Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams* — Source link

Matthias Sutter

Institutions: University of Innsbruck

Published on: 01 Jan 2009 - The Economic Journal (Blackwell Publishing Ltd)

Topics: Deception

Sutter, Matthias

Working Paper

Deception through telling the truth?!: Experimental evidence from individuals and teams

Working Papers in Economics and Statistics, No. 2007-26

Provided in Cooperation with:
Institute of Public Finance, University of Innsbruck

University of Innsbruck

Working Papers
in
Economics and Statistics

Deception through telling the truth?!
Experimental evidence from individuals and teams

Matthias Sutter

2007-26

Deception through telling the truth?!

Experimental evidence from individuals and teams*

Matthias Sutter#

Abstract. Informational asymmetries abound in economic decision making and often provide an incentive for deception through telling a lie or misrepresenting information. In this paper I use a cheap-talk sender-receiver experiment to show that telling the truth should be classified as deception too if the sender chooses the true message with the expectation that the receiver will not follow the sender's (true) message. The experimental data reveal a large degree of 'sophisticated' deception through telling the truth. The robustness of my broader definition of deception is confirmed in an experimental treatment where teams make decisions.

JEL-classification: C72, C91, D82

Keywords: Deception, Expectations, Team decision making, Individual decision making, Experiment

This version: 23rd October 2007

* I would like to thank Vincent Crawford, Martin Dufwenberg, Uri Gneezy, Navin Kartik, Martin Kocher, Stephan Kroll, Topi Miettinen and Jörg Oechssler for very helpful comments. Leonardo Felli and three referees provided excellent advice on how to improve the paper. Olga Mellizo was of great help in language editing. Financial support from the Austrian Science Foundation (FWF-Project No. P16617), the Max Planck Society, the Center of Experimental Economics at the University of Innsbruck (sponsored by *Raiffeisen Landesbank Tirol*), the Tyrolean Science Foundation (TWF) and the German Science Foundation (through the Leibniz-Award to Axel Ockenfels) is gratefully acknowledged.

University of Innsbruck and University of Göteborg. Address for correspondence: University of Innsbruck, Department of Public Finance, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. e-mail: matthias.sutter@uibk.ac.at

1 Introduction

Many real-life situations are characterized by informational asymmetries between interacting parties. Obviously, such situations may provide an incentive for either party to exploit the informational asymmetries to one's own advantage. This may then imply the use of deception. Think of opposing parties in war when one party pretends to invade the other's territory at a different location than where the attack is actually carried out (as it happened when the Allied Forces invaded Normandy in 1944, when they incurred substantial costs to pretend to Germans to invade in Calais; see Crawford, 2003, for a short account), think of managers concealing how to restructure or downsize a company that shall be taken over, think of electoral competition where spreading false facts about the opponents' plans or hiding one's own plans may gain additional votes (see Heidhues and Lagerlof, 2003), think of football goalkeepers who falsely signal in a penalty situation to jump into a particular corner in order to induce the player to shoot to the other corner (Chiappori et al., 2002), or think of poker players bluffing about the quality of their cards (Charness and Dufwenberg, 2005). In all such situations, and many more, sending wrong information may ultimately pay off for the sender.

A unifying assumption of the literature on strategic information transmission in games with informational asymmetries is that deception is typically meant to involve any sort of lying or misrepresentation of information (Crawford and Sobel, 1982; Vrij, 2001). According to this standard presumption telling the truth cannot be considered as an attempt to deceive others. Or should it? In this paper I use a cheap-talk sender-receiver experiment to challenge this narrow view and show that telling the truth should be classified as intended deception[1] if the sender chooses the true message with the expectation that the receiver will not follow the

[1] In the following I will focus on the intention to deceive, even if I speak of deception. Note, however, that one can make a distinction between the *intention to deceive* and *deception*. The latter refers to a successful attempt at misleading the receiver, whereas the former concerns intentions alone – irrespective of the actual outcome of an interaction.

sender's (true) message. This sort of 'sophisticated' deception through telling the truth is, in fact, frequently observed both when individuals and teams make decisions. Hence, this paper shall create an awareness that any attempt to measure deception in sender-receiver games as the frequency with which the sender sends the wrong message about the consequences of a given action for the receiver will fail to capture the full amount of deception. Likewise, interpreting truth-telling as an act of honesty will be misleading in many cases.

Recently, deception has caught quite some interest in economics, because deception is part of many interactions, for instance when managers, workers, politicians or consumers lie for strategic advantage and make use of private information in a dishonest way (Crawford, 2003). If there is a conflict of interest, then game theory would predict that zero-cost messages are not informative in such situations, and should be disregarded (Crawford and Sobel, 1982). However, Crawford (2003) has shown why such messages are nevertheless sent and are, in fact, influential for behaviour. If there are two types of players (sophisticated ones and boundedly rational ones) then deception can be used by the former type to exploit the latter type. Whereas this approach assumes that there is no negative outcome associated with lying per se, recent experimental research has challenged this assumption. Since the presence of informational asymmetries and the exploitation thereof is much more difficult to control with field data than in a controlled laboratory setting, experimental studies have abound in recent years. Gneezy (2005) has shown in his experiment that the willingness of subjects to lie to others is a matter of weighing costs and benefits. Subjects deceive others more often the higher the possible gains from deception and the lower the associated losses for the deceived subjects. Deception has been defined and measured by Gneezy (2005) as the relative frequency with which the sender in a cheap-talk sender-receiver game sends the wrong message about the consequences of a given action for the receiver.[2] This narrow definition has been also used in subsequent papers by Charness and Dufwenberg (2005), Hurkens andKartik

[2] In fact, Gneezy (2005) discusses also several other definitions of deception and lies. Yet, when analyzing his data he uses the definition at the top of p. 386 that implies the definition to which I refer to.

