [1084]	64.74 [536]
Player 1 interested	82.64 [985]	83.71 [491]	36.18 [984]	33.27 [502]

[number of observations]

In order to check whether distributional motives contribute to disinterested lying, we have performed a paired t-test in which we have compared the frequencies of senders' lies in the case in which they are disinterested to the frequency of "irrational" lies, i.e., lies when the sender is interested and the receiver is disinterested (in which case neither direct nor distributional preferences can induce lies). We find that these frequencies are significantly different (both for the last 20 rounds and for all rounds the p-values are 0.000, two-tailed paired t-tests), which is consistent with the explanation based on the role of distributional concerns.

[22] All the tests reported in the paper consider that each matching group of 4 players delivers one independent observation. The frequencies in Table 3 are significantly different from 50 (all the p-values are 0.000, two-tailed t-tests).

In Table 4 we present the frequencies of actions chosen by the receivers in response to the content of the message they received, both for the last 20 rounds and for all rounds. Again, we observe that the modal choices in each case corresponds to the prescriptions of the informative equilibrium: the choice of *left* when the message says that the receiver is interested (*Message – Int*), with a frequency slightly above 50%, and the choice of *right* when the message says that he is not interested (*Message – NoInt*), with a frequency slightly below 60%.[23] In both cases action *center* is chosen around one third of the times.

Table 4. Player 2's behavior. Frequencies of choices

	Left		Center		Right	
	All Rounds	*Rounds 21-40*	*All Rounds*	*Rounds 21-40*	*All Rounds*	*Rounds 21-40*
Message – NoInt	9.25 [234]	8.63 [111]	33.58 [850]	35.93 [462]	**57.17 [1447]**	**55.44 [713]**
Message – Int	**53.65 [831]**	**53.58 [404]**	34.60 [536]	35.94 [271]	11.75 [182]	10.48 [79]

[number of observations], the theoretical predictions of the informative equilibrium are in bold

In terms of payoffs, we find that the average payoff of senders and receivers is, respectively, 50.44 and 83.43 (51.30 and 84.07 in the second half of the experiment). We find that these payoffs are significantly lower (at the 1% level) than the expected payoffs in the informative equilibrium and higher than the expected payoffs in the babbling equilibrium. With respect to the opposite equilibrium, we find that only the average payoff of senders is significantly different, at the 1% level.[24]

In Table 14 in Appendix C we report the behavior at the matching group level. More specifically, we report the average behavior of the participants in each matching

[23] In order to statistically check whether the choice of *right* corresponds to the modal play in the information set "*Message – NoInt*" and that the choice of *left* is the modal play in the information set "*Message – Int*", we test whether the frequencies of these choices are significantly above an arbitrary benchmark (set at 45%), while the frequencies of the two alternative choices are significantly below it. We find that all the differences to the benchmark are statistically significant at the 1% level (two-tailed t-test). More precisely, when all rounds are considered, all the p-values are 0.000 and, when we focus on the last 20 rounds, the p-values of the comparisons of the frequencies of choices of *right*, *center* and *left* to the benchmark are, respectively, 0.005, 0.000 and 0.005 (0.000, 0.004 and 0.001) in the information set "*Message – NoInt*" ("*Message – Int*").

[24] All the p-values in the comparison to the informative equilibrium are 0.000 (two-tailed t-tests). The p-values in the comparison to the babbling equilibrium are always 0.000 for the sender, and 0.002 and 0.005 for the receiver (considering all rounds and the second half of the experiment, respectively). The p-values in the comparison to the opposite equilibrium are always 0.000 for the sender, and 0.142 and 0.504 for the receiver (considering all rounds and the second half of the experiment, respectively).

group when acting as senders (left panel), and when acting as receivers (right panel). It is important to note that some of the groups present a modal play that does not resemble that in any of the equilibria. This may be due to the fact that, in addition to heterogeneity, the presence of social preference and lying aversion traits may play a role. For instance, we see that some groups display full revelation (i.e., the senders' modal play is to send a truthful message in all information sets, which can be sustained as an equilibrium in the presence of sufficiently high lying costs).[25] We explore this possibility in more detail next.

5.1 Social preferences

We examine here to what extent social preferences and lying aversion are associated with the observed behavior in the game. In Table 5 we report the distribution of social preferences (panel A) and lying costs (panel B) in our pool of experimental subjects. We see that almost 58% of the subjects are classified as envious, 20% of them are classified as non-pro-social and 52% have high lying cost (>12).

Table 5. Distribution of social preferences and lying costs

A) *Social preferences*

	Strongly Envious $\alpha_i \in [0.5, \infty)$	*Weakly Envious* $\alpha_i \in [0,0.5]$	*Non envious* $\alpha_i = 0$	*Total*
Strongly pro-social $\beta_i \in [0.5,1)$	22 (43.14) [30.14]	3 (5.88) [7.5]	26 (50.98) [31.71]	51 (100) [26.15]
Weakly pro-social $\beta_i \in [0,0.5]$	28 (26.67) [38.36]	32 (30.48) [80]	45 (42.86) [54.88]	105 (100) [53.85]
Non pro-social $\beta_i = 0$	23 (58.97) [31.51]	5 (12.82) [12.5]	11 (28.21) [13.41]	39 (100) [20]
Total	73 (37.44) [100]	40 (20.51) [100]	82 (42.05) [100]	195 (100) [100]

B) *Lying costs*

$L_i = 0$	$L_i \in [0, 12]$	$L_i \in [12, 24]$	$L_i \in [24, 36]$	$L_i \in [36, 48]$	$L_i \in [48, \infty)$	*Total*
25 (17.48)	44 (30.77)	37 (25.87)	19 (13.29)	1 (0.70)	17 (11.89)	143 (100)

(% over total row), [% over total column]. Nine subjects on panel A and twenty one subjects on panel B are not considered because their choices were inconsistent (see footnotes 18 and 19).

[25] See, for instance, groups #8, #9, #17, #20, #32, #39 and #49.

We derive then the best response for each individual when acting as sender, according to the utility function specified in equation (1), with respect to the (average) behavior of receivers in her matching group (contingent on each of the two potential messages – see Table 14 in Appendix C). To this aim, we use the values of α_i (envy) and β_i (pro-sociality), measured by the social preference test, and the lying cost L_i, measured by the lying aversion test.[26] The fit of these best responses with the data, by information set, is reported in Table 6. More precisely, this table shows the proportion of choices in our data, in each information set, that are best responses to the actual average behavior of receivers in the same group, for the social preference and lying aversion parameters measured for each individual. The first column reports the results for the whole experiment, the second one when considering only the last 20 rounds.

Table 6. Frequency of senders' best response behavior considering social preferences and lying costs

Info. Set	All Rounds	Rounds 21 -40
P1 not int. & *P2 not int.*	0.735	0.749
P1 not int. & *P2 int.*	0.708	0.713
P1 int. & *P2 not int.*	0.862	0.866
P1 int. & *P2 int.*	0.727	0.751
Total	0.757	0.768

A visual inspection of Table 6 reveals that around three quarters of the observations are consistent with best response behavior, given the actual preferences of the individuals elicited from the tests. Such frequency slightly increases if we focus on the second half of the experiment. This is a first indicator that an equilibrium benchmark in which players exhibit both social preference and lying aversion traits provides a good approximation to the observed data.

To further explore the impact of social preferences and lying aversion on subjects' behavior, we perform a logit estimation of the probability that player 1 sends a true message. The explanatory variables we consider are the social preference and lying

[26] Note that, for each participant, the outcome of the social preference test provides an interval for the values of parameters α_i and β_i (see the last two columns of Table 2 and panel A of Table 5), and the lying aversion test provides an interval for the lying cost L_i (see panel B of Table 5). We assess choices as best responses if they are optimal for some values of the parameters in the intervals determined by the subject's answers to the tests.

aversion variables and dummies that determine whether the sender and the receiver are interested in the object. We specify a dummy for each information set (*Info1*, *Info2*, *Info3* and *Info4*).[27] Similarly, we define *Prosoc* (*Env*), a dummy that takes value 1 if player 1 is pro-social (envious), either weakly or strongly. Following Gibson et al (2013) we also include in the regression the variable *Lying_av*, that takes integer values ranging from 0 to 5, where 0 corresponds to no lying costs and 5 corresponds to the highest level of lying cost.[28] [29]*Round* (from 1 to 40) and *Risk*, which describes the choice made by the subject in the risk test, are also included as explanatory variables. In the regression the variables *Env* and *Prosoc* are estimated only in the information sets where, according to the utility function specified in (1), they play a role. Thus, *Env* is estimated on all information sets while *Prosoc* only in the info set where both players are interested.

In Table 7 we report the marginal effects of being envious, displaying pro-social attitudes and being lying averse on the probability of sending a truthful report for all the rounds (top panel) and for the final 20 rounds (bottom panel). The full estimation is reported in Table 15 in Appendix D. More precisely, on the left panel of Table 7 we present the marginal effect of envy (i.e., of *Env* = 1 vs. *Env* = 0), the marginal effect of pro-sociality is shown in the central panel (i.e., of *Prosoc* = 1 vs. *Prosoc* = 0) and the marginal effect of lying aversion is reported on the right panel.

The first conclusion we can draw from the inspection of the marginal effects in Table 7 is that social preference attitudes are associated with the departures we observed in Table 3 in the behavior of the sender from the one indicated by the informative equilibrium, departures which persist in the final rounds of the experiment. In particular, when the sender is not interested such equilibrium prescribes that senders report the truth but we saw this happened only 2/3 of the time. The results on the left-hand side of the table (i.e., regarding the marginal effect of *Envy*) show that envy reduces the probability of telling the truth in these two information sets and the effect is significant. More

[27] *Info1* takes value 1 if player 1 is not interested and player 2 is not interested, *Info2* takes value 1 if player 1 is not interested and player 2 is interested, *Info3* takes value 1 if player 1 is interested and player 2 is not interested, and *Info4* takes value 1 if player 1 is interested and player 2 is interested.

[28] The variable *Lying_av* takes value 5 if the subject chooses to report the truth in all the five situations/questions of the lying aversion test (see the instructions in Appendix B). The variable takes then value 4 if she tells the truth in situations 2, 3, 4 and 5; value 3 if she tells the truth in situations 3, 4 and 5; value 2 if she tells the truth in situations 4 and 5; value 1 if she tells the truth in situation 5; and value 0 if the subject never tells the truth. In all other cases, the subject is classified as inconsistent (with respect to lying aversion) and her observations are not used in the estimation.

[29] We note that the correlations between the variable *Lying_av* and the social preference variables (*Envy* and *Pro-sociality*) are negligible and not significant.

precisely, when considering all rounds, we find that the probability of sending true messages for subjects with an envious attitude decreases by 13 percentage points when both players are not interested, and 6 percentage points when only the sender is not interested (although the latter is not statistically significant). If we restrict attention to the second half of the experiment we see larger decreases, of 14 and 12 percentage points, respectively, both statistically significant.

