University of Heidelberg

Department of Economics

Discussion Paper Series | No. 581

Who never tells a lie?

Christoph Vanberg

February 2015

Who never tells a lie?

Christoph Vanberg*

Abstract

Erat and Gneezy (2012) conduct an experiment to test whether people avoid lying in a situation where doing so would lead to a Pareto improvement. They conclude that many people exhibit such a "pure lie aversion." I argue that the experiment does not provide a reliable test for such an aversion, and that the evidence does not support the authors' conclusion. I conduct two new experiments which are explicitly designed to test for a 'pure' aversion to lying, and find no evidence for the existence of such a motivation. I discuss the implications of the findings for moral behavior and rule following more generally.

Keywords: Lying; Deception; Morality; Ethics; Experiments

*Department of Economics, University of Heidelberg

1 Introduction

One of the most basic moral rules in most cultures is that one should not lie. The importance of this moral principle is reflected in stories told to children, such as the one about the boy who cried wolf. Some moral philosophers follow Immanuel Kant (1785) in suggesting that rules of this kind should be followed in all instances, irrespective of the consequences to be expected in any particular case. According to this position, it would be wrong to lie even in a case where doing so would lead to consequences that would be generally regarded as desirable. Given the importance of this and other basic moral principles, it is interesting to ask whether (or to what extent) common moral attitudes and behaviors coincide with this Kantian position. That is, to what extent do people regard lying as 'wrong' and therefore avoid lying, irrespective of the consequences to be expected from doing so?

Together with various coauthors, Uri Gneezy has investigated this question using simple laboratory experiments (Gneezy 2005; Erat and Gneezy 2012; Gneezy et al 2013). In general, the message emerging from this literature has been that most experimental subjects take expected consequences into account when deciding whether to lie. In particular, people are more likely to lie if the personal benefit to be gained from doing so is larger, and if the harm imposed on others is smaller. With one exception, these experiments have investigated lying in settings where the person telling a lie can benefit at a cost to someone else.

In a recent contribution, Erat and Gneezy (2012, henceforth EG) aimed to investigate cases where lying can be materially beneficial to another person ('white lies'), and possibly to both the liar and another person ('Pareto white lies'). Testing for an aversion to the latter type of lie is particularly interesting. Since the material consequences of telling such a lie are unambiguously desirable, such an aversion could be attributed only to a "pure" cost of lying. Therefore this could provide a good test for the existence of 'Kantian' moral attitudes. As the authors explain,

> "[P]eople who are reluctant to tell Pareto white lies demonstrate lie aversion independent of social preferences for outcomes. Such people refrain from lying not (merely) because of the consequences, but because they simply view lying as a bad act in itself. This provides the best test of a pure cost of lying in line with a moral stand." (ibid. p. 724)

To test for such a 'pure' aversion to lying, EG conduct an experiment in which one participant (the 'sender') is given the opportunity to send either a true or a false message to another (the 'receiver). The experimental design will be described in detail in Section 2. As we will see, the aim of that design is to create the impression, on the part of the sender, that *both* subjects will benefit (receive a larger monetary payment) if she lies (sends a false message).

Thus, EG place subjects into a situation that is *intended* to induce the belief that telling a lie will lead to mutually beneficial consequences. The main finding of the experiment is that 35% of participants tell the truth even in this situation. Based on this observation, the authors draw the following conclusion:

> "We find that a large fraction of participants are reluctant to tell even a Pareto white lie, demonstrating a pure lie aversion independent of any social preferences for outcomes." (ibid. p. 723)

They elaborate

> "This result represents clean evidence of lie aversion, without any confounding of distributional preferences. Thus, the aversion to lying cannot be (fully) explained by the harm it causes others, as the aversion is present even when everyone gains from the lie." (ibid. p. 726)