(2006), Rode (2006) or Sánchez-Pagés and Vorsatz (2007). Yet, in this paper I argue that the narrow definition is not appropriate for measuring the full amount of deception.

Though my experimental data corroborate the main findings of Gneezy (2005) and Hurkens and Kartik (2006) with respect to the influence of costs and benefits on the frequency of deception, my results also indicate that they report only a lower bound of the frequency of deception in strategic interaction. By taking into account what I call 'sophisticated' deception through telling the truth, the frequency of deception increases by about 20 to 30 percentage points. Given this substantial increase, it seems warranted not to confine deception to simply telling a lie.

The rest of the paper is organized as follows: Section 2 introduces the experimental design and procedure. Section 3 presents the experimental results with individuals as decision makers. Section 4 reports on a robustness test of my classification of deception by observing the discussion in teams before teams send a message. The analysis of the videotaped discussions allows tracking down the true motives for a given pattern of behaviour[3] and it reveals that the vast majority of teams that send the true message actually intends to deceive the receiver team. Section 5 concludes the paper.

2 Experimental design

The cheap-talk game. There are two players of which only the sender is informed about the monetary consequences of two different options, A and B. The sender's only choice in the game is to send one of the following two messages to the receiver:

Message A: "Option A will earn you more money than Option B."

Message B: "Option B will earn you more money than Option A."

[3] The method of video-experiments has been applied successfully to uncover, for instance, the main motives for bargaining behaviour in a so-called power-to-take experiment (see Bosman et al., 2006).

Knowing the sent message, the receiver has to pick one of the two options which is then implemented for payment. Note that after his decision, the receiver is only informed about the monetary payoff from the chosen option, but not about the payoff in the other option, nor about the sender's payoff in any of the options. Hence, the receiver cannot judge whether the sender has told the truth or not.

Procedure. Table 1 summarizes the 3 different treatments (T1, T2, T3) used in this experiment (and also in Gneezy, 2005).[4] Treatments differ with respect to the relative gains for the sender and the relative losses for the receiver if Option B instead of Option A is implemented. Note that Option B is always more favourable for the sender, while Option A is always more favourable for the receiver.

TABLE 1 – PAYOFFS (IN €) IN THE DIFFERENT TREATMENTS

Treatment	Option	Payoff to Sender	Payoff to Receiver
T1	A	5	6
	B	6	5
T2	A	5	15
	B	6	5
T3	A	5	15
	B	15	5

At the beginning of the experimental sessions subjects received written instructions (see the Appendix) and clarifying questions were answered privately by the experimenter. After

[4] For the main purpose of this paper – i.e., to examine under which conditions telling the truth may be classified as deception – it might have been sufficient to pick just one of Gneezy's (2005) original treatment. However, since the frequency of lying has been found by Gneezy (2005) to depend on the costs and benefits to the sender and/or the receiver, the same might hold true for the relative frequency of deception through telling the truth. This possibility can only be checked by running all three treatments.

senders had made their decisions and while receivers had time to consider their decision, senders were asked to answer the following two questions in turn.[5]

Question Q1: "Which option do you expect the receiver to choose?"

Question Q2: "Out of 100 receivers, how many do you think follow the sender's message on average?"

Experimental sessions were run from June to August 2005 at the Max Planck Institute of Economics in Jena. A total of 570 participants yielded 96 sender-receiver-pairs each in treatments T1 and T2, and 93 pairs in T3. Recruitment was done via ORSEE (Greiner, 2004), and sessions were computerized using zTree (Fischbacher, 2007). Average session length was 25 minutes.

3 Experimental results

Table 2 reports the results of the experiment. Panel A presents unconditional relative frequencies, whereas Panel B interacts the senders' messages with their answers to question Q1. The figures in row [1] of Panel A indicate the relative frequency of lying, i.e. sending the incorrect Message B. It is lowest in treatment T2 – where it costs the receiver 10 € but gains the sender only 1 € more – intermediate in T1, and highest in T3 – where the sender can gain 10 € more at the same cost for the receiver. Considering all three treatments, the frequencies of sending Message B are significantly different across treatments ($\chi^2 > 10$, $d.f. = 2$, $p < 0.01$). However, there is no significant difference between treatments T1 and T2 in my experiment (even though the overall averages point in the same direction as in Gneezy, 2005). Hence, it seems that in my subject pool the treatment differences with respect to the sender's payoffs in

[5] Answers to these questions were not monetarily incentivized. There is evidence that eliciting expectations with or without monetary rewards for accuracy does not yield significantly different results (see Grether, 1992, or Camerer and Hogarth, 1999).

Option B (compare T3 vs. T1 and T2) are relatively more important than the treatment differences with respect to the receiver's payoffs in Option B (compare T1 vs. T2 and T3).