Table 7. Marginal effects of envy, pro-sociality and lying aversion on the probability of sending a true message by values of Int1 and Int2.

	Info. set	Marginal effect of Envy	Marginal effect of Pro-sociality	Marginal effect of Lying aversion
All Rounds	*P1 not int.* & *P2 not int.*	-0.135*** (0.039)		-0.004 (0.019)
	P1 not int. & *P2 int.*	-0.065 (0.044)		0.003 (0.021)
	P1 int. & *P2 not int.*	0.035 (0.051)		-0.002 (0.014)
	P1 int. & *P2 int.*	0.099 (0.123)	-0.106 (0.137)	0.042* (0.023)
Rounds 21-40	*P1 not int.* & *P2 not int.*	-0.142*** (0.045)		-0.004 (0.022)
	P1 not int. & *P2 int.*	-0.126** (0.051)		-0.006 (0.022)
	P1 int. & *P2 not int.*	0.015 (0.056)		-0.005 (0.017)
	P1 int. & *P2 int.*	0.084 (0.138)	-0.161 (0.142)	0.068*** (0.022)

[***, **, * significant at the 1%, 5%, and 10% level, respectively] with robust standard errors at the matching group level. For each information set, the marginal effects are computed at round 20 (All Rounds) and 30 (Rounds 21-40) and at the average of all other covariates.

Such findings are somewhat surprising. As shown above, at the informative equilibrium, where information transmission is maximal, the distribution of the gains is asymmetric and favors the receiver. This may suggest the conjecture that the more limited level of information transmission observed in the experiment could at least be partly explained by envy traits in subjects, while non-pro-sociality should not play a role. Additionally, the results reported in Table 7 show that in the absence of conflicts of interest, while social preferences indeed contribute to explaining the untruthful reports by senders (the relevant trait is envy), lying aversion (right panel) does not play a role.

Next, we turn our attention to the situations in which the informative equilibrium prescribes the sender to lie. This happens when both the sender and the receiver are interested in the object. The results for this case, reported in the fourth line of Table 7, show that the marginal effects of envy and pro-sociality on the probability of sending a true message are not significantly different from zero. Instead, in this information set (where both the sender and the receiver are interested in the object) we find that lying aversion significantly increases the probability that the sender chooses a truthful message. In particular, when considering all rounds, we find that the probability of sending a true message increases by 0.04 for each level of lying aversion. Thus, a fully lying-averse subject is around 20 percentage points more likely to send a true message compared to a subject with no lying aversion (our lying cost has 5 levels, times 4%). In the second half of the experiment, these effects become stronger: the probability increase per level of lying aversion rises to 0.07, meaning that a fully lying-averse subject is around 35 percentage points more likely to send a true message than one with no lying aversion.

Our findings are thus in line with those by Brandts and Charness (2003), who argue that truth-telling, when not individually advantageous, is often the product of a social norm against lying, rather than the result of distributional preferences, i.e. the tension between pure self-interest and aversion to lying. It is then the aversion to lying that contributes to explaining the departure from the prescribed equilibrium behavior in this information set, where 1/3 of the subjects sent a truthful message instead of lying.

Altogether, we note that the general picture which emerges from Table 7 is that distributional preferences (envy and pro-sociality) and social norms (lying aversion) could play a role or not depending on the specific context in which the subject is called to decide. While envy contributes to explain subjects' behavior in the information sets where they are not interested, lying aversion plays a role in events where subjects face a conflict of interest. This observation provides clear support to our conjecture that a model including both social traits and lying aversion is the most appropriate to explain observed behavior.

5.2 Senders' behavioral strategies

In this section we further investigate the heterogeneity of senders' choices building on the econometric model in Costa-Gomes, Crawford and Broseta (2001), suitably adapted to our framework. In this analysis, we allow for different possible types of the subjects' preferences. Each sender is then assumed to best respond, on the basis of some appropriately identified preferences, to the receivers' behavior in her group, with departures from it arising from random mistakes (see Appendix E for a description of the estimation technical details). The mixture model is estimated allowing every preference type to have a different error rate.

In particular, we consider the following preference types: (i) the one given by equation (1), which includes both lying aversion and social preferences (and was used before to derive best replies in Table 6); (ii) a type that only allows for lying aversion (i.e., $\alpha_i = \beta_i = 0$); (iii) a type featuring only social preferences (i.e., $L_i = 0$); (iv) another type in which both traits (social preferences and lying aversion) are present but, in the spirit of Charness and Rabin (2002), the impact of envy is reduced.[30] We consider then one last type, (v), according to which subjects best respond according to a standard utility function.

Prior to the estimation of the model, to get a clearer sense of the different implications for behavior of the four preference types we considered, where social traits and/or lying aversion play a role, in Table 8 we present for each of them the distribution across senders of their theoretical best responses. Each column of Table 8 corresponds to a different preference type. Using this type and the parameter values (regarding social preferences and/or lying aversion) elicited for every individual, we find the best reply of each sender (sending a true message, sending a false message or sending any message) to the average behavior of receivers in her matching group, in every information set. The values reported in the table are the frequencies of senders by type of best response.

The results in Table 8 suggest that, if we fail to consider types featuring the contemporaneous presence of social preferences and lying aversion (i.e., if we only consider types (ii) and (iii)), we cannot account for the senders' observed behavior at the information set level (described in Table 3), whereas by including both traits (i.e., including types (i) and (iv)) we can. For instance, according to type (i) the sender's best response in the first information set (when none of the players is interested) is to send a false message in 23.9% of the cases, to send a true message in 40.3% of the cases and in

[30] In this last specification we treat weakly envious subjects as if they were not envious ($\alpha_i = 0$) and strongly envious subjects as if they were weakly envious ($\alpha_i \in [0, 0.5]$). This specification, in which the impact of envy is still present but reduced, was added because it allows to increase the fit of the mixture model.

the remaining 35.8% both messages (true and false) are best responses. These numbers align quite well with the observed frequency of true messages by senders in this information set, as reported in Table 3 (67.58%). Similar observations can be made in the other information sets. Likewise, with type (iv), in which again both social preferences and lying aversion are taken into account, but with a lower impact for envy, we can again say that the frequencies of best responses are consistent with the observed data.

Table 8. Distribution of senders' best responses, according to different preference types

Info. Set	Message	(i) Both soc. pref. and lying av.	(ii) Only lying aversion	(iii) Only social preferences	(iv) Both soc. pref. and lying av. with 'reduced envy'
P1 not int.	False	23.9		56.4	8.2
&	True	40.3	82.5	1.5	62.7
P2 not int.	Both	35.8	17.5	42.1	29.1
P1 not int.	False	26.9		56.4	8.2
&	True	38.8	82.5	1.5	59.7
P2 int.	Both	34.3	17.5	42.1	32.1
P1 int.	False			2.1	
&	True	96.3	100.0	94.4	100.0
P2 not int.	Both	3.7		3.6	
P1 int.	False	63.6	77.6	96.9	58.2
&	True	9.1	10.5	2.1	13.4
P2 int.	Both	27.3	11.9	1.0	28.4

In contrast, if we examine type (iii), in which only social preferences are considered, the theoretical best response of the majority of subjects is to lie in the first two information sets. The observation of a significantly lower amount of lying in the data can be attributed to the presence of lying averse individuals who are less willing to lie. If instead only lying aversion is taken into account, as in type (ii), most senders would tell the truth, in all information sets, again not in line with what we see in the data. Thus, these types cannot account for the observed senders' behavior at the information set level (see Table 3).

In Table 9 we report then the estimated parameter values of the mixture model, in which each sender's type is drawn from a common distribution over types and remains constant for all the periods in which a player acts as a sender. A sender's type is a strategy that is a best response, across all information sets, to the actual behavior of the receivers in her group, computed according to the conditions (i)-(v) we described.[31] We assume that a type k sender chooses the actions prescribed by strategy k with probability $1 - \varepsilon_k$ and makes a mistake with the residual probability $\varepsilon_k \in [0,1]$, describing type k 's error rate. When she makes a mistake, she sends any of the two possible reports with equal probability. We estimate the model for all 40 periods.[32]

Table 9. Error-rate model

Type	*Probability*	*Error rate*
(i) Both social preferences and lying aversion	0.222*** (0.047)	0.024*** (0.049)
(ii) Only lying aversion	0.121*** (0.040)	0.764 (0.255)
(iii) Only social preferences	0.186*** (0.045)	0.120*** (0.150)
(iv) Both soc. pref. and lying av. with 'reduced envy'	0.157*** (0.048)	0.046*** (0.118)
(v) Standard preferences	0.314*** (0.064)	0.271*** (0.105)

Standard errors within brackets
Significance: Probabilities (with respect to 0) and Error rates (with respect to 1): ***, **, * at the 1%, 5%, 10% level, resp. The Log Likelihood of the estimation is -899.755.

As Table 9 shows, the specifications that take into account both social traits and lying aversion, i.e. (i) and (iv), jointly account for almost 40% of the subjects. The two specifications that consider only lying aversion (ii) or only social preferences (iii) fit then the behavior of around 30% of the subjects (as we see from the table, the behavior of subjects classified in specification (ii) is quite noisy). The behavior of the remaining 30% of subjects is best represented by the specification with standard preferences (v). Thus, the analysis in this section supports the role of the combination of social preferences and lying aversion to explain subjects' behavior not only at the information set level but also across information sets.

[31] In the case of standard preferences, to partially solve for indifferences in those information sets where the sender is not interested, we assume that those subjects with high lying costs (higher than 12) best respond by sending the true message, and those with lower lying costs send any message. Alternative specifications provide similar results.

[32] The results in Table 9 are robust to the consideration of specifications in which either we do not include type (iv) or type (i) -see Table 16 in Appendix E. The log likelihood of the reported estimation is -899.755, which is higher than the values obtained if either we do not include type (iv) or type (i) (equal, respectively, to -968.018 and -992.414). Given that the reported specification only requires to estimate two additional parameters with respect to these alternatives, it is preferred according to the Akaike information criterion.

5.3. Senders without material concerns for the receivers' response

We have seen that subjects send false reports in situations in which they do not have any material interest in doing so, as their payoff is not affected by the choices made by receivers. We then found that social preferences play a significant role in explaining this fact.

One significant advantage of our design is that both the sender and the receiver are aware of each other's interests, hence strategic considerations have an important role as we argue happens in various realistic applications. We should point out, however, that in the game we analyzed strategic considerations come both from the sender's material interest for the object, as well as from social preferences, that we established play a role.

To assess the importance of these strategic considerations, we consider an additional treatment, labelled *No conflict*, in which the sender is never interested in the object and hence does not have any material motive to tell lies. This simplifies a bit the inference problem of the receiver, which is still nontrivial if he anticipates the sender may have social preferences. Specifically, in this treatment the receiver's objective is to match the state and the sender's payoff is independent of the state and the receiver's action. The game is a reduced form of the game considered so far (see Table 1), where the sender is never interested in the object (and, hence, has no conflict of interest). In this modified game the sender observes the state and sends a message to the receiver who subsequently chooses an action (*Left*, *Center* or *Right*). The payoffs are reported in Table 10.