This result of EG's experiment has received substantial attention in the literature.[1] Many authors cite EG as evidence of a 'pure' aversion to lying. If correct, the result would be of substantive relevance to our understanding, not just of truth telling, but of moral behavior more generally. In particular, it would seem to indicate that many people do subscribe to a Kantian ethical position and avoid violating general moral principles irrespective of the consequences to be expected from doing so. Therefore,

[1] Published in *Management Science* in 2012, google scholar counts 120 citations as of January 2015. Many authors explicitly state that EG show a reluctance to tell even Pareto white lies (e.g. Cappelen et al 2013, Chen and Houser 2013, Rosaz and Villeval 2012, to name just a few).

the evidence EG provide, as well as the interpretation they offer, deserve to be considered carefully.

This paper calls that interpretation into question. I will argue that EG's experimental design does not provide a reliable test for a 'pure' aversion to lying. In particular, that design is likely to have fostered considerable confusion as to the material consequences to be expected from lying. As a result, *those subjects to whom EG attribute a 'pure' aversion to lying are likely to have believed that telling the truth would lead to better material consequences for both subjects.* As I will show, this belief was supported by a lack of common knowledge about the rules of the game, as well as inaccurate and contradictory wording of instructions.

As a consequence of these features, the experiment conducted by EG cannot be used to test the conjecture that (many) people avoid lying even when doing so would result in a Pareto improvement. To do so, it is necessary to conduct an experiment in which subjects actually have reason to expect, with a sufficient degree of confidence, that lying will result in a Pareto improvement. The experiments I report on here are designed to achieve this, while otherwise confronting subjects with essentially the same choice as in EG's experiment.

To avoid a possible misunderstanding, it is important to reiterate and elaborate on this point briefly. The aim of the present paper is *not* to replicate the experiment that EG have run, or to call into question the results obtained under their design.[2] The concern I am raising is that EG's experiment does not provide a reliable test for a 'pure' aversion to lying. Therefore my goal is to provide a more reliable test of the claim, put forth by EG, that (many) people exhibit such an aversion. Thus, I am questioning the *conclusion* that EG draw from their results, and I will treat that conclusion as a *conjecture* to be tested herein. In order to do so, it is necessary to conduct experiments *different* from the one that EG have conducted.

I therefore conduct two new experiments explicitly designed to test the hypothesis that a significant part of the population avoids lying, even when doing so would result in a Pareto improvement. Subjects in both experiments face essentially the same choice as in EG. Both experiments are designed such that the material consequences of lying are considerably less ambiguous. As we will see, the results obtained differ substantially from those of EG, and cast doubt on the existence of a 'pure' aversion to lying. More generally, they suggest that subjects' moral attitudes rarely, if ever, coincide with the Kantian position which demands that ethical rules be followed in all instances and irrespective of consequences.

The remainder of the paper is organized as follows. Section 2 discusses EG's experiment and explains reasons to doubt their interpretation of subject behavior.

[2]If the experiment run by EG were repeated, I fully expect that the results obtained would be the same. Either way, the data collected would not constitute evidence either for or against the *hypothesis* we are concerned with, as EG's design does not provide a reliable test of that hypothesis.

Section 3 presents my experimental design and argues that it is appropriate for testing the validity of that interpretation. Section 4 describes the results and concludes that there is no evidence for an aversion to telling Pareto white lies. Section 5 concludes.

2 Erat and Gneezy's (2012) Experiment

EG conduct a game involving 2 players, labeled 'sender' and 'receiver'. At the beginning of the game, the sender is informed that the experimenter has rolled a die and that the outcome was '2'. He is also told that the receiver does not know the outcome of the die roll. The sender is asked to send a message of the form 'The number rolled was X'. This message is transmitted to the receiver, who must then choose a number between 1 and 6. I will refer to this number as the receiver's 'guess,' though EG do not use this terminology.