Row [2] of Table 2 reports the answers of senders to question Q1. It shows that across treatments senders expect the receiver to implement the option that is indicated in the chosen message as yielding the higher payoffs in about 70% of cases. There is no significant difference in these expectations across treatments. Row [3] of Table 2 then shows the relative frequency of receivers actually following the sender's message, and there is again no significant treatment effect.[6] Across treatments receivers follow the sender's message in 72% of cases. Thus, senders' expectations match receivers' actions remarkably well in the aggregate.[7]

TABLE 2 – EXPERIMENTAL RESULTS (RELATIVE FREQUENCIES)

Panel A Treatment	T1	T2	T3
	$N = 96$	$N = 96$	$N = 93$
[1] Sender chooses Message B	0.44	0.35	0.59
[2] Q1: Sender expects the receiver to implement chosen message	0.73	0.70	0.66
[3] Receiver implements message	0.67	0.75	0.74
[4] Sender expects receiver to choose Option B	0.67	0.61	0.89
Panel B. Sender's message and expected option (in Q1)			
[A] Message A and Option A (*Benevolent truth-teller*)	0.31	0.36	0.09
[B] Message A and Option B (*Sophisticated truth-teller*)	0.25	0.28	0.32
[C] Message B and Option B (*Liar*)	0.41	0.33	0.56
[D] Message B and Option A (*Benevolent liar*)	0.02	0.02	0.03

[6] Note that receivers had no information about the possible monetary payoffs in any of the treatments. Rather, receivers had exactly the same information conditions in all treatments. Hence, it would have been discomforting to find significant differences between treatments concerning the receivers' frequency to implement the received message.

[7] It is noteworthy that senders and receivers seem to be very well locally calibrated both in Gneezy's (2005) experiment in Israel as well as mine in Germany. Although the relative frequency of senders expecting the receiver to follow the message differs slightly (about 80% in Gneezy's experiment, and 70% here), both populations are very well calibrated as the fraction of receivers following the message is very close to (and statistically not significantly different from) the senders' expectations.

The details with respect to the sender's chosen message, however, matter significantly for the corresponding expectations, as is shown in Panel B of Table 2 and illustrated in Figure 1. The grey bars in Figure 1 indicate the relative frequency with which senders expect the receivers to implement their message (question Q1) *if* the sender chooses the *true* Message A. On the contrary, the black bars are based on senders who send the *wrong* Message B. Obviously, there is a marked difference in expectations, depending on the actually chosen message.

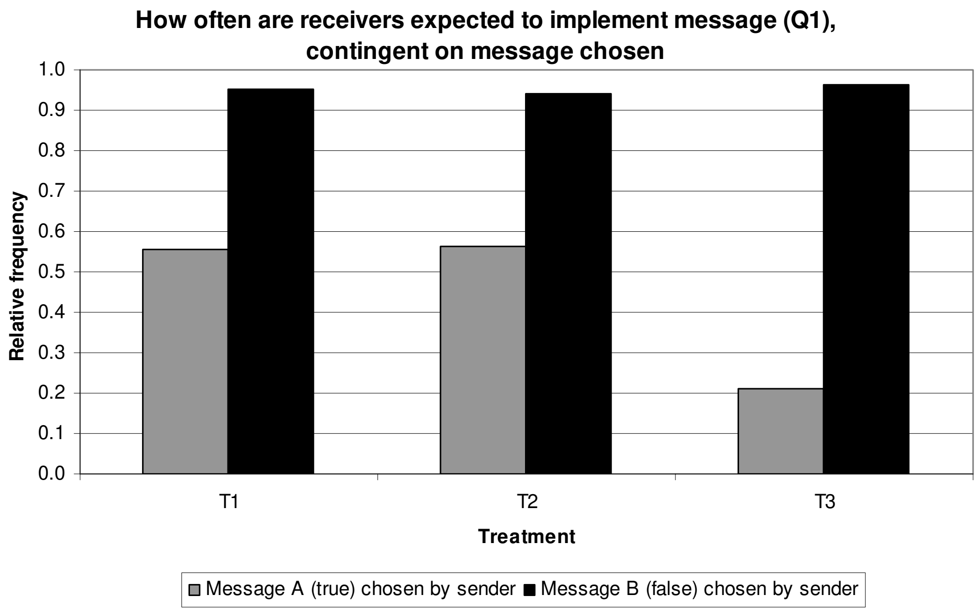

FIGURE 1 – EXPECTATIONS OF SENDERS (Q1) DEPENDING UPON CHOSEN MESSAGE

When sending the false Message B, senders expect in about 95% of cases the receiver to implement the chosen message (black bars in Figure 1).[8] If senders choose the true Message A, then they expect receivers to implement Option A in only 56% of cases both in treatments T1 and T2, and in only 21% (!) of cases in treatment T3 (grey bars). The latter value is

[8] These relative frequencies can also be derived by dividing the entries in row [C] of Table 2 by the sum of entries in rows [C] and [D]. Note that entries in Table 2 are rounded to the second decimal, which might lead to small differences in calculating the relative frequencies that are analogous to those shown in Figure 1.

particularly noteworthy as it is significantly smaller than the values in treatments T1 and T2 ($\chi^2 > 10$ in both comparisons, *d.f.* = 1, $p < 0.01$). One plausible explanation for this finding might be the following: Senders that expect their message to be implemented have more to lose (10 €) in treatment T3 than in treatments T1 and T2 (1 €). This should make such senders less likely to send the true Message A, as it actually does (see row [A] in Table 2). Consequently, Message A is also chosen less often (compare the sums of rows [A] and [B] in Table 2 across treatments). This results in a lower conditional frequency of expecting the message to be implemented contingent on sending the true Message A in treatment T3 (0.21) compared to T1 (0.56) and T2 (0.56).[9]

In Table 2 I have called those senders who send Message A and expect the receiver to implement Option A "*benevolent truth-tellers*". Senders who send Message A, but expect their receiver to implement Option B are called "*sophisticated truth-tellers*". There is a strong and significant difference between benevolent and sophisticated truth-tellers in each treatment concerning their answers to question Q2 (Mann Whitney U-test, $p < 0.05$ in each treatment). The latter question asked for the average relative frequency with which receivers follow the sender's message, i.e. implement the option that is said to be better in the chosen message.