Table 10. Payoffs in treatment *No conflict*

Player 2 not interested			*Player 2 interested*		
Left	Center	Right	Left	Center	Right
20, 20	20, 60	20, 100	20, 160	20, 120	20, 50

We run three sessions of treatment *No conflict* at the laboratory of experimental economics of the University of Siena (LabSi) in December 2021. A total of 36 subjects participated in these sessions, yielding 9 groups of 4 subjects.

In the following, we compare the behavior of the sender and the receiver between treatment *No conflict* and our original treatment (labelled *Baseline*).[33] In Table 11 we consider the sender. We report the frequencies of true messages in the *Baseline* treatment

[33] All the results reported in this section are robust if we exclude from the comparison the data of the sessions of the *Baseline* treatment run in 2014 (see Footnote 21).

(focusing on the information sets in which the sender is not interested) and in the *No conflict* treatment. We observe that the frequencies of true messages are quite different in the two treatments. Specifically, in the *No conflict* treatment there is significantly more truth-telling, especially when the receiver is interested.

Table 11. Player 1. Frequencies of true messages - comparison between treatments

	All Rounds			Rounds 21-40		
	P2 not interested	*P2 interested*	Total	*P2 not interested*	*P2 interested*	Total
Baseline (with P1 not int.)	67.58 [1027]	63.56** [1084]	65.51** [2111]	68.69 [511]	64.74** [536]	66.67* [1047]
No conflict	77.95 [381]	76.40 [339]	77.22 [720]	76.62 [201]	81.13 [159]	78.61 [360]

[number of observations]
***, **, * significant at the 1%, 5%, and 10% level, respectively, at two-sample t-test.

In Table 12 we examine then the behavior of the receiver. We report the frequencies of actions *Left*, *Center* and *Right* in the *Baseline* treatment and in the *No conflict* treatment.

Table 12. Player 2. Frequencies of choices - comparison between treatments

		All Rounds			Rounds 21-40		
		Left	Center	Right	Left	Center	Right
Baseline	*Message NoInt*	9.25 [234]	33.58** [850]	57.17 [1447]	8.63 [111]	35.93* [462]	55.44 [713]
	Message Int	53.65* [831]	34.60* [536]	11.75 [182]	53.58* [404]	35.94* [271]	10.48 [79]
No conflict	*Message NoInt*	13.26 [50]	20.42 [77]	66.31 [250]	11.41 [21]	21.20 [39]	67.39 [124]
	Message Int	64.72 [222]	25.36 [87]	9.91 [34]	67.61 [119]	22.73 [40]	9.66 [17]

[number of observations]
***, **, * significant at the 1%, 5%, and 10% level, respectively, at two-sample t-test.

We observe that the frequencies of responses are rather different in quite a few of the cases between the *Baseline* treatment, in which there could be a conflict of interest, and the *No conflict* treatment, where this conflict is never present. It is especially notable the large decrease in "moderate" center responses both when the message is *Interested* as well as when it is *Not-Interested*. This difference is persistent over time. Since in the *No conflict* treatment there is more truth telling by the sender, the receiver responds by choosing less often *Center* (which is optimal when the sender is babbling), and for instance, by increasing the choice of *Left* when the message received is *Interested* (which is optimal if the message is credible).

Thus, these results suggest that the fact that the receiver is aware of a potential conflict of interest with the sender does matter for the behavior of both players and induces more deception by senders. In this sense, our study allows to gain some understanding for the determinants of behavior in situations in which the interests of senders and receivers might be misaligned, a consideration that is absent in former literature that studies deception (e.g., Gneezy (2005), or Erat and Gneezy (2012)).

6. Conclusion

The traditional experimental economics literature on cheap talk has generally painted a positive view on the ability of communication to improve social welfare. Because of humans' moral tendency to avoid falsehood, individuals eschew lying even in circumstances where it would materially benefit them. This improves the chances of communication to coordinate social behavior. In this paper we introduce a counterpoint to that generally held belief. Many humans also exhibit a tendency to act to avoid inequality that goes against them. We show those individuals may be willing to overcome the aversion to lying if this could reduce a payoff gap with respect to others. This is true even when the lie does not benefit them.

We were able to find this because our experimental design, unlike the commonly used one, allows for the possibility that no material conflict of interest is present, but truth-telling – if believed – yields a material benefit for the receiver, not for the sender. We were able to verify that the individuals who choose deception in such situations tend to be envious, because we measure independently the envy and pro-sociality of our experimental participants.

We believe our results open interesting new avenues of research. One could investigate whether our results also hold in the field, for tasks that are common in everyday life. Would a trader at the government bonds desk pass really valuable information about a stock to a colleague in the stocks desk who thereby might enjoy an important promotion? What would be the personal characteristics of such a bond trader (and the stocks one) that would make that kind of communication more likely?

Declaration of interest: The authors declare that they have no relevant material or financial interests that relate to the research described in this article.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, *87*(4), 1115-1153.

Agranov, M., Dasgupta, U., & Schotter, A. (2020). Trust me: Communication and competition in psychological games. Working paper.

Amato, C., Gino, F., Montinari, N., & Sacco, P. (2020). Cheating, inequality aversion, and appealing to social norms. *Journal of Economic Behavior & Organization*, *179*, 767-776.

Barron, Stuber, & van Veldhuizen (2019). Motivated motive selection in the lying-dictator game. Discussion Paper, SP II 2019-303.

Bartling, B., Fehr, E., Maréchal, M. A., & Schunk, D. (2009). Egalitarianism and competitiveness. *The American Economic Review*, *99*(2), 93-98.

Blume, A., DeJong, D. V., Kim, Y. G., & Sprinkle, G. B. (1998). Experimental evidence on the evolution of meaning of messages in sender-receiver games. *The American Economic Review*, 1323-1340.

Blume, A., DeJong, D. V., Kim, Y. G., & Sprinkle, G. B. (2001). Evolution of communication with partial common interest. *Games and Economic Behavior*, *37*(1), 79-120.

Bolton, P., Freixas, X., & Shapiro, J. (2007). Conflicts of interest, information provision, and competition in the financial services industry. *Journal of Financial Economics*, *85*(2), 297-330.

Brandts, J., & Charness, G. (2003). Truth or consequences: An experiment. *Management Science*, *49*(1), 116-130.

Cabrales, A., Feri, F., Gottardi, P., & Meléndez-Jiménez, M. A. (2020). Can there be a market for cheap-talk information? An experimental investigation. *Games and Economic Behavior*, 121, 368-381.

Cai, H., & Wang, J. T. Y. (2006). Overcommunication in strategic information transmission games. *Games and Economic Behavior*, *56*(1), 7-36.

Charness, G., & Gneezy, U. (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry*, *48*(1), 133-146.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, *117*(3), 817-869.

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, *69*(5), 1193-1235.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431-1451.

Dickhaut, J. W., McCabe, K. A., & Mukherji, A. (1995). An experimental study of strategic information transmission. *Economic Theory*, *6*(3), 389-403.

Erat, S., & Gneezy, U. (2012). White lies. *Management Science*, *58*(4), 723-733.

Farina, A., Fréchette, G. R., Ispano, A., Lizzeri, A., & Perego, J. (2024). The selective disclosure of evidence: An experiment. Mimeo.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817-868.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Fréchette, G. R., Lizzeri, A., & Perego, J. (2021). Rules and commitment in communication: An experimental analysis. Mimeo.

Gibson, R., Tanner, C., & Wagner, A. (2013). Preferences for truthfulness: Heterogeneity among and within individuals. *The American Economic Review*, 103(1). 532–548.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384-394.

Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying Aversion and the Size of the Lie. *AmericanEconomic Review*, 108(2), 419–453.

Gneezy, U., Saccardo, S., Serra-Garcia, M., & van Veldhuizen, R. (2020). Bribing the self. *Games and Economic Behavior*, *120*, 311-324.

Jacobsen, C., Fosgaard, T. R., & Pascual-Ezama, D. (2018). Why do we lie? A practical guide to the dishonesty literature. *Journal of Economic Surveys*, *32*(2), 357-387.

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, *76*(4), 1359-1395.

Kartik, N., Ottaviani, M., & Squintani, F. (2007). Credulity, lies, and costly talk. *Journal of Economic Theory*, *134*(1), 93-116.

Kawagoe, T., & Takizawa, H. (1999). Instability of babbling equilibrium in cheap talk games. Saitama University.

Kawagoe, T., & Takizawa, H. (2005). Why lying pays: Truth bias in the communication with conflicting interests. *Available at SSRN 691641*.

Kerschbamer, R., Neururer, D., & Gruber, A. (2019). Do altruists lie less? *Journal of Economic Behavior & Organization*, *157*, 560-579.

Lafky, J., Lai, E. K., & Lim, W. (2022). Preferences vs. strategic thinking: An investigation of the causes of overcommunication. *Games and Economic Behavior 136*, 92–116.

Maggian, V., & Villeval, M.C. (2016). Social preferences and lying aversion in children. *Experimental Economics*, *19*, 663-685.

Sánchez-Pagés, S., & Vorsatz, M. (2007). An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior*, *61*(1), 86-112.

Sobel, J. (2013). Giving and receiving advice, in *Advances in Economics and Econometrics*, D. Acemoglu, M. Arellano, and E. Dekel (eds.).

Thordal-Le Quement, M. (2016). The (human) sampler's curses. *American Economic Journal: Microeconomics 8*(4), 115-48.

Vranceanu, R., & Dubart, D. (2019). Deceitful communication in a sender-receiver experiment: Does everyone have a price? *Journal of Behavioral and Experimental Economics 79*, 43-52.

Wang, J. T. Y., Spezio, M., & Camerer, C. F. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *American Economic Review*, *100*(3), 984-1007.

APPENDIX

A) EQUILIBRIUM ANALYSIS OF THE SENDER-RECEIVER GAME

In Table 13 below, we present the expected payoffs to players 1 and 2 associated with all of the pure strategy profiles and identify the Bayes-Nash equilibria (in grey), i.e., the candidates for perfect Bayesian equilibria. We then check which Bayes-Nash equilibria are perfect Bayesian equilibria. In the rows of Table 13, we list all of the strategy profiles of player 1 and in the columns those of player 2. A strategy of player 1 (the sender) is a vector of four components $\left(c_{ni,ni}, c_{ni,i}, c_{i,ni}, c_{i,i}\right) \in \{0,1\}^4$, where $c_{ni,ni}$ is the sender's choice in the information set in neither the sender nor the receiver are interested in the object, $c_{ni,i}$ is the choice in the information set in which the sender is not interested and the receiver is interested, $c_{i,ni}$ represents the choice in the information set in which the sender is interested and the receiver is not, and $c_{i,i}$ represents the choice in the information set in which both of them are interested. A choice $c = 1$ indicates that the sender sends a true message, while $c = 0$ indicates that the sender sends a false message. A strategy of player 2 (the receiver) is a vector of two components $\left(c_{m\_ni}, c_{m\_i}\right) \in \{L, C, R\}^2$. The component c_{m_ni} represents player 2's choice in the information set in which player 1's message says that player 2 is not interested in the object (i.e., the color reported in the message does not coincide with player 2's assigned color); c_{m_i} represents player 2's choice in the information set in which player 1's message says that player 2 is interested in the object. For each information set, the choice $c = L$ indicates that player 2 chooses action *left*, the choice $c = C$ indicates that player 2 chooses action *center*, and the choice $c = R$ indicates that player 2 chooses action *right*.