The sender is told that the players' payoffs are determined as follows. If the receiver chooses 2 ('guesses correctly'), both receive 20 dollars. Otherwise (i.e. if she 'guesses wrong'), both receive 30 dollars. Finally, the sender is given the following information concerning the receiver's knowledge

> "Finally, only you will be informed of the particular monetary value connected to each message. The other participant will not be informed

> of these monetary values. However, he or she will be told that you have been informed of the monetary value connected to each message."

These last statements are actually inaccurate (see below). Despite the repeated mentioning of "messages," they are meant to convey the fact that the receiver is not informed about the monetary consequences connected to an accurate or inaccurate *guess*. In particular, only the sender knows that both players receive a higher payoff if the receiver "guesses wrong".

The authors find that 35% of senders tell the truth in this situation (choose the message "The number rolled was 2"). They conclude from this that these senders avoid lying despite the fact that doing so would result in a Pareto improvement. This interpretation of the data is summarized as

> "RESULT 2. A significant fraction of senders do not lie *even when lying results in a Pareto improvement*." (ibid. p. 726, emphasis added)

As explained above, this interpretation of the experimental results appears to be widely accepted, and has received considerable attention in the literature (see footnote 1). However, there are strong reasons to doubt that interpretation. In particular, the assertion that "lying results in a Pareto improvement" relies entirely on the assumption that senders expect receivers to "follow" their message - i.e. to choose the number corresponding to it.

Closer examination of EG's experimental design as well as the wording of instructions reveals that this assumption is unjustified. Indeed it seems highly likely that some participants may have felt sufficiently certain that the receiver will *not* follow their message to have believed that telling the truth would be materially advantageous for both.[3]

In order to establish this point, it is necessary to consider the instructions EG provided to senders. These are reprinted in Appendix 1. I will identify two reasons why the assumption that senders expect receivers to follow is unjustified. First, if the instructions are interpreted in a way that is consistent with EG's own description of the experiment, then the receiver has no reason to follow the message and therefore the sender has no reason to expect that she will do so. Second, the contradictory and inaccurate wording mentioned above is likely to have supported another interpretation of the instructions, according to which the receiver would have every reason *not* to follow the message, and so the sender would be quite confident that she will do so.

I will begin with the first issue. In describing their experimental design, EG explain the situation as follows:

[3]The authors dismiss this objection on the grounds that the sender would have to be quite certain that the receiver will not follow for this argument to apply. My concern, however, is that the design and instructions made it quite likely that this may have been true for many senders.

> "The sender was told (...) that the only information the receiver would have regarding the actual outcome of the roll of the die was her message. There were two payment options, A and B. The sender knew the payoffs (to both players) associated with each option, and she was told the receiver would not know the payoffs. Finally, the sender was told that if the receiver chose the real outcome of the die roll, payment option A would be implemented. Otherwise, both would be paid according to option B."

Suppose that a sender interprets the experimental instructions in a way that is consistent with this description. Then, she believes that the receiver simply *does not know anything* about the payoff consequences of guessing correctly or wrong. In particular, the receiver does not even know that both participants will receive the same payoff. Given her complete ignorance, there does not then appear to be any rational reason for the receiver to follow the sender's message or not. Indeed, she may well assume that the players' interests are *opposed*, in which case it seems plausible that she will *not* follow (though again there would be no *rational* argument to support either choice). Thus, if the instructions are interpreted in a way that is consistent with the intended experimental design, it is not at all obvious that senders should expect receivers to follow.[4]

[4]Interestingly, the *receivers'* instructions were worded in such a way as to suggest that interests are aligned. This may explain why most receivers did in fact follow the message. However that

A second interpretation of the instructions is possible because the text describing the receiver's knowledge of the game repeatedly refers to the 'monetary value connected to each *message*' (emphasis added). As the sender knows, payoffs depend on the correctness of the receiver's *guess*, and *not on the message*. This contradictory information is very likely to have fostered confusion. In fact, the text might well be interpreted as simply a restatement of the fact that the receiver does not know the *number rolled*. Under this interpretation, *the receiver knows that both will receive a higher payoff if she guesses wrong*. Since she does not know the outcome of the die roll, she does know *which messages* are correct and which one's are wrong. In that case, the natural thing for a sender to do is to tell the truth ('the number rolled is 2'), fully confident that the receiver will *not* follow the message (choose a number different from 2), since she knows that this will lead to a higher payoff for both.[5]

Based on these considerations, it does not appear unlikely that the 35% of subjects who told the truth in EG's experiment did so *believing that the receiver would not follow*. If we interpret the data using this assumption, we would come to a different conclusion:

---

fact is not relevant when it comes to assessing senders' expectations.