One can see from the grey bars in Figure 2 that *benevolent truth-telling* expect about two thirds of receivers to implement the suggested option (which is significantly larger than 50%; binomial test, $p < 0.05$). The black bars reveal that *sophisticated truth-telling* expect only about one third of receivers to implement the suggested option (which is significantly smaller than 50%; binomial test, $p < 0.05$).[10]

[9] Another possible explanation for the differences in contingent expectations might be what is known in the psychology literature as desirability bias or wishful thinking (Hogarth, 1987), which states that the estimated probability of a desirable future event is positively correlated to the importance of the desirable event to the forecaster. Since senders in treatment T3 can gain 10 € instead of 1 € in treatments T1 and T2, this might induce a wishful thinking bias that makes truthful senders in T3 much more "pessimistic" about the frequency with which receivers follow their message.

[10] It might be interesting to note that the answers to questions Q1 and Q2 are very well, if not perfectly, related on the individual level. Across all treatments, 78% (92%) of senders who expect their message not to be implemented (Q1) expect strictly less than 50% (up to 50%) of the responder population to follow the senders'

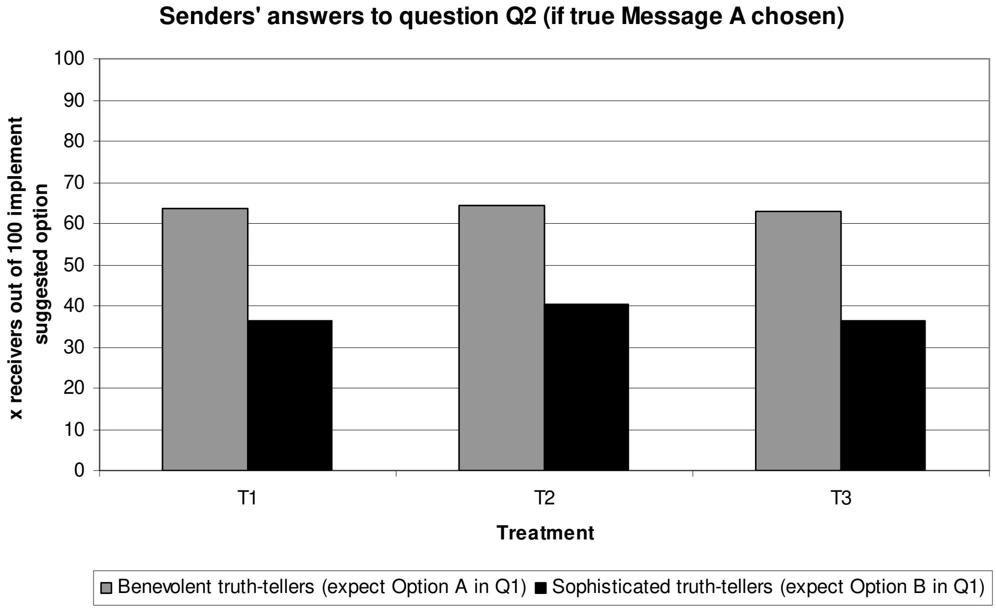

FIGURE 2 – TYPES OF TRUTHFUL SENDERS AND EXPECTATIONS (Q2)

If I regard the *sophisticated truth-tellers* as subjects who actually intend to deceive the receiver (by sending the true Message A!), then it seems that the frequency of senders picking the wrong Message B is not an adequate measure of the total frequency of deception in this game. Therefore, I propose the following

Broader definition of deception: *Deception includes all cases where a sender sends either of the two messages, but expects Option B to be implemented by the receiver.*

This definition then encompasses practically all senders with Message B (of which about 95% expect the receiver to implement Option B), but additionally those senders with Message A who expect the receiver to implement Option B. Applying this broader definition I arrive at the figures in row [4] of Table 2. It shows that 67% of senders in treatment T1 expect Option B to be chosen by the receiver. The corresponding figures are 61% in T2, and 89% in T3. The share of senders expecting to get Option B is significantly larger in treatment T3 (with gains

messages (Q2). 86% (96%) of senders expecting their message to be implemented (Q1) guess that strictly more than 50% (at least 50%) of responders implement the senders' messages.

of 10 €) than in the other treatments T1 and T2 (with gains of 1 € only) ($\chi^2 > 10$ in both comparisons, *d.f.* = 1, $p < 0.01$). There is again no significant difference between treatments T1 and T2, confirming the earlier result (from row [1]) that the differences in own payoffs across treatments seem to have a stronger (and the only statistically significant) impact than the differences in receivers' payoffs across treatments.

Subtracting the figures in row [1] of Table 2 (narrow definition of deception) from those in row [4] (broader definition) shows that the broader definition of deception increases the relative frequency of deception by about 20 to 30 percentage points. Hence, relying solely on row [1] as an indicator of deception would have largely underestimated its extent.

4 A robustness test of the broader definition of deception – An experiment with teams as decision makers

My suggested measure for the overall frequency of deception hinges critically on the assumption that *sophisticated truth-tellers* (who send Message A, but expect Option B to be implemented) really intend to deceive and exploit the receiver. There are three possible alternative explanations, though, that might lead to the same behaviour, but would not be an indication of deception. First, there might be senders who tell the truth (Message A) regardless of what they expect, because telling the truth is the "moral choice" that is separated from expectations. Second, there might be senders who, in fact, intended to benefit the receiver (wishing that he implemented Option A), but expected the receiver to misread their true intention. If such senders were averse to lying (i.e. sending Message B) they would stick to Message A even though they expected the receiver to invert the message.[11] Third, there might be some senders who expected the receiver to choose a particular option independent of

[11] If such senders did not care about lying, yet expected their message to be inverted and wanted to benefit the receiver, then they should send Message B. However, it should be clear from Figure 1 that less than 5% of senders who send Message B expect their message to be inverted. Hence, the fraction of *benevolent liars* (notation introduced in Table 2) is very small.

the sender's message.[12] Hence, if a sender expected the receiver to choose Option B irrespective of the sent message, it would be unreasonable to interpret the combination of "Send Message A and expect Option B" as deception.