Each cell of Table 13 contains a vector with the (*ex-ante*) expected payoffs to player 1 and player 2, associated with the respective strategies of players 1 and 2 indicated by row and column. The expected payoffs are computed using the payoffs contained in Table 1, taking into account that each of the four possible combinations of players 1 and 2 being interested/not interested in the object has an *ex ante* probability of 1/4.

In this table, we also mark in bold the best responses of players 1 and 2 and fill in grey those cells in which the strategies of players 1 and 2 are mutual best responses, i.e., the (pure strategy) Bayes-Nash equilibria.

Table 13. Bayes-Nash equilibria

	(L,L)	(L,C)	(L,R)	(C,L)	(C,C)	(C,R)	(R,L)	(R,C)	(R,R)
(1, 1, 1, 1)	(**20.0**, 67.5)	(32.5, 55.0)	(45.0, 35.0)	(32.5, 92.5)	(**45.0**, 80.0)	(57.5, 60.0)	(45.0, **110**)	(57.5, 97.5)	(**70.0**, 77.5)
(1, 1, 1, 0)	(**20.0**, 67.5)	(20.0, 57.5)	(20.0, 40.0)	(**45.0**, 90.0)	(**45.0**, 80.0)	(45.0, 62.5)	(**70.0**, **105**)	(**70.0**, 95.0)	(**70.0**, 77.5)
(1, 1, 0, 1)	(**20.0**, 67.5)	(**45.0**, 70.0)	(**70.0**, 57.5)	(20.0, 77.5)	(**45.0**, 80.0)	(**70.0**, 67.5)	(20.0, 87.5)	(45.0, **90.0**)	(**70.0**, 77.5)
(1, 0, 1, 1)	(**20.0**, 67.5)	(32.5, 65.0)	(45.0, 62.5)	(32.5, **82.5**)	(**45.0**, 80.0)	(57.5, 77.5)	(45.0, **82.5**)	(57.5, 80.0)	(**70.0**, 77.5)
(0, 1, 1, 1)	(**20.0**, 67.5)	(32.5, 65.0)	(45.0, 55.0)	(32.5, 82.5)	(**45.0**, 80.0)	(57.5, 70.0)	(45.0, **90.0**)	(57.5, 87.5)	(**70.0**, 77.5)
(1, 1, 0, 0)	(**20.0**, 67.5)	(32.5, 72.5)	(45.0, 62.5)	(32.5, 75.0)	(**45.0**, 80.0)	(57.5, 70.0)	(45.0, 82.5)	(57.5, **87.5**)	(**70.0**, 77.5)
(1, 0, 1, 0)	(**20.0**, 67.5)	(20.0, 67.5)	(20.0, 67.5)	(**45.0**, **80.0**)	(**45.0**, **80.0**)	(45.0, **80.0**)	(**70.0**, 77.5)	(**70.0**, 77.5)	(**70.0**, 77.5)
(0, 1, 1, 0)	(**20.0**, 67.5)	(20.0, 67.5)	(20.0, 60.0)	(**45.0**, 80.0)	(**45.0**, 80.0)	(45.0, 72.5)	(**70.0**, **85.0**)	(**70.0**, **85.0**)	(**70.0**, 77.5)
(1, 0, 0, 1)	(**20.0**, 67.5)	(**45.0**, 80.0)	(**70.0**, **85.0**)	(20.0, 67.5)	(**45.0**, 80.0)	(**70.0**, **85.0**)	(20.0, 60.0)	(45.0, 72.5)	(**70.0**, 77.5)
(0, 1, 0, 1)	(**20.0**, 67.5)	(**45.0**, **80.0**)	(**70.0**, 77.5)	(20.0, 67.5)	(**45.0**, **80.0**)	(**70.0**, 77.5)	(20.0, 67.5)	(45.0, **80.0**)	(**70.0**, 77.5)
(0, 0, 1, 1)	(**20.0**, 67.5)	(32.5, 75.0)	(45.0, 82.5)	(32.5, 72.5)	(**45.0**, 80.0)	(57.5, **87.5**)	(45.0, 62.5)	(57.5, 70.0)	(**70.0**, 77.5)
(1, 0, 0, 0)	(**20.0**, 67.5)	(32.5, 82.5)	(45.0, **90.0**)	(32.5, 65.0)	(**45.0**, 80.0)	(57.5, 87.5)	(45.0, 55.0)	(57.5, 70.0)	(**70.0**, 77.5)
(0, 1, 0, 0)	(**20.0**, 67.5)	(32.5, **82.5**)	(45.0, **82.5**)	(32.5, 65.0)	(**45.0**, 80.0)	(57.5, 80.0)	(45.0, 62.5)	(57.5, 77.5)	(**70.0**, 77.5)
(0, 0, 1, 0)	(**20.0**, 67.5)	(20.0, 77.5)	(20.0, 87.5)	(**45.0,** 70.0)	(**45.0**, 80.0)	(45.0, **90.0**)	(**70.0**, 57.5)	(**70.0**, 67.5)	(**70.0**, 77.5)
(0, 0, 0, 1)	(**20.0**, 67.5)	(**45.0**, 90.0)	(**70.0**, **105**)	(20.0, 57.5)	(**45.0**, 80.0)	(**70.0**, 95.0)	(20.0, 40.0)	(45.0, 62.5)	(**70.0**, 77.5)
(0, 0, 0, 0)	(**20.0**, 67.5)	(32.5, 92.5)	(45.0, **110**)	(32.5, 55.0)	(**45.0**, 80.0)	(57.5, 97.5)	(45.0, 35.0)	(57.5, 60.0)	(**70.0**, 77.5)

We find that there are 10 Bayes-Nash equilibria that can be grouped into 3 classes (*informative* equilibria, *babbling* equilibria and *opposite* equilibria). The equilibria within each class are informationally equivalent and only differ: (i) in the use of colors, as each equilibrium has a reverse one, and (ii) in the case of the babbling and opposite equilibria, also in the choice of player 2 in one of his information sets, which can be either *left* or *center*. The equilibria for this game are as follows:

1.- *Informative equilibrium*:

- $((1,1,1,0),(R,L))$
- $((0,0,0,1),(L,R))$

The expected payoffs for players 1 and 2 are 70 and 105, respectively.

2.- *Babbling equilibrium*:

- $((1,0,1,0),(C,C\ or\ L))$
- $((0,1,0,1),(C\ or\ L,C))$

The expected payoffs for players 1 and 2 are 45 and 80, respectively.

3.- *Opposite equilibrium*:

- $((0,1,1,0),(R, L\ or\ C))$
- $((1,0,0,1),(L\ or\ C, R))$

The expected payoffs for players 1 and 2 are 70 and 85, respectively.

We now check that all the Bayes-Nash equilibria we found in Table 13 also constitute (weak) perfect Bayesian equilibria, hereafter PBE.

1.- *Informative equilibrium*

As both profiles are informationally equivalent (they only differ in the use of colors), let us consider the profile $((1,1,1,0),(R, L))$.

1.i) *Player 2's beliefs*

If the message of player 1 says that player 2 is not interested in the object, then, given the strategy of player 1, the beliefs of player 2 derived from Bayes' rule assign equal probability (1/3) to the following three events: (i) player 2 is not interested in the object and player 1 is interested, (ii) neither player 2 nor player 1 is interested in the object, and (iii) both players 2 and 1 are interested in the object.

If the message of player 1 says that player 2 is interested in the object, then, given the strategy of player 1, the beliefs of player 2 assign probability 1 to the following event: player 2 is interested in the object, and player 1 is not interested.

1.ii) *Sequential rationality of player 2*

If the message of player 1 says that player 2 is not interested in the object, then, given the beliefs above, the expected payoffs of player 2 associated with the choices L, C and R are 110/3, 200/3 and 260/3, respectively (see Table 7). Thus, the choice $c_{m\_ni} = R$ prescribed by the strategy of player 2 is sequentially rational.

If the message of player 1 says that player 2 is interested in the object, then, given the beliefs above, the expected payoffs of player 2 associated with the choices L, C and R are 160, 120 and 50, respectively. Thus the choice $c_{m\_i} = L$ prescribed by the strategy of player 2 is sequentially rational.

1.iii) *Sequential rationality of player 1*

If player 1 is not interested in the object, his payoff to player 1 is 20, regardless the choices of players 1 and 2. Thus, the choices $c_{ni,ni} = 1$ and $c_{ni,i} = 1$ prescribed by the strategy of player 1 are sequentially rational.

If player 1 is interested in the object and player 2 is not interested, then, given the strategy of player 2, (R, L), the payoffs to player 1 associated with sending a true and false message are, respectively, 120 and 20. Thus the choice $c_{i,ni} = 1$ prescribed by the strategy of player 1 is sequentially rational.

If both player 1 and player 2 are interested in the object, then, given the strategy of player 2, (R, L), the payoffs to player 1 associated with sending a true and a false message are, respectively, 20 and 120 (see Table 1). Thus the choice $c_{i,i} = 0$ prescribed by the strategy of player 1 is sequentially rational.

Hence, $((1,1,1,0), (R, L))$ is a PBE, and therefore, $((0,0,0,1), (L, R))$ is also a PBE.

2.- *Babbling equilibrium*

Because we have two sets of profiles that are informationally equivalent (they only differ in the use of colors), let us focus on the profiles $\left((1,0,1,0), (C, C)\right)$ and $((1,0,1,0), (C, L))$

2.i) *Player 2's beliefs*

In this case, the message of player 1 is not correlated with the state of the world: it always says that player 2 is not interested. Thus, if the message of player 1 says that player 2 is not interested in the object, then the beliefs of player 2 using Bayes' rule assign equal probability (1/4) to each of the four possible events regarding whether players 1 and 2 are interested in the object.

If the message of player 1 says that player 2 is interested in the object (which does not happen on the equilibrium path), then beliefs cannot be determined by Bayes' rule and are specified below.

2.ii) *Sequential rationality of player 2*

If the message of player 1 says that player 2 is not interested in the object, then, given the beliefs above, the expected payoffs to player 2 associated with the choices L, C and R are 270/4, 320/4 and 310/4, respectively. Thus, the choice $c_{m\_ni} = C$ prescribed by the strategy of player 2 is sequentially rational.