[5]After reading EG's instruction, many participants of an informal survey argued either that the message is entirely 'irrelevant' to the receiver's choice (because she knows nothing about consequences), or that the receiver would probably not follow, either because she would 'distrust' the sender, or because she would know that it's best to guess wrong.

"Result 2′ A significant fraction of senders tell the truth when telling the truth results in a Pareto improvement."

Naturally, EG's data cannot be used to assess the validity, either of "Result 2" or of "Result 2′" - both of these "Results" constitute *interpretations* of the data based on different assumptions about what senders expect receivers to do. Since the experiment does not *reliably* induce the expectation that receivers will follow, it does not provide a reliable test of what senders do when lying leads to a Pareto improvement.

In order to conduct such a test, we need to look at a situation in which senders actually expect receivers to follow. The experiments reported on here were designed to do achieve exactly that, while ultimately presenting senders with essentially the same choice that they make in EG. The next section describes the design and how it achieves this aim.

3 Experimental Design and Procedures

Design

I conducted two experiments explicitly designed to test the hypothesis that (many) people avoid lying even when doing so would result in a Pareto improvement. Both experiments involve sender-receiver games similar to the one conducted by EG. Senders in both experiments face the same choice between accurate or inaccurate messages concerning the roll of a die. Other than this, the experiments differ substantially from EG, such that senders face less ambiguity with respect to receiver behavior. This section describes only one of these experiments in detail. The other treatment is discussed in Appendix 4. Results from both treatments are essentially identical.

As in Erat and Gneezy (2012), subjects played a sender-receiver game. First, the sender rolled a 6-sided die.[6] He then sent a message of the form 'The number rolled was X'. Finally, the receiver chose a number between 1 and 6. If the chosen number corresponded to the number rolled, both participants received a certain common payoff A. If not, both received another common payoff B. As in EG, only the sender was informed about the concrete values A and B. Unlike in their experiment, however, the fact that both subjects would receive the same payoff was common knowledge. All subjects received the same instructions, such that senders knew perfectly what information was available to receivers.

An important feature of this design is that the players commonly know that their material interests are aligned, though still the receiver does not know whether it is better to guess correctly or not. My expectation was that this should make senders more confident that the receiver will indeed follow their message. (Naturally,

[6]The roll of the die was simulated by a computer.

there is still no purely *rational* argument for doing so.) In order to emphasize the commonality of interest, the instructions referred to the participants as a 'team'.

In order to further increase the predictability of the receiver's behavior, the game was repeated 8 times in fixed pairs. The payoff values A and B varied from round to round, such that sometimes it was better to 'guess correctly' ($A > B$) and sometimes it was better to 'guess wrong' ($B > A$). The roles (sender/receiver) were switched after 4 rounds. I expected this repetition to further reduce uncertainty as to the receiver's behavior. In order to elicit relevant observations not affected by prior coordination, the first round was set up such that $B > A$, i.e. it was better to guess a wrong number.

Finally, senders were asked whether they expect the receiver to follow their message or not. This elicitation was not incentivized.[7] All subjects were paid in cash for one randomly chosen round.