In order to discriminate between these alternative explanations and my broader definition of deception I ran additional sessions with 342 subjects in late August 2005. These participants were randomly assigned to *teams* of two or three subjects each, with the same payoffs *per capita* in a given treatment as in the sessions with individual decision makers. I obtained 22 pairs of sender and receiver teams for each of the treatments T1 and T2, and 24 such pairs for treatment T3. The sessions were conducted in the VideoLab of the Max Planck Institute of Economics in Jena. The discussions of sender teams were videotaped in order to detect the true intention behind sending Message A.[13]

Table 3 summarizes the results of the experiment with teams as decision makers. Row [1] in Panel A indicates that teams send Message B significantly less often than individual senders (see Table 2), namely in only 23% of cases in T1 and T2, and in 25% in T3 ($p < 0.05$ in any comparison between teams and individuals for a given treatment). Thus, at first sight it seems as if teams were the more "honest" decision makers.

Yet, the fact that teams send the wrong Message B less often than individuals is driven by different expectations about the receiver team's decision: Averaging over the figures in row [2] shows that sender teams expect the receiver team to follow their message in only 44% of cases. This relative frequency is significantly smaller than the corresponding 70% of individual senders ($\chi^2 > 10$, $p < 0.01$), and not significantly different from a random draw (binomial test). Hence, teams are far more pessimistic concerning the likelihood of receivers

[12] Using the strategy method, Hurkens and Kartik (2006) show that about one third of their receivers ignored the received message, but implemented a particular option that was independent of the message. Yet, it is important to stress that only one out of 30 senders expected his message to be ignored. Hence, almost all senders believe that their message has a systematic influence on the option finally implemented by the receiver.

[13] Sessions with individuals had also been run in the VideoLab of the Max Planck Institute of Economics in order to guarantee the same local environment for individuals and teams. Upon entering the VideoLab, participants were routinely informed that they might be videotaped, but that the tapes would only be used for scientific purposes. In fact, only sender cabins in the team condition were videotaped.

to follow the senders' message. In other words, senders think that receivers are less trusting or that receivers regard the senders as less trustworthy than this is the case with individual decision makers.

TABLE 3 – DECISIONS OF TEAMS (RELATIVE FREQUENCIES)

Panel A	Treatment	T1	T2	T3
		N = 22	*N* = 22	*N* = 24
[1] Sender team chooses Message B		0.23	0.23	0.25
[2] Q1: Sender team expects receiver team to implement chosen message		0.55	0.50	0.29
[3] Receiver team implements message		0.64	0.45	0.67
[4] Sender team expects receiver team to choose Option B		0.68	0.73	0.88
Panel B. Sender team's message and expected option (in Q1)				
[A] Message A and Option A (*Benevolent truth-teller*)		0.32	0.27	0.08
[B] Message A and Option B (*Sophisticated truth-teller*)		0.45	0.50	0.67
[C] Message B and Option B (*Liar*)		0.23	0.23	0.21
[D] Message B and Option A (*Benevolent liar*)		0.00	0.00	0.04

Concerning receivers' actual actions we see from row [3] of Table 3 that receiver teams follow the message on average in 59% of cases. This value is significantly smaller than the 72% of individual receivers ($\chi^2 = 4.4$, $p < 0.05$), which indicates that receiver teams are, in fact, more sceptical about the truthfulness of the sender's message than individual receivers.

Matching messages and expectations we find that all teams, except one, that send the wrong Message B expect the receiver team to implement Option B. This means that those teams sending a lie are pretty confident that the lie will work out. When sending Message A, however, sender teams expect the receiver team to implement Option A only in 15 out of 52 cases (i.e. 29%). That means that in the other 37 cases (i.e. 71%) with Message A, the sender team expects the receiver team to choose Option B. The crucial question is whether these teams intend to deceive the receiver in order to get more money (as implied by my broader definition of deception) or whether the behaviour of these teams has to be subsumed under any of the above discussed three alternative explanations.

The videotapes of the 37 sender teams that sent Message A, but expected Option B, were scrutinized with respect to the driving motivation for the decision to send Message A. The communication of team members reveals that 34 out of these 37 sender teams (i.e. 92%) explicitly mention the following argument: by sending the true Message A the sender team will earn more money because the receiver team will not believe the received message, but rather invert it and implement Option B.[14]

Consequently, sending Message A and expecting Message B can almost always be classified as an attempt to deceive the receiver team. In fact, only three out of 37 sender teams sent Message A because they wanted to avoid lying, without any recourse to the desirability of the receiver team inverting the message. None of the sender teams ever mentioned that the receiver team might take a particular option completely independent of the received message, ruling out the third alternative explanation of the pattern "Send Message A and expect Option B".