If the message of player 1 says that player 2 is interested in the object, then we can find beliefs such that both the choices (i) $c_{m\_i} = C$ and (ii) $c_{m\_i} = L$ are sequentially rational (the beliefs are free in this case). For instance, if the beliefs in this information set are the same as in the former one (i.e., all four possible events have the same probability), then the choice $c_{m\_i} = C$ is sequentially rational. Alternatively, if the beliefs in this information set assign probability 1 to the event in which player 2 is interested in the object and player 1 is not, then the choice $c_{m\_i} = L$ is sequentially rational.

2.iii) *Sequential rationality of player 1*

If player 1 is not interested in the object, then the payoff to player 1 is 20, regardless of the choices of players 1 and 2. Thus the choices $c_{ni,ni} = 1$ and $c_{ni,i} = 0$ prescribed by the strategy of player 1 are sequentially rational.

If player 1 is interested in the object and player 2 is not interested, then we have the following:

- Given the strategy of player 2 in the first equilibrium in this class, $\text{i.e.}, (C, C)$, the payoff of player 1 is 70 regardless of whether he sends a true or a false message. Thus, the choice $c_{i,ni} = 1$ prescribed by the strategy of player 1 is sequentially rational.
- Given the strategy of player 2 in the second equilibrium in this class, $\text{i.e.}, (C, L)$, the payoffs of player 1 associated with sending a true and a false message are 70 and 20, respectively. Thus, the choice $c_{i,ni} = 1$ prescribed by the strategy of player 1 is sequentially rational.

If both players 1 and 2 are interested in the object, then we have the following:

- Given the strategy of player 2 in the first equilibrium in this class, $\text{i.e.}, (C, C)$, the payoff of player 1 is 70 regardless of whether he sends a true or a false message. Thus, the choice $c_{i,ni} = 0$ prescribed by the strategy of player 1 is sequentially rational.
- Given the strategy of player 2 in the second equilibrium in this class, $\text{i.e.}, (C, L)$, the payoffs of player 1 associated with sending a true and a false message are,20 and 70, respectively. Thus, the choice $c_{i,ni} = 0$ prescribed by the strategy of player 1 is sequentially rational.

Hence, the profiles $\big((1,0,1,0), (C, L\ or\ C)\big)$ are PBE, and therefore, $((0,1,0,1), (L\ or\ C, C))$ also are PBE.

3.- *Opposite equilibrium*

As we have two sets of profiles that are informationally equivalent (they only differ in the use of colors), let us focus on the profiles $((0,1,1,0), (R, L))$ and $((0,1,1,0), (R, C))$

3.i) *Player 2's beliefs*

If the message of player 1 says that player 2 is not interested in the object, then, given the strategy of player 1, the beliefs derived from Bayes' rule of player 2 assign equal probability (1/2) to the following two events: (i) player 2 is not interested and player 1 is interested in the object, and (ii) both players 2 and 1 are interested in the object.

If the message of player 1 says that player 2 is interested in the object, then, given the strategy of player 1, the beliefs of player 2 assign equal probability (1/2) to the following two events: (i) neither player 2 nor player 1 is interested in the object, and (ii) player 2 is interested in the object and player 1 is not interested.

3.ii) *Sequential rationality of player 2*

If the message of player 1 says that player 2 is not interested in the object, then, given the beliefs above, the expected payoffs of player 2 associated with the choices L, C and R are 90/2, 140/2 and 160/2, respectively. Thus the choice $c_{m\_ni} = R$ prescribed by the strategy of player 2 is sequentially rational.

If the message of player 1 says that player 2 is interested in the object, then, given the beliefs above, the expected payoffs of player 2 associated with the choices L, C and R are 180/2, 180/2 and 150/2, respectively. Thus, the choices $c_{m\_i} = L$ and $c_{m\_i} = C$ are sequentially rational.

3.iii) *Sequential rationality of player 1*

If player 1 is not interested in the object, then his payoff is 20, regardless of the choices of players 1 and 2. Thus, the choices $c_{ni,ni} = 0$ and $c_{ni,i} = 1$ prescribed by the strategy of player 1 are sequentially rational.

If player 1 is interested in the object and player 2 is not interested, then we have the following:

- Given the strategy of player 2 in the first equilibrium in this class, i.e.(R, L)), the payoffs of player 1 associated with sending a true and a false message are 120 and 20, respectively. Thus, the choice $c_{i,ni} = 1$ prescribed by the strategy of player 1 is sequentially rational.
- Given the strategy of player 2 in the second equilibrium in this class, i.e., (R, C), the payoffs of player 1 associated with sending a true and a false message are 120 and 70, respectively. Thus, the choice $c_{i,ni} = 1$ prescribed by the strategy of player 1 is sequentially rational.

If both player 1 and player 2 are interested in the object, then we have the following:

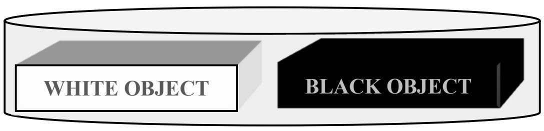
- Given the strategy of player 2 in the first equilibrium in this class, i.e., (R, L), the payoffs of player 1 associated with sending a true and a false message are 20 and 120, respectively. Thus, the choice $c_{i,i} = 0$ prescribed by the strategy of player 1 is sequentially rational.
- Given the strategy of player 2 in the second equilibrium in this class, i.e., (R, C), the payoffs of player 1 associated with sending a true and a false message are 70 and 120, respectively. Thus, the choice $c_{i,i} = 0$ prescribed by the strategy of player 1 is sequentially rational.

Hence, $((0,1,1,0), (R, L\ or\ C))$ is a PBE, and therefore, $((1,0,0,1), (L\ or\ C, R))$ is also a PBE.

B) EXPERIMENTAL INSTRUCTIONS

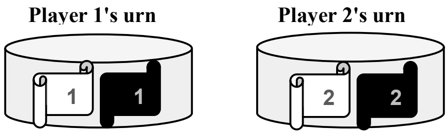
The aim of this experiment is to study how individuals make decisions in certain contexts. The instructions are simple. If you follow them carefully you will earn a non-negligible amount of money in cash (euros) at the end of the experiment. During the experiment, your earnings will be in ECUs (experimental currency units). Individual payments will remain private, as nobody will know the other participants' payments. Any communication among you is strictly forbidden and will result in immediate exclusion from the experiment.

1. The experiment consists of 40 rounds. In each round, you will be randomly assigned to a group of 2 participants (including yourself). This group is determined randomly at the beginning of the round. Therefore, the group you are assigned to changes at each round. In this room, there are 4 participants (including yourself) who are potential members of your group. That is, at every round, your group is selected among these 4 participants, each of them being equally likely to be in your group. You will not know the identities of any of these participants. In each round, you will only interact with the other participant in your group, and your payoff will only depend on your choice and the choice of the other participant in your group.
2. In each round, one of the two participants in your group will have the role of player 1 and the other one will have the role of player 2. The roles will be randomly assigned, and both participants in a group are equally likely to have each role assigned. At the beginning of the round, each participant will be informed of his/her assigned role.
3. At the beginning of the round, the computer randomly draws one object from an (virtual) urn containing two objects: one white object and one black object. Each object is picked with equal probability (50%).



   The color of the object is revealed to player 1 but not to player 2 in your group.
4. At each round, each player is assigned a color. At the beginning of the round, the color assigned to each player is determined in the following way. There is one (virtual) urn for each player, containing two pieces of paper: one white and one black. The computer randomly (and independently) draws one piece of paper from each urn. In

each urn, each piece of paper is picked with equal probability (50%). The piece of paper selected for each player determines that player's assigned color.



In each group, player 1 is informed both of his/her assigned color for the round and of the color assigned to player 2. Player 2 is only informed of his/her assigned color but not of the color assigned to player 1.

5. At each round, in each group, player 1 will be the first to make his/her decisions, knowing the color of the object drawn from the computer, his/her assigned color, and the assigned color of player 2. Player 1 has to decide what message to send to player 2 regarding the color of the object (which is unknown by player 2). The message can be either "*The object is white*" or "*The object is black*". Thus, the message can contain the true color or the false one.
6. Then, player 2, being informed of his/her assigned color (but neither of the color of the object nor of the color assigned to player 1), observes the content of the message sent by player 1 and decides which action to take: *Left*, *Center* or *Right*.
7. Round payoffs. At each round, the payoff to each player depends on whether the color of the object did or did not match his/her assigned color and on the action chosen by player 2:
   i. At each round, the payoff to player 1 is determined as follows.
      - If the color of the object is equal to the color assigned to player 1, then his/her payoff depends on the choice of player 2 in the following way:
        - 20 ECUs if player 2 has chosen *Left*
        - 70 ECUs if player 2 has chosen *Center*
        - 120 ECUs if player 2 has chosen *Right*
      - If the color of the object is different from the color assigned to player 1, then his/her payoff is 20 ECUs, regardless of the action chosen by player 2.

   ii. At each round, the payoff of player 2 is determined as follows.
      - If the color of the object is equal to the color assigned to player 2, then action *Left* provides him/her with a higher payoff than action *Center* or *Right*, and action *Center* provides him/her with a higher payoff than action *Right*.

- If the color of the object is different from the color assigned to player 2, then action *Right* provides him/her with a higher payoff than action *Center* or *Left*, and action *Center* provides him/her with a higher payoff than action *Left*.
- The payoff of player 2 also depends on the correspondence between his/her assigned color and the color assigned to player 1: the payoff for player 2 when his/her assigned color is the same than the assigned color of player 1 is *lower* than in the case in which his/her assigned color is different from the color assigned to player 1.

The four tables below provide the payoffs of player 1 and player 2 in all possible situations:

- The top-left table corresponds to the cases in which both players have the same assigned color, which is different from the color of the object;
- The top-right table corresponds to the cases in which the color of the object is equal to the color assigned to player 2 but different from the color assigned to player 1;
- The bottom-left table corresponds to the cases in which the color of the object is equal to the color assigned to player 1 but different from the color assigned to player 2;
- The bottom-right table corresponds to the cases in which both players have the same assigned color, which is equal to the color of the object.