Procedures

The experiment was conducted at a large University in Germany. Subjects were students from a variety of disciplines. Participants were recruited using the online

[7] In the other treatment, the receiver was allowed to *tell* the sender, prior to the choice of message, whether she intends to follow the message or not. After observing that no sender told the truth in that treatment, this feature was eliminated. See Appendix 4.

recruitment system ORSEE (Greiner 2004). The experiment was computerized using z-Tree (Fischbacher 2007). We conducted two sessions involving 16 and 20 subjects. Upon entering the laboratory, subjects were randomly assigned to visually isolated computer terminals. Instructions (reproduced in Appendix 2) were handed out. After subjects had read the instructions, they were read aloud by the experimenter. Next, subjects filled out a brief questionnaire testing comprehension of the rules (reproduced in Appendix 3). The experiment began after all subjects had correctly filled out the questionnaire.[8] The payoff values A and B were either 12 and 6 EUR or 6 and 12 EUR. Each subject played 4 rounds as a sender, 2 in which $A > B$ and 2 in which $B > A$.[9] After the experiment was completed, one of the 8 rounds was randomly chosen for payment. Subjects filled out a post-experimental questionnaire asking about their attitude towards lying. Finally, participants were paid in private and left the laboratory.

[8]Only very few subjects made a mistake on the questionnaire. In such cases they were asked to reconsider their answers. All subjects were then able to correct their errors immediately.

[9]The instructions informed subjects that the payoffs would differ from round to round, but no information was given as to the underlying distribution.

4 Results

4.1 Data

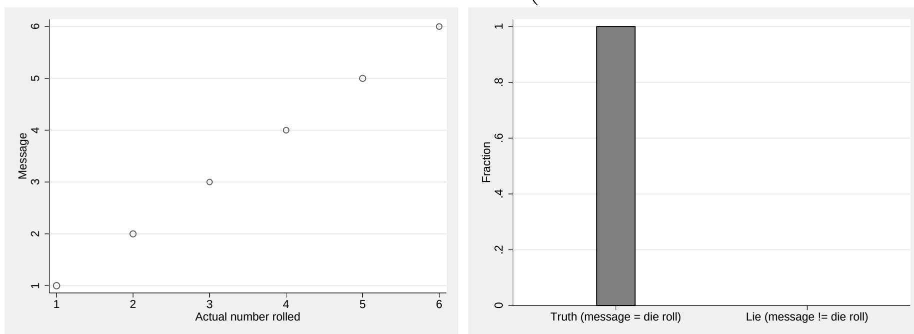
Each of the 36 participants made 4 choices as a sender, giving us 144 cases overall. In 72 of these cases, both participants would have received a higher payoff (12 instead of 6 EUR) if the receiver correctly guessed the number rolled. In the other 72 cases, they would have received a higher payoff if the guess was wrong. Naturally, these observations are not independent. Given the fixed matching scheme, we have 18 statistically independent observations, this being the number of pairs observed.

4.2 When it was better to guess correctly

First, let us look at the 72 cases where $A > B$. In that situation, both the sender and the receiver receive 12 EUR if the receiver guesses correctly and 6 EUR if not. In 68 of those cases (94%), senders said that they expected the receiver would follow the message. Given this expectation, the (joint) payoff maximizing choice for the sender would be to tell the *truth*, i.e. to transmit the number actually rolled.

The left panel of Figure 1, depicts the distribution of messages given the numbers actually rolled. Note that all points are located along the 45 degree line, suggesting a strong correlation between the messages sent and the actual roll of the die. Another

Figure 1: Truth telling when A>B (better to guess correctly)



Left panel: The distribution of messages given actual rolls of the die.

Right panel: The relative frequency of truth and lies.