Although individual and team senders differ quite a lot with respect to sophisticated truth-telling, i.e., with respect to the frequency of sending the wrong Message B or the expectations about the receiver's actions (compare Table 2 and Table 3 with respect to rows [1] and [2]), individual and team senders do not differ when my broader definition of deception is applied. Checking the frequency of sender teams expecting Option B to be chosen reveals that there are no longer any significant differences between individuals and teams (compare the entries in row [4] of Tables 2 and 3). Hence, the overall amount of deception in a given treatment

[14] Interestingly, some sender teams discussed that their strategic reasoning might lead to an infinite number of reasoning steps. If the receiver team expected the sender team to send the true message because the sender team expected the receiver team not to believe the message, then the receiver team should in fact implement the received message. That, in turn, would make it more profitable for the sender team to send the wrong Message B instead of the true Message A. Yet, if sender teams expected receiver teams to anticipate this move and therefore invert the message, then sender teams would again prefer the true Message A, and so on. All teams discussing this possibility ended up in sending Message A.

does not depend on the type of decision maker being an individual or a team. Moreover, it is again treatment T3 – where senders can gain the most from receiving Option B instead of Option A – in which sender teams expect the receiver team to choose Option B most often (i.e. in 88% of cases; with $p < 0.05$ compared to T1 or T2). Hence, the size of the possible gains from deception has a similar influence on team decision making as it has been found to have on individual decision making.

5 Conclusion

Without doubt, sending a wrong message in a cheap-talk sender-receiver game can be interpreted as a deliberate attempt to deceive the receiver. In fact, my data show that 95% of senders choosing the wrong Message B do so while expecting the receiver to follow their message. Like Gneezy (2005) or Hurkens and Kartik (2006), I have also found that individual senders lie significantly more often when they can gain relatively more by doing so and less often (though not significantly so in my experiment) when the receiver loses relatively more. This obviously rather stable pattern of behaviour has been explained in a recent note by Charness and Dufwenberg (2005) by their theory of guilt aversion (Charness and Dufwenberg, 2006). In a nutshell, this theory states that a player i suffers from guilt to the extent that he believes that player $j \neq i$ receives a lower payoff than i believes j believes she will receive. If subjects are guilt averse, they might behave as if they have a cost of lying in contexts where they believe their lies will mislead. Since the payoffs for the receiver differ across treatments, the theory of guilt aversion can also explain the differences in the frequency of deception across the three different treatments, as has been documented in Gneezy (2005) and in this paper.

The main point of my paper, however, concerns the classification of what a lie or deception is, and therefore challenges the methodological approach of measuring deception in the papers by Gneezy (2005), Charness and Dufwenberg (2005) or Hurkens and Kartik (2006). I have argued that the concept of deception as measured in these papers is too narrow to capture the full amount of deception prevalent in the two-person game under consideration[15], and probably also in other games where misrepresentation of information is possible. I have proposed that telling the truth should be counted as an act of deception when the sender expects the receiver *not* to follow the sender's message *and* when the true message is sent for precisely this reason. The experimental data of my first experiment with individual decision-makers has been compatible with such an interpretation of sophisticated truth-telling as a form of deception. The analysis of videotaped team-decisions has then confirmed this interpretation by showing that the true message is very often sent in order to deceive the receiver team. This form of deception through telling the truth has been shown to be driven by payoff considerations – which have been found to be important already in the seminal paper by Gneezy (2005) – and by the expectations about the receiver's actions, which has been the essential addition of my paper.

I would like to conclude by noting that this paper also adds to the literature on individual versus team decision-making – besides contributing to our understanding and the measurement of deception. I have found that teams resort to sophisticated truth-telling as a means of deception significantly more often than individuals do, thus confirming earlier findings that teams behave more strategically than individuals (Bornstein and Yaniv, 1998; Cooper and Kagel, 2005; Kugler et al., 2007). This is of particular importance when expectations about other players' actions matter. This paper's sender-receiver game is therefore closely related to the literature on level-k thinking or step-level reasoning models (Camerer et al., 2004; Costa-Gomes and Crawford, 2006). In general, such models examine a decision maker's best response given an (expected) distribution of other decision makers'

[15] It should be noted here that the occurrence of sophisticated truth-telling depends on the two-option design implemented here and on the orthogonal incentives of senders and receivers (see Farrell and Rabin, 1996). Including a third option – as in Rode (2006) – may eliminate the incentives for truth-telling as a sophisticated selfish action.

level of reasoning. One of the best known applications of such reasoning models is the guessing- or beauty-contest game (Nagel, 1995) where decision makers have to choose a number from an interval $I \equiv [0,100]$ and the winner is the decision maker whose number is closest to p times the average of all chosen numbers. Kocher and Sutter (2005) and Kocher et al. (2006) have shown that teams apply significantly higher levels of reasoning in guessing-games than individuals do, making them better in anticipating others' moves. Such a behavioural difference between individuals and teams might explain why I have observed more deception through sophisticated truth-telling by teams than by individuals, because deception through sophisticated truth-telling requires one more step of reasoning than a straightforward lie that counts on the receiver to believe it. Note, of course, that in the sender-receiver game of this paper any levels of thinking that are two steps apart result in the same behaviour.

Another noteworthy difference between individuals and teams is that receiver teams are less trusting and sender teams anticipate that correctly. This finding is related to recent evidence of a lower level of trust among teams than among individuals in simple trust experiments (where no informational asymmetries prevail; see Kugler et al., 2007; Cox, 2002). This paper has documented that (the expectation of) mutual trust is also an important factor for the behaviour in sender-receiver games where senders may have an incentive to deceive receivers. My results have shown that teams appear to be more honest at first sight, but that this is only driven by a larger degree of sophisticated truth-telling. The latter is motivated by more pessimistic expectations about how likely receiver teams will believe the sender's message. These findings reinforce the necessity to consider deception through telling the truth as a realistic option to which decision makers may resort in many important real-life situations in politics, sports, or economics.