(**Object:** White **Player 2:** Black **Player 1:** Black) **or** (**Object:** Black **Player 2:** White **Player 1:** White)			
Action of player 2	Left	Center	Right
Payoff to player 1	20	20	20
Payoff to player 2	20	60	100

(**Object:** White **Player 2:** White **Player 1:** Black) **or** (**Object:** Black **Player 2:** Black **Player 1:** White)			
Action of player 2	Left	Center	Right
Payoff to player 1	20	20	20
Payoff to player 2	160	120	50

(**Object:** White **Player 2:** Black **Player 1:** White) **or** (**Object:** Black **Player 2:** White **Player 1:** Black)			
Action of player 2	Left	Center	Right
Payoff to player 1	20	70	120
Payoff to player 2	30	90	120

(**Object:** White **Player 2:** White **Player 1:** White) **or** (**Object:** Black **Player 2:** Black **Player 1:** Black)			
Action of player 2	Left	Center	Right
Payoff to player 1	20	70	120
Payoff to player 2	60	50	40

8. At the end of each round, prior to proceeding to the next round, all the players are informed about current and past rounds: assigned role, color of the object, color assigned to each player in the group, the message of player 1, the action chosen by player 2 and the payoff of each player.
9. Payments. At the end of the experiment, you will be paid the earnings that you obtained in 8 rounds (out of 40). These rounds will be randomly selected by the computer: 4 rounds will be selected from the rounds in which you were assigned the role of player 1 and the other 4 rounds will be selected from the rounds in which you were assigned the role of player 2. The earnings that you have obtained in the selected rounds will be converted into cash at the exchange rate of 40 ECUs = 1 euro and will be paid to you in private.

Instructions of the risk test

Now you have the possibility to invest your show-up fee (5 euros). You can decide not to invest or to invest only a share of the 5 euros.

This investment has the following characteristics:

1) With a 50% probability, it pays two and a half times the amount invested
2) With 50% probability, it pays zero

The uninvested amount will be paid to you

How much do you want to invest? (from 0 to 5 euros)

Instructions of the social preferences test

For each row of the following table, select the preferred option, either option A or option B.

Note that there are 4 rows in the table and only one row will be randomly chosen for payment. Each row has an equal chance of being chosen.

In the selected line, the selected option will determine your earnings and that of another participant randomly chosen from those present:

- If you have selected option A, you will receive 2 euros and the other participant will receive 2 euros.
- If you have selected option B, you will receive the earnings specified in the line for option B.

Example: if you choose option B in the second row and this row is selected for payment, you will receive 3 euros and the other participant will receive 1 euro.
Important: the other participant will never know your identity and you will never know the identity of the other participant.

When everyone has selected your preferred options, the computer will randomly choose 2 participants and the row that will determine the payment. The option selected by the first participant in the chosen row will determine the payment of the two participants. The other participants will not be paid for this part of the experiment.

1	Option A:	o	2 euro for you 2 euro for the other participant	Option B:	o	2 euro for you 1 euro for the other participant
2	Option A:	o	2 euro for you 2 euro for the other participant	Option B:	o	3 euro for you 1 euro for the other participant
3	Option A:	o	2 euro for you 2 euro for the other participant	Option B:	o	2 euro for you 4 euro for the other participant
4	Option A:	o	2 euro for you 2 euro for the other participant	Option B:	o	3 euro for you 5 euro for the other participant

Instructions of the lying aversion test (implemented in a subset of the sample)

Now you can earn additional money. Please read carefully the following description.
Imagine you are in the following situation:

The earnings that you can get depend on the outcome that you report regarding a coin toss.

In the following, there will be five possible situations for each of which you have to make a choice. In each of the situations you will be asked to report the outcome of the coin toss. You can report either HEADS or TAILS.

Once all the participants of the experiment have completed this task, we will randomly pick one of these 5 situations and one participant in this room (randomly selected) will be paid the money corresponding to his/her choice in the selected situation.

Also imagine that the real result of the coin toss is TAILS and that only you have been able to observe it.

What outcome do you report?

(Please choose one alternative in each one of the five situations.)

SITUATION 1:

- O TAILS - In this case, your earnings will be 0.30 EUR
- O HEADS - In this case, your earnings will be 1.50 EUR

SITUATION 2:

- O TAILS - In this case, your earnings will be 0.60 EUR
- O HEADS - In this case, your earnings will be 1.50 EUR

SITUATION 3:

- O TAILS - In this case, your earnings will be 0.90 EUR
- O HEADS - In this case, your earnings will be 1.50 EUR

SITUATION 4:

- O TAILS - In this case, your earnings will be 1.20 EUR
- O HEADS - In this case, your earnings will be 1.50 EUR

SITUATION 5:

- O TAILS - In this case, your earnings will be 1.50 EUR
- O HEADS - In this case, your earnings will be 1.50 EUR

C) HETEROGENEITY

Table 14. Behavior by matching group

		A) Player 1's behavior (freq. of true mess.)					*B) Player 2's behavior (freq. of choices)*					
		All Rounds		Rounds 21-40		P1's message	All Rounds			Rounds 21-40		
Group		P2 not int.	P2 int.	P2 not int.	P2 int.	says that P2 is:	Left	Center	Right	Left	Center	Right
1	P1 not int.	40.00	50.00	60.00	66.67	Not interested	26.19	23.81	50.00	21.74	13.04	65.22
	P1 int.	43.75	23.81	55.56	25.00	Interested	26.32	26.32	47.37	17.65	47.06	35.29
2	P1 not int.	56.00	100.00	45.45	100.00	Not interested	2.13	23.40	74.47	4.76	19.05	76.19
	P1 int.	86.67	13.04	71.43	21.43	Interested	69.70	21.21	9.09	73.68	10.53	15.79
3	P1 not int.	75.00	70.59	80.00	71.43	Not interested	5.26	14.04	80.70	6.45	16.13	77.42
	P1 int.	96.00	27.78	100.00	20.00	Interested	60.87	39.13	0.00	66.67	33.33	0.00
4	P1 not int.	65.00	52.38	45.45	50.00	Not interested	6.38	14.89	78.72	0.00	8.33	91.67
	P1 int.	59.09	35.29	63.64	20.00	Interested	54.55	21.21	24.24	43.75	43.75	12.50
5	P1 not int.	60.00	41.18	57.14	60.00	Not interested	52.08	33.33	14.58	40.91	50.00	9.09
	P1 int.	60.00	38.89	71.43	44.44	Interested	34.38	50.00	15.63	11.11	72.22	16.67
6	P1 not int.	47.37	90.00	55.56	81.82	Not interested	2.17	15.22	82.61	0.00	19.23	80.77
	P1 int.	90.91	21.05	100.00	9.09	Interested	55.88	38.24	5.88	42.86	50.00	7.14
7	P1 not int.	76.92	54.55	58.33	70.00	Not interested	11.76	19.61	68.63	8.70	34.78	56.52
	P1 int.	76.47	46.67	100.00	62.50	Interested	37.93	44.83	17.24	41.18	47.06	11.76
8	P1 not int.	73.68	67.86	70.00	61.54	Not interested	14.63	24.39	60.98	18.18	36.36	45.45
	P1 int.	56.25	47.06	66.67	50.00	Interested	46.15	38.46	15.38	55.56	33.33	11.11
9	P1 not int.	86.36	63.16	83.33	66.67	Not interested	0.00	32.61	67.39	0.00	45.00	55.00
	P1 int.	77.78	71.43	71.43	80.00	Interested	76.47	17.65	5.88	80.00	15.00	5.00
10	P1 not int.	54.55	61.11	55.56	33.33	Not interested	18.00	30.00	52.00	26.67	23.33	50.00
	P1 int.	80.00	25.00	91.67	20.00	Interested	36.67	46.67	16.67	50.00	40.00	10.00
11	P1 not int.	47.83	42.86	53.33	37.50	Not interested	22.64	20.75	56.60	22.22	29.63	48.15
	P1 int.	78.95	11.76	75.00	11.11	Interested	55.56	11.11	33.33	46.15	0.00	53.85
12	P1 not int.	70.00	80.95	55.56	92.31	Not interested	0.00	25.49	74.51	0.00	34.78	65.22
	P1 int.	86.96	18.75	100.00	14.29	Interested	72.41	27.59	0.00	88.24	11.76	0.00
13	P1 not int.	77.78	63.64	75.00	77.78	Not interested	7.41	44.44	48.15	11.11	44.44	44.44
	P1 int.	83.33	22.73	90.00	23.08	Interested	26.92	61.54	11.54	30.77	61.54	7.69
14	P1 not int.	66.67	46.43	55.56	52.94	Not interested	9.43	35.85	54.72	8.33	41.67	50.00
	P1 int.	82.35	29.41	71.43	14.29	Interested	48.15	40.74	11.11	43.75	43.75	12.50
15	P1 not int.	86.67	58.33	75.00	50.00	Not interested	21.15	15.38	63.46	20.83	25.00	54.17
	P1 int.	75.86	41.67	60.00	40.00	Interested	35.71	57.14	7.14	6.25	81.25	12.50
16	P1 not int.	68.00	59.09	64.29	66.67	Not interested	7.55	37.74	54.72	3.85	42.31	53.85
	P1 int.	80.00	16.67	77.78	12.50	Interested	55.56	18.52	25.93	64.29	21.43	14.29
17	P1 not int.	85.71	75.00	100.00	81.82	Not interested	6.38	34.04	59.57	4.00	32.00	64.00
	P1 int.	91.30	57.89	100.00	54.55	Interested	60.61	33.33	6.06	73.33	20.00	6.67
18	P1 not int.	85.71	54.17	100.00	62.50	Not interested	9.09	36.36	54.55	15.63	21.88	62.50
	P1 int.	100.00	52.94	100.00	37.50	Interested	36.00	64.00	0.00	37.50	62.50	0.00
19	P1 not int.	83.33	88.89	92.31	90.00	Not interested	3.51	12.28	84.21	3.57	14.29	82.14
	P1 int.	100.00	17.65	100.00	20.00	Interested	60.87	34.78	4.35	66.67	25.00	8.33
20	P1 not int.	85.71	71.43	88.89	80.00	Not interested	12.50	29.17	58.33	13.64	22.73	63.64
	P1 int.	94.12	61.90	88.89	66.67	Interested	65.63	28.13	6.25	72.22	22.22	5.56
21	P1 not int.	85.19	77.27	91.67	83.33	Not interested	1.82	14.55	83.64	0.00	8.00	92.00
	P1 int.	85.00	9.09	75.00	25.00	Interested	68.00	20.00	12.00	66.67	26.67	6.67
22	P1 not int.	87.50	100.00	100.00	100.00	Not interested	4.00	24.00	72.00	3.85	19.23	76.92
	P1 int.	92.00	31.58	92.31	40.00	Interested	90.00	3.33	6.67	100.00	0.00	0.00
23	P1 not int.	63.16	69.57	80.00	56.25	Not interested	4.35	58.70	36.96	0.00	66.67	33.33
	P1 int.	88.24	42.86	88.89	20.00	Interested	52.94	38.24	8.82	61.54	30.77	7.69
24	P1 not int.	76.47	57.89	100.00	54.55	Not interested	2.08	22.92	75.00	3.57	21.43	75.00
	P1 int.	68.42	44.00	83.33	36.36	Interested	62.50	34.38	3.13	100.00	0.00	0.00
25	P1 not int.	64.71	61.90	57.14	45.45	Not interested	4.08	14.29	81.63	3.57	10.71	85.71
	P1 int.	85.00	40.91	90.91	27.27	Interested	64.52	35.48		66.67	33.33	0.00
26	P1 not int.	46.67	55.56	50.00	47.06	Not interested	2.13	63.83	34.04	4.17	75.00	20.83
	P1 int.	88.24	38.10	80.00	44.44	Interested	39.39	48.48	12.12	18.75	68.75	12.50
27	P1 not int.	38.89	61.90	25.00	61.54	Not interested	0.00	20.00	80.00	0.00	25.00	75.00
	P1 int.	90.91	47.37	66.67	16.67	Interested	71.43	20.00	8.57	65.00	30.00	5.00
28	P1 not int.	18.18	50.00	18.18	12.50	Not interested	13.51	48.65	37.84	17.39	60.87	21.74
	P1 int.	60.00	35.29	54.55	20.00	Interested	41.86	39.53	18.60	23.53	58.82	17.65