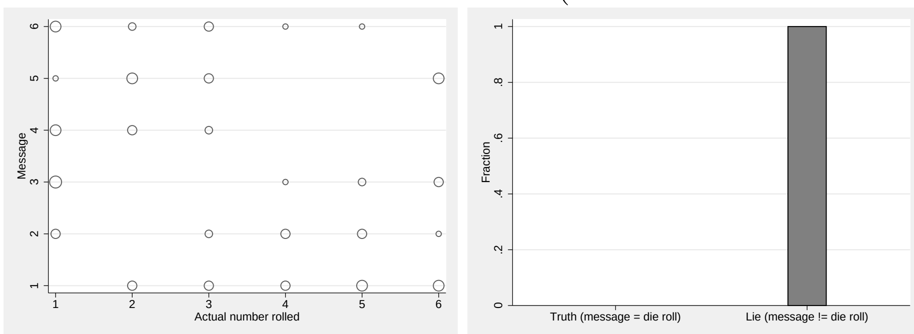
way to visualize this result is shown in the right panel, which contains a histogram of the frequencies of lying and truth telling in this condition. Naturally, these data are consistent with the notion that some or all subjects dislike telling lies.

4.3 When it was better to guess wrong

Next, consider the 72 cases where $B > A$. In that situation, both the sender and the receiver receive 12 EUR if the receiver guesses *wrong* and 6 EUR if he guesses correctly. In 67 of those cases (93%), senders said that they expected the receiver would follow their message. (This fraction is not statistically different from the expectation in the $A > B$ conditions.) Given this belief, the (joint) payoff maximizing choice for the sender would be to *lie*, i.e. to transmit a number different from the one actually rolled.

The left panel of Figure 2, displays the observed distribution of messages given the

Figure 2: Truth telling when B>A (better to guess wrong)



Left panel: The distribution of messages given actual rolls of the die.

Right panel: The relative frequency of truth and lies.

number rolled. In contrast to Figure 1, we can see that all observations are located away from the 45 degree line, suggesting that the message sent was reliably different from the actual roll of the die. A histogram displaying the frequency of truth telling and lying is provided in the right panel of Figure 2. Naturally, these data are not at all consistent with the notion that any of the participants dislike telling lies under all circumstances.

5 Conclusion

Erat and Gneezy (2012) conduct an experiment to test whether people are reluctant to tell a lie even when doing so results in a Pareto improvement. According to the authors, this approach "provides the best test of a pure cost of lying in line with a moral stand." Based on their results, EG conclude that indeed a significant number of people exhibit such a 'pure' aversion to lying. This conclusion has received substantial attention and appears to be widely accepted. I have argued that this interpretation of the evidence is doubtful because EG's experimental design and the wording of instructions are likely to have fostered the belief that telling the truth would lead to a Pareto improvement.

I conducted two new experiments *similar* to but deliberately *distinct* from EG. My experiments are designed to explicitly test the claim, put forth by EG, that (many) people avoid telling even Pareto white lies. The main difference is that the material consequences from lying are far less ambiguous in the experiments reported on here. Although this reduction in ambiguity is achieved by substantial modifications in the overall design, subjects ultimately make the same choice between telling the truth or lying about the outcome of a die roll. A subject who exhibits 'pure' lying aversion in EG's sense should therefore refrain from lying in these experiments as much as they (ostensibly) did in theirs.

Out of the 72 participants who took part in the two experiments, *each and every* one chose to lie when doing so implied a Pareto improvement in material terms. This evidence suggests that a truly 'pure' aversion to lying (in Erat and Gneezy's sense) is quite rare indeed, and possibly non-existent. More broadly, it indicates that few if any subjects adhere to a Kantian ethic according to which general moral rules must be followed under all circumstances, even when a violation would bring about consequences that are unanimously preferred.

Before concluding, it may be important to reconsider the implications of this result for the question of whether honest behavior is typically motivated by a general disposition or commitment to moral rules. Superficially, the evidence presented here might be taken to favor instead the idea that such behavior is governed entirely by an evaluation of its foreseeable consequences. This would be premature. What the evidence suggests is merely that subjects' attitudes about moral rules are not *completely* insensitive to the foreseeable consequences of rule violations. Under normal circumstances, common moral rules prohibit behaviors that inflict *harm* on others. Though Kantian moral philosophers may argue that such rules should be followed even in exceptional instances where no harm is done, it seems possible that people can be simultaneously *committed* to such rules in general, but ready to make exceptions in instances where it is obvious that the rule does not serve any purpose. In other words, the evidence presented here is inconsistent with a Kantian ethical attitude, but consistent with a 'pragmatic' preference for moral rule following.