References

Bornstein, Gary and Yaniv, Ilan (1998), Individual and group behaviour in the ultimatum game: Are groups more 'rational' players?. *Experimental Economics* 1: 101-108.

Bosman, Ronald, Henning-Schmidt, Heike and van Winden, Frans (2006), Exploring group behavior in a power-to-take video experiment. *Experimental Economics* 9: 35-51.

Camerer, Colin F., Ho, Teck-Hua and Chong, Juin-Kuan (2004), A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119: 861-898.

Camerer, Colin and Hogarth, Robin M. (1999), The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty* 19: 7-42.

Charness, Gary and Dufwenberg, Martin (2005), Deception: The role of guilt. Working Paper. University of Arizona, Tucson.

Charness, Gary and Dufwenberg, Martin (2006), Promises and partnerships. *Econometrica* 74: 1579-1602.

Chiappori, Pierre-Andre, Levitt, Stephen D. and Groseclose, Timothy (2002), Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review* 92: 1138-1151.

Cooper, David J. and Kagel, John H. (2005), Are two heads better than one? Team versus individual play in signaling games. *American Economic Review* 95: 477-509.

Costa-Gomes, Miguel and Crawford, Vincent P. (2006), Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review* 96: 1737-1768.

Cox, James C. (2002), Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. In: Zwick, Rami and Rapoport, Amnon (eds.), *Advances in Experimental Business Research*. Kluwer Academic Publishers.

Crawford, Vincent P. (2003), Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. *American Economic Review* 93: 133-149.

Crawford, Vincent P. and Sobel, Joel (1982), Strategic information transmission. *Econometrica* 50: 1431-1451.

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

Fischbacher, Urs (2007). “z-Tree: Zurich Toolbox for ready-made economic Experiments.” *Experimental Economics* 10, 171-178.

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

Greiner, Ben (2004), An Online Recruitment System for Economic Experiments. In: Kurt Kremer and Volker Macho (eds.): *Forschung und wissenschaftliches Rechnen 2003*. GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, 79-93.

Grether, David M. (1992), Testing bayes rule and the representative heuristic: Some experimental results. *Journal of Economic Behavior and Organization* 17: 623-638.

Heidhues, Paul, Lagerlof, J. (2003), Hiding information in electoral competition. *Games and Economic Behavior* 42: 48-74.

Hogarth, Robin M. (1987), *Judgment and Choice*. New York: Wiley.

Hurkens, Sjaak and Kartik, Navin (2006), (When) Would I lie to you? Comment on “Deception: The role of consequences”. University of California San Diego. Working Paper.

Kocher, Martin, Strauss, Sabine and Sutter, Matthias (2006) Individual or team decision making - causes and consequences of self-selection. *Games and Economic Behavior* 56: 259-270.

Kocher, Martin and Sutter, Matthias (2005), The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *Economic Journal* 115: 200-223.

Kugler, Tamar, Bornstein, Gary, Kocher, Martin and Sutter, Matthias (2007), Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic Psychology*, forthcoming.

Nagel, Rosemarie (1995), Unravelling in guessing games: An experimental study. *American Economic Review* 85: 1313-1326.

Rode, Julian (2006), Truth and trust in communication. An experimental study of behaviour under asymmetric information. University Pompeau Fabra Barcelona. Working Paper.

Sanchéz-Pages, Saniago and Vorsatz, Marc (2007), An experimental study of truth-telling in a sender-receiver game. *Games and Economic Behavior* 61: 86-112.

Vrij, Albert (2001), *Detecting Lies and Deceit: The Psychology of Lying and the Implications for Professional Practice.* New York: John Wiley and Sons.

Appendix A1 – Instructions for senders
Welcome to this experiment on decision making. Please read the following instructions carefully. If you have any questions, an experimenter will answer them privately.

2 people interact with each other once
In this experiment, you and another person will interact with each other for exactly one round. Each of you will have to make one single decision after which the experiment will end and you both will be paid in cash.
Both of you are randomly assigned to one of two possible roles: either the role of Person 1 or the role of Person 2. Your role in this experiment is **Person 1**.

Your decision
There are two options that can be chosen in this experiment. Both options yield different payoffs for you and the Person 2 that is paired with you. The options are as follows:

In treatment T1 it read
Option A: 5 € for you (= Person 1) and 6 € for Person 2
Option B: 6 € for you (= Person 1) and 5 € for Person 2

In treatment T2 it read
Option A: 5 € for you (= Person 1) and 15 € for Person 2
Option B: 6 € for you (= Person 1) and 5 € for Person 2

In treatment T3 it read
Option A: 5 € for you (= Person 1) and 15 € for Person 2
Option B: 15 € for you (= Person 1) and 5 € for Person 2

Person 2 will make the choice which option to implement. Yet, note that Person 2 **will not** be informed about the monetary consequences of both options. Instead, Person 2 will only receive one of the two following messages:

Message 1: „Option A will earn you more money than Option B.“
Message 2: „Option B will earn you more money than Option A.“

You have to **choose which message** you want to send to Person 2.

Information conditions and payoffs
After Person 2 has implemented one of the two options, you and Person 2 will be paid the corresponding amount of money. Note that Person 2 will neither be informed about your payoff in the chosen option nor about the possible payoffs (for both persons) in the other option. Person 2 will only be informed about his monetary payoff earned in the implemented option.

After Person 1 had sent her message, she was asked the following two questions:
While Person 2 is deciding about which option to implement, would you please answer the following two questions:
Q1: “Which option do you expect the receiver to choose?”
Q2: “Out of 100 receivers, how many do you think follow the sender’s message on average?”