		A) Player 1's behavior (freq. of true mess.)					*B) Player 2's behavior (freq. of choices)*					
		All Rounds		Rounds 21-40		P1's message	All Rounds			Rounds 21-40		
Group		P2 not int.	P2 int.	P2 not int.	P2 int.	says that P2 is:	Left	Center	Right	Left	Center	Right
29	P1 not int.	90.00	93.33	100.00	92.86	Not interested	0.00	66.67	33.33	0.00	77.27	22.73
	P1 int.	89.47	19.05	80.00	33.33	Interested	48.57	42.86	8.57	61.11	27.78	11.11
30	P1 not int.	84.62	68.00	100.00	84.62	Not interested	1.82	41.82	56.36	3.70	48.15	48.15
	P1 int.	90.91	20.00	90.00	11.11	Interested	64.00	28.00	8.00	53.85	46.15	0.00
31	P1 not int.	55.00	47.83	53.85	54.55	Not interested	6.00	38.00	56.00	4.17	41.67	54.17
	P1 int.	90.48	50.00	80.00	33.33	Interested	53.33	40.00	6.67	50.00	50.00	0.00
32	P1 not int.	85.00	55.00	83.33	55.56	Not interested	1.85	20.37	77.78	0.00	22.22	77.78
	P1 int.	95.00	55.00	100.00	60.00	Interested	84.62	11.54	3.85	92.31	7.69	0.00
33	P1 not int.	90.91	90.48	100.00	90.00	Not interested	1.79	14.29	83.93	0.00	13.33	86.67
	P1 int.	94.12	10.00	100.00	7.69	Interested	95.83	4.17	0.00	100.00	0.00	0.00
34	P1 not int.	89.47	100.00	100.00	100.00	Not interested	0.00	8.77	91.23	0.00	0.00	100.00
	P1 int.	100.00	0.00	100.00	0.00	Interested	100.00	0.00	0.00	100.00	0.00	0.00
35	P1 not int.	84.62	58.82	72.73	77.78	Not interested	19.61	33.33	47.06	18.18	45.45	36.36
	P1 int.	72.22	52.63	63.64	44.44	Interested	65.52	27.59	6.90	66.67	22.22	11.11
36	P1 not int.	96.15	73.33	100.00	71.43	Not interested	22.03	11.86	66.10	9.09	9.09	81.82
	P1 int.	72.22	19.05	100.00	20.00	Interested	71.43	9.52	19.05	71.43	14.29	14.29
37	P1 not int.	77.78	59.09	77.78	50.00	Not interested	12.96	35.19	51.85	13.79	31.03	55.17
	P1 int.	88.89	31.82	88.89	28.57	Interested	53.85	26.92	19.23	45.45	36.36	18.18
38	P1 not int.	61.90	60.00	75.00	58.33	Not interested	1.89	64.15	33.96	0.00	64.29	35.71
	P1 int.	75.00	10.53	77.78	9.09	Interested	59.26	37.04	3.70	58.33	33.33	8.33
39	P1 not int.	58.82	60.00	55.56	58.33	Not interested	14.89	38.30	46.81	9.52	38.10	52.38
	P1 int.	86.36	52.38	90.91	87.50	Interested	42.42	36.36	21.21	47.37	36.84	15.79
40	P1 not int.	66.67	66.67	58.33	66.67	Not interested	10.00	52.00	38.00	3.57	53.57	42.86
	P1 int.	92.00	57.89	100.00	37.50	Interested	23.33	43.33	33.33	25.00	41.67	33.33
41	P1 not int.	42.31	44.44	50.00	54.55	Not interested	0.00	48.98	51.02	0.00	54.17	45.83
	P1 int.	88.89	22.22	80.00	20.00	Interested	45.16	48.39	6.45	37.50	50.00	12.50
42	P1 not int.	84.21	83.33	81.82	100.00	Not interested	1.96	41.18	56.86	0.00	36.00	64.00
	P1 int.	100.00	37.50	100.00	30.00	Interested	48.28	48.28	3.45	53.33	40.00	6.67
43	P1 not int.	58.82	52.17	57.14	41.67	Not interested	10.20	28.57	61.22	8.00	44.00	48.00
	P1 int.	76.19	36.84	75.00	38.46	Interested	51.61	45.16	3.23	60.00	40.00	0.00
44	P1 not int.	50.00	66.67	58.33	61.54	Not interested	6.82	68.18	25.00	8.70	69.57	21.74
	P1 int.	76.47	35.00	69.23	0.00	Interested	55.56	33.33	11.11	52.94	35.29	11.76
45	P1 not int.	73.91	43.75	81.25	28.57	Not interested	12.50	52.08	35.42	3.85	53.85	42.31
	P1 int.	81.25	64.00	100.00	64.29	Interested	40.63	50.00	9.38	50.00	35.71	14.29
46	P1 not int.	57.14	42.86	50.00	66.67	Not interested	2.44	73.17	24.39	0.00	72.22	27.78
	P1 int.	42.86	54.17	50.00	53.33	Interested	28.21	61.54	10.26	31.82	68.18	0.00
47	P1 not int.	62.50	66.67	75.00	60.00	Not interested	11.11	42.22	46.67	13.64	63.64	22.73
	P1 int.	73.91	50.00	70.00	58.33	Interested	37.14	51.43	11.43	11.11	72.22	16.67
48	P1 not int.	36.84	19.23	40.00	28.57	Not interested	33.33	50.00	16.67	46.15	53.85	0.00
	P1 int.	66.67	20.00	71.43	22.22	Interested	46.15	30.77	23.08	50.00	21.43	28.57
49	P1 not int.	85.00	83.33	90.00	86.67	Not interested	13.51	27.03	59.46	9.52	28.57	61.90
	P1 int.	62.50	70.00	87.50	57.14	Interested	46.51	34.88	18.60	52.63	36.84	10.53
50	P1 not int.	28.57	50.00	37.50	40.00	Not interested	7.84	43.14	49.02	10.71	64.29	25.00
	P1 int.	93.75	26.67	87.50	14.29	Interested	72.41	20.69	6.90	66.67	25.00	8.33
51	P1 not int.	63.16	41.18	60.00	55.56	Not interested	10.71	50.00	39.29	15.38	34.62	50.00
	P1 int.	95.45	40.91	88.89	33.33	Interested	45.83	41.67	12.50	50.00	42.86	7.14

D) ECONOMETRIC ANALYSIS

Table 15. Determinants of the probability that player 1 sends a true message

	All Rounds	Rounds 21-40
Env#Info1	-0.741***	-0.866**
	(0.261)	(0.363)
Env#Info2	-0.304	-0.652**
	(0.220)	(0.306)
Env#Info3	0.267	0.122
	(0.349)	(0.438)
Env#Info4	0.415	0.419
	(0.678)	(0.757)
Prosoc#Info4	-0.509	-0.737
	(0.719)	(0.769)
Prosoc#Envy#Info4	0.0693	-0.249
	(0.763)	(0.873)
Info1	2.176***	2.489***
	(0.717)	(0.725)
Info2	1.710**	2.168***
	(0.735)	(0.819)
Info3	2.570***	2.842***
	(0.760)	(0.832)
Lying_av#Info1	-0.0192	-0.0255
	(0.107)	(0.136)
Lying_av#Info2	0.0119	-0.0324
	(0.0968)	(0.115)
Lying_av#Info3	-0.0114	-0.0409
	(0.106)	(0.140)
Lying_av#Info4	0.185*	0.311***
	(0.100)	(0.0940)
Risk	0.0716	0.121**
	(0.0462)	(0.0515)
Period	-0.00174	-0.00249
	(0.00457)	(0.0122)
Constant	-1.051	-1.237
	(0.702)	(0.950)
Observations	2,691	1,352
Wald chi2(15)	257.09	221.38
Prob > chi2	0.0000	0.0000
Pseudo R2	0.1298	0.1658
Log pseudolikelihood	-1535.6359	-739.46185

***, **, * significant at the 1%, 5%, and 10% level, respectively. With robust standard errors at the matching group level.
Info1 takes value 1 if player 1 is not interested and player 2 is not interested, *Info2* takes value 1 if player 1 is not interested and player 2 is interested, *Info3* takes value 1 if player 1 is interested and player 2 is not interested, and *Info4* takes value 1 if player 1 is interested and player 2 is interested.

E) MIXTURE MODEL ESTIMATION

A strategy for a sender is a vector $(c_{ni,ni}, c_{ni,i}, c_{i,ni}, c_{i,i}) \in \{0,1\}^4$, as defined in Appendix A. For the estimation of the mixture model, let $i = 1, \ldots, N$ index the different players and $k = 1, \ldots, K$ index our types. We assume that a type-k player normally makes a type k decision, but in each period, he makes an error with probability $\varepsilon_k \in [0, 1]$, constituting type k's *error rate*, in which case he chooses to send a true or a false message with equal probability $\frac{1}{2}$. For a type-k player, the probability of a type k decision in any information set is then $1 - \frac{1}{2}\varepsilon_k$. Hence, the probability of a non-type-k decision is $\frac{\varepsilon_k}{2}$. We assume that errors are independently and identically distributed across periods and players and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_K)$.

The likelihood function can be constructed as follows. Let T_i denote the total number of periods in which player i acted as sender. Next, let x_{ik} denote the number of player i's decisions that equal type k's in periods in which he acts as a sender and $x_i = (x_{i1}, \ldots, x_{iK})$, $x = (x_1, \ldots, x_i, \ldots, x_N)$. Let p_k denote the common probability that a player is of type k, $\sum_{k=1}^{K} p_k = 1$ and $p = (p_1, \ldots, p_K)$. As each period has one type-k decision and one non-type-k decision, the probability of observing a particular sample with x_{ik} type-k decisions when player i is type k can be written as follows:

$$L_k^i(\varepsilon_k | x_{ik}) = \left[1 - \frac{1}{2}\varepsilon_k\right]^{x_{ik}} \left[\frac{1}{2}\varepsilon_k\right]^{T_i - x_{ik}}$$

Weighting the right-hand side by p_k, summing over k, taking logarithms, and summing over i yields the log-likelihood function for the entire sample:

$$lnL(p, \varepsilon | x) = \sum_{i=1}^{N} \ln \sum_{k=1}^{K} p_k L_k^i(\varepsilon_k | x_{ik})$$

This function is maximized by the EM algorithm, as proposed in the seminal paper by Dempster, Laird and Rubin, 1977 (Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, *39*(1), 1-38).