References

Cappelen, A., E. Sorensen, B. Tungodden 2013. When do we lie? *Journal of Economic Behavior & Organization* **3**, 258-265.

Chen, J. and D. Houser 2013. Promises and Lies: An Experiment on Detecting Deception. *GMU working paper.*

ERAT, S. AND U. GNEEZY 2012. White Lies. *Management Science* **58**, 723-733.

FISCHBACHER, U. 2007. Zurich toolbox for readymade economic experiments. *Experimental Economics* **10**, 171-178.

GNEEZY, U. 2005. Deception: The role of consequences. *American Economic Review* **95**, 384-394.

GNEEZY, U., B. ROCKENBACH, M. SERRA-GARCIA 2013. Measuring Lying Aversion. *Journal of Economic Behavior and Organization* **93**, 293-300.

GREINER, B. 2004. An Online Recruitment System for Economic Experiments. *Forschung und wissenschaftliches Rechnen* 2003. GWDG Bericht **63**, 79-93.

KANT, I. 1785. *Fundamental Principles of the Metaphysics of Morals.*

ROSAZ, J. AND M.C. VILLEVAL 2012. Lies and biased evaluation: A real effort experiment. *Journal of Economic Behavior and Organization* **84**, 537-549.

Appendix 1: Erat and Gneezy (2012) Instructions

The following instructions were provided to senders:

"Before starting this experiment, we have rolled a 6-sided die, and obtained the outcome 2. The other participant will not be informed that the outcome of the die roll was 2. However, he or she will be told that you have been informed the outcome of the die roll.

We now ask you to send a message to the other participant. The message will correspond to a number from 1 to 6. There are six possible messages:

Message 1: 'The outcome from the roll of the 6-sided die is 1.'
Message 2: 'The outcome from the roll of the 6-sided die is 2.'
Message 3: 'The outcome from the roll of the 6-sided die is 3.'
Message 4: 'The outcome from the roll of the 6-sided die is 4.'
Message 5: 'The outcome from the roll of the 6-sided die is 5.'
Message 6: 'The outcome from the roll of the 6-sided die is 6.'

Your message will be shown to the other participant, and then he or she will be asked to choose a number between 1 and 6.

The choice of the number if the other participant will determine the payments in the experiment.

- If he or she chooses the actual outcome of the roll of the dice, 2, then you will receive $20 and he or she will receive $20.
- If he or she chooses a number different than the actual outcome, you will receive $30 and he or she will receive $30.

Finally, only you will be informed of the particular monetary value connected to each message. The other participant will not be informed of these monetary values. However, he or she will be told that you have been informed of the monetary value connected to each message."

Appendix 2: Instructions for main treatment

The following is a translation of the printed instructions. Original German instructions are available upon request.

<u>Instructions (Page 1 of 2)</u>

Basic Procedure

Teams

Each participant will be randomly assigned to a **Team**. Each **Team** consists of **2** Participants: one **Roller** und one **Chooser**. At the end of the experiment, the two members of a team will receive the same **common payoff**.

Rounds

The experiment will last for 8 rounds. **The composition of the teams will remain fixed for all 8 rounds**. After 4 rounds, the members of a team will switch roles (Roller and Chooser). At the end of the experiment, **one round** will be randomly chosen for payment. Your payoff will correspond to the common payoff of your team in the round chosen.

Procedure within a round

1. The roller **rolls** (with a die) a number between 1 and 6.
   (The roll of the die will be simulated by the computer.)
   Only the roller will know the number rolled.

2. The chooser **chooses** a number between 1 and 6.
   The common payoff of the team depends upon whether the chosen number corresponds to the number rolled.