Appendix A2. Instructions for receivers

Welcome to this experiment on decision making. Please read the following instructions carefully. If you have any questions, an experimenter will answer them privately.

2 people interact with each other once

In this experiment, you and another person will interact with each other for exactly one round. Each of you will have to make one single decision after which the experiment will end and you both will be paid in cash.

Both of you are randomly assigned to one of two possible roles: either the role of Person 1 or the role of Person 2. Your role in this experiment is **Person 2**.

Two possible options and your decision

Two possible monetary payoffs are available to you and Person 1 in this experiment. The payoffs depend on the option chosen by **you**. We have already showed the two options to Person 1. The only information that you will have before making your choice is the message that Person 1 has sent to you.

Person 1 will choose one of the two following messages:

Message 1: „Option A will earn you more money than Option B."
Message 2: „Option B will earn you more money than Option A."

You will be informed about the message chosen by Person 1 on the screen.

We will then ask you to choose either Option A or Option B. Your choice will determine the payoffs in the experiment. You will never know what sums were actually offered in the option not chosen (i.e. whether the message sent by Person 1 is true or not). Furthermore, you will never know the amount that Person 1 has earned in the chosen option or would have earned in the option not chosen by you.

We will pay both persons the amount of money that corresponds to the option you chose.

Thank you for your participation in the experiment.

University of Innsbruck – Working Papers in Economics and Statistics
Recent papers

2007-26	**Matthias Sutter:** Deception through telling the truth?! Experimental evidence from individuals and teams.
2007-25	**Andrea M. Leiter, Harald Oberhofer and Paul A. Raschky:** Productive disasters? Evidence from European firm level data.
2007-24	**Jesus Crespo Cuaresma:** Forecasting euro exchange rates: How much does model averaging help?
2007-23	**Matthias Sutter, Martin Kocher and Sabine Strauß:** Individuals and teams in UMTS-license auctions.
2007-22	**Jesus Crespo Cuaresma, Adusei Jumah and Sohbet Karbuz:** Modelling and Forecasting Oil Prices: The Role of Asymmetric Cycles.
2007-21	**Uwe Dulleck and Rudolf Kerschbamer:** Experts vs. discounters: Consumer free riding and experts withholding advice in markets for credence goods.
2007-20	**Christiane Schwieren and Matthias Sutter:** Trust in cooperation or ability? An experimental study on gender differences. *Conditionally accepted for publication in: Economics Letters.*
2007-19	**Matthias Sutter and Christina Strassmair:** Communication, cooperation and collusion in team tournaments – An experimental study.
2007-18	**Michael Hanke, Jürgen Huber, Michael Kirchler and Matthias Sutter:** The economic consequences of a Tobin-tax – An experimental analysis.
2007-17	**Michael Pfaffermayr:** Conditional beta- and sigma-convergence in space: A maximum likelihood approach.
2007-16	**Anita Gantner:** Bargaining, search, and outside options. *Revised version forthcoming in: Games and Economic Behavior.*
2007-15	**Sergio Currarini and Francesco Feri:** Bilateral information sharing in oligopoly.
2007-14	**Francesco Feri:** Network formation with endogenous decay.
2007-13	**James B. Davies, Martin Kocher and Matthias Sutter:** Economics research in Canada: A long-run assessment of journal publications. *Revised version forthcoming in: Canadian Journal of Economics.*
2007-12	**Wolfgang Luhan, Martin Kocher and Matthias Sutter:** Group polarization in the team dictator game reconsidered. *Revised version forthcoming in: Experimental Economics.*
2007-11	**Onno Hoffmeister and Reimund Schwarze:** The winding road to industrial safety. Evidence on the effects of environmental liability on accident prevention in Germany.
2007-10	**Jesus Crespo Cuaresma and Tomas Slacik:** An "almost-too-late" warning mechanism for currency crises.
2007-09	**Jesus Crespo Cuaresma, Neil Foster and Johann Scharler:** Barriers to technology adoption, international R&D spillovers and growth.
2007-08	**Andreas Brezger and Stefan Lang:** Simultaneous probability statements for Bayesian P-splines.
2007-07	**Georg Meran and Reimund Schwarze:** Can minimum prices assure the quality of professional services?.
2007-06	**Michal Brzoza-Brzezina and Jesus Crespo Cuaresma:** Mr. Wicksell and the global economy: What drives real interest rates?.
2007-05	**Paul Raschky:** Estimating the effects of risk transfer mechanisms against floods in Europe and U.S.A.: A dynamic panel approach.
2007-04	**Paul Raschky and Hannelore Weck-Hannemann:** Charity hazard - A real hazard to natural disaster insurance.
2007-03	**Paul Raschky:** The overprotective parent - Bureaucratic agencies and natural hazard management.

2007-02 **Martin Kocher, Todd Cherry, Stephan Kroll, Robert J. Netzer and Matthias Sutter:** Conditional cooperation on three continents.

2007-01 **Martin Kocher, Matthias Sutter and Florian Wakolbinger:** The impact of naïve advice and observational learning in beauty-contest games.

University of Innsbruck

Working Papers in Economics and Statistics

2007-26

Matthias Sutter

Deception through telling the truth?! Experimental evidence from individuals and teams

Abstract
Informational asymmetries abound in economic decision making and often provide an incentive for deception through telling a lie or misrepresenting information. In this paper I use a cheap-talk sender-receiver experiment to show that telling the truth should be classified as deception too if the sender chooses the true message with the expectation that the receiver will not follow the sender's (true) message. The experimental data reveal a large degree of 'sophisticated' deception through telling the truth. The robustness of my broader definition of deception is confirmed in an experimental treatment where teams make decisions.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)