In Table 16 we report the results of the mixture model estimation if we exclude either type (iv) or type (i) from the specification discussed in the text (*cf.* Table 9). The results are robust. The log likelihoods obtained (-968.018 and -992.414, respectively) are lower than the one obtained from the specification including both types (-899.755). Thus, given that the specification reported in Table 9 only requires to estimate two additional parameters with respect to these alternatives, it is preferred according to the Akaike information criterion.

Table 16. Mixture model without types (i) or (iv)

	Without (iv)		Without (i)	
Type	*Probab.*	*Error rate*	*Probab.*	*Error rate*
(i) Both social preferences and lying aversion	0.267*** (0.045)	0.029*** (0.032)		
(ii) Only lying aversion	0.156*** (0.044)	0.695 (0.213)	0.161*** (0.047)	0.646** (0.180)
(iii) Only social preferences	0.151*** (0.041)	0.313*** (0.165)	0.181*** (0.042)	0.232*** (0.122)
(iv) Both soc. pref. and lying av. with 'reduced envy'			0.328*** (0.058)	0.187*** (0.081)
(v) Standard preferences	0.426*** (0.055)	0.174*** (0.057)	0.330*** (0.064)	0.078*** (0.088)
Log Likelihood	– 968.016		– 992.414	

Standard errors within brackets
Significance: Probabilities (with respect to 0) and Error rates (with respect to 1): ***, **, * at the 1%, 5%, 10% level, resp.

F) COMPARISON OF CHOICES IN 2014 AND 2021-23

In the following, we compare the behavior of the sender and the receiver across the sessions run in 2014 and the sessions run in 2021-23. In Table 17 we consider the sender and report the frequencies of true messages in all rounds and in the last 20 rounds. Although we observe some significant differences when the sender is interested, they disappear in the second half of the experiment.

Table 17. Sender comparison (sessions 2014 – sessions 2021-23)

		All rounds		Rounds 21-40	
		P2 not interested	*P2 interested*	*P2 not interested*	*P2 interested*
Sessions 2021-2023	*P1 not int*	68.84	63.25	70.72	64.49
	P1 interested	84.61**	36.6	84.38	32.66
Sessions 2014	*P1 not int*	63.16	64.97	61.11	65.93
	P1 interested	74.19	34.39	80.85	35.51

***, **, * significant at the 1%, 5%, and 10% level, respectively, at two-sample t test.

In Table 18 we examine then the behavior of the receiver. We report the frequencies of actions *Left*, *Center* and *Right* in the sessions run in 2014 and the sessions run in 2021-23. We observe that, although there exist some differences when the message says *Not-Interested*, they basically disappear in the last half of the experiment (only the difference in the frequency of *Center* remains marginally significant).

Table 18. Receiver comparison (sessions 2014 – sessions 2021-23)

		All Rounds			Rounds 21 – 40		
		Left	Center	Right	Left	Center	Right
Sessions 2021-2023	Message NoInt	8.22*	36.04**	55.74	7.66	38.31*	54.02
	Message Int	54.74	34.80	10.46	55.03	35.23	9.73
Sessions 2014	Message NoInt	13.68	22.95	63.37	12.81	25.62	61.57
	Message Int	49.54	33.85	16.62	48.10	38.61	13.29

***, **, * significant at the 1%, 5%, and 10% level, respectively, at two-sample t test.

G) EQUILIBRIUM ANALYSIS WITH SOCIAL PREFERENCES AND LYING AVERSION FOR THE HOMOGENEOUS CASE

In the following analysis we consider the homogeneous scenario with social preferences and lying aversion, in which a player's utility –equation (1)– becomes:

$$U_i(\pi_i, \pi_j) = v_i(\pi_i, \pi_j) - I_i \cdot L,$$

with

$$v_i(\pi_i, \pi_j) = \pi_i - \alpha \cdot max\{\pi_j - \pi_i, 0\} - \beta \cdot max\{\pi_i - \pi_j, 0\},$$

where $0 \leq \beta < 1$, $\beta \leq \alpha$, $L \geq 0$, π_i and π_j are the monetary payoffs of, respectively, individuals i and j, and I_i is an indicator function taking the value 1 when the individual sends a false message (0 otherwise). Table 19 below shows the value of $v_i(\pi_i, \pi_j)$, i.e., the utilities absent of lying aversion costs, for all possible outcomes.

Table 19. Utilities with social preferences

Player 2's choice	Left	Center	Right	Left	Center	Right
Player 1 not interested	20 20	$20-40\alpha$ $60-40\beta$	$20-80\alpha$ $100-80\beta$	$20-140\alpha$ $160-140\beta$	$20-100\alpha$ $120-100\beta$	$20-30\alpha$ $50-30\beta$
Player 1 Interested	$20-10\alpha$ $30-10\beta$	$70-20\alpha$ $90-20\beta$	120 120	$20-40\alpha$ $60-40\beta$	$70-20\beta$ $50-20\alpha$	$120-80\beta$ $40-80\alpha$
	Player 2 not interested			*Player 2 interested*		

Under these assumptions, we have the following set of equilibria (notes: in the first column we report the equilibrium strategies and in the second column we report the conditions under which such strategies constitute equilibria. The equilibria have been characterized using the software *Mathematica*. The calculus is available from the authors upon request).

1. *Informative equilibria*:

$((1,1,1,0),(R,L))$	$0 \leq \alpha \leq 1$ $0 \leq \beta \leq min\left\{\frac{1}{8}(10-7\alpha), 3-3\alpha\right\}$ $110\alpha \leq L \leq 100+40\alpha-80\beta$
$((1,1,1,0),(L,L))$ $((0,0,0,1),(L,L))$	$4 < \alpha$ $9-2\alpha \leq \beta$ $L = 0$
$((1,1,1,0),(C,L))$	$\frac{3}{4} \leq \alpha \leq \frac{9}{2}$ $3-3\alpha \leq \beta \leq 9-2\alpha$ $40\alpha \leq L \leq 50+40\alpha-20\beta$
$((0,0,0,1),(L,R))$	$\alpha = 0$ $\beta = 0$ $L = 0$

2. *Babbling equilibria*:

$((1,0,1,0),(C,C))$ $((0,1,0,1),(C,C))$	$0 \leq \alpha < 4$ $\beta \geq \frac{2\alpha-5}{3}$ $L = 0$
$((1,0,1,0),(C,L))$	$0 \leq \alpha < 4$ $\beta \geq \frac{2\alpha-5}{3}$ $L = 40\alpha$
$((0,1,0,1),(L,C))$	$\alpha = 0$ $\beta = 0$ $L = 0$
$((1,0,1,0),(L,L))$ $((0,1,0,1),(L,L))$	$\frac{5}{2} \leq \alpha$ $\beta \leq \frac{2\alpha-5}{3}$ $L = 0$

3. *Opposite equilibria*:

$((0,1,1,0),(R,L))$ $((0,1,1,0),(R,C))$ $((1,0,0,1),(L,R))$ $((1,0,0,1),(C,R))$	$\alpha = 0$ $\beta = 0$ $L = 0$
$((0,1,1,0),(L,L))$ $((1,0,0,1),(L,L))$	$\frac{5}{2} \leq \alpha$ $\beta \leq \frac{2\alpha-5}{3}$ $L = 0$
$((0,1,1,0),(C,L))$	$\frac{1}{3} \leq \alpha < 4$ $\frac{2\alpha-5}{3} \leq \beta \leq 3\alpha - 1$ $L = 40\alpha$
$((0,1,1,0),(C,C))$ $((1,0,0,1),(C,C))$	$\frac{1}{3} \leq \alpha < 4$ $\frac{2\alpha-5}{3} \leq \beta \leq 3\alpha - 1$ $L = 0$

4. *Full revelation*:

$((1,1,1,1),(R,L))$	$0 \leq \alpha$ $0 \leq \beta \leq \frac{1}{8}(5+2\alpha)$ $L \geq \max\{100 + 40\alpha - 80\beta, 110\alpha\}$
$((1,1,1,1),(R,C))$	$\frac{5}{6} \leq \alpha < \frac{3}{2}$ $\frac{1}{8}(5+2\alpha) \leq \beta$ $L \geq 70\alpha$

5. *Other equilibria*:

$((1,0,1,1),(R,L))$	$\frac{8}{11} \leq \alpha$ $\frac{10-7\alpha}{8} \leq \beta \leq \frac{1+2\alpha}{4}$ $max\{100+40\alpha-80\beta, 80\alpha\} \leq L \leq 110\alpha$
$((1,0,1,1),(R,C))$	$\frac{1}{2} \leq \alpha < \frac{3}{2}$ $\frac{1+2\alpha}{4} \leq \beta$ $40\alpha \leq L \leq 70\alpha$
$((1,1,0,0),(C,C))$ $((0,0,1,1),(C,C))$	$\frac{3}{10} \leq \alpha \leq \frac{3}{2}$ $\frac{3-6\alpha}{4} \leq \beta \leq \frac{4}{9}$ $L = 0$
$((1,1,0,0),(C,R))$	$\frac{5}{6} \leq \alpha \leq \frac{3}{2}$ $\frac{5}{6} \leq \beta$ $L \leq 60\beta - 50$
$((0,0,1,1),(R,L))$	$\frac{3}{2} \leq \alpha$ $\beta \geq max\left\{\frac{5-2\alpha}{4}, \frac{4}{9}\right\}$ $100+40\alpha-80\beta \leq L \leq 80\alpha$
$((0,0,1,1),(R,C))$	$\frac{1}{2} \leq \alpha \leq \frac{3}{2}$ $\beta \geq max\left\{\frac{5-2\alpha}{4}, \frac{4}{9}\right\}$ $50-60\beta \leq L \leq 40\alpha$

Figure 2 focuses on the informative equilibria and illustrates their robustness to the introduction of lying costs. In particular, it presents region plots of the ranges of parameters (α, β) in which the informative equilibrium exists, considering values $L \in \{0, 20, 40, 60, 80, 100, 120, 140\}$.

In Figure 2 we use different colors (grey intensities) to refer to different informative equilibria that exist with social preferences regarding the receiver's response (*cf.* Footnote 14). For example, the darkest (black) areas in the plots of Figure 2 corresponds to the informative equilibrium that already exists without social preferences (in which the receiver's response is *left* if the report says that he is interested and *right* otherwise). This equilibrium exists only for $(\alpha, \beta) = (0,0)$ in case $L = 0$ and, as we can see, the size of the black area initially increases in L, then decreases and, eventually, the area vanishes when L is large enough.

Figure 2. Robustness of the informative equilibria to the introduction of lying costs