The **common payoff** of the team is determined as follows:

If number chosen = number rolled -> **both** participants receive **A EUR**

If number chosen = different number -> **both** participants receive **B EUR**

The concrete numbers **A** and **B** will differ from round to round.

<u>Instructions (Page 2 of 2)</u>

Important Details

What the participants know

Only the **roller** in the team will know the number rolled, and only he will know the concrete payment values **A** and **B**.

The **chooser** in the team will not learn the number rolled and the payment values until the end of the round, i.e. after he has chosen a number.

The opportunity to send messages

After the roller has rolled the die, and **before** the chooser chooses a number, the roller can send one of the following 6 messages:

- “The number rolled is 1”
- “The number rolled is 2”
- “The number rolled is 3”
- “The number rolled is 4”
- “The number rolled is 5”
- “The number rolled is 6”

This message is immediately transmitted to the chooser.

Appendix 3: Questionnaire

The following is an English translation of the control questionnaire.

Control Questions (Page 1/2)

(1) Assume: A = 12 EUR und B = 6 EUR. The number rolled is 3. The chooser chooses the number 2. How large are the payoffs?

ANSWER: The roller receives: ____ EUR

The chooser receives: ____ EUR

(2) Assume: A = 12 EUR und B = 6 EUR. The number rolled is 3.

- Which number would the chooser have to choose in order to receive the largest payoff?

  Answer: _______

- What would the payoffs be in this case?

  Payoff roller: _____

  Payoff chooser: _____

(3) Assume: A = 12 EUR und B = 6 EUR. The number rolled is 4. Suppose that the chooser will pick the number which the roller reports to him. What payoffs would result if the roller sends the following messages? (If the information provided is insufficient to determine the result, mark the field with an "X".)

- „The number rolled is 1." Payoffs: Roller: ____ Chooser: ____
- „The number rolled is 5." Payoffs: Roller: ____ Chooser: ____
- „The number rolled is 4." Payoffs: Roller: ____ Chooser: ____

Control Questions (Page 2/2)

(4) Under what circumstances will the roller receive a larger payoff than the chooser? (Please mark all correct answers!)

- If the chooser chooses a number different from the number rolled.
- If the chooser picks the number that the roller reports to him.
- If the chooser reports the number that he rolled.
- Under no circumstances.

(5) Assume: A = 6 EUR und B = 12 EUR. The number rolled is 4. Suppose that the chooser will NOT pick the number which the roller reports to him. What payoffs would result if the roller sends the following messages? (If the information provided is insufficient to determine the result, mark the field with an "X".)

- „The number rolled is 1." Payoffs: Roller: ____ Chooser: ____
- „The number rolled is 5." Payoffs: Roller: ____ Chooser: ____
- „The number rolled is 4." Payoffs: Roller: ____ Chooser: ____

Appendix 4: Additional treatment

Prior to conducting the treatment described in the main text, I conducted two sessions with a slightly different design. The main difference between that design and the one described above was that the receiver was permitted to send a message to the sender, explicitly stating whether she intends to follow the sender's message. 40 subjects participated in this treatment. As in the main treatment, subjects interacted in fixed pairs and each subject was in the role of sender four times. Each subjects faced two situations in which it would have been better to guess correctly ($A > B$) and two in which it would be better to guess wrong ($B > A$). Unlike the main treatment, the first round involved $A > B$, i.e. it was better to guess correctly.

With only four exceptions out of 160 cases (discussed in a moment), results from this treatment are identical to those from the main treatment: All receivers stated that they would follow the sender's message. In all cases where it was better to guess correctly, senders told the truth. In all cases where it was better to guess wrong, senders lied. The four exceptions involved receivers who indicated that they would not follow the message. In two of these, it was better to guess correctly, and senders chose to lie. In two others, it was better to guess wrong, and senders told the truth. Thus, *in all cases, senders chose the message that maximized the material payoff given the receiver's stated intention.*