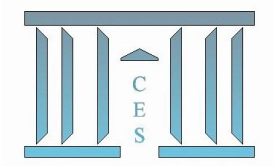
Truth-telling under Oath

Nicolas Jacquemet, Stéphane Luchini, Julie Rosaz, Jason Shogren

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers



Truth-telling under Oath

Nicolas JACQUEMET, Stéphane LUCHINI,

Julie ROSAZ, Jason F. SHOGREN

2015.68

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Truth-telling under Oath*

Nicolas Jacquemet† Stéphane Luchini‡ Julie Rosaz § Jason F. Shogren¶

July 2015

Abstract

A growing experimental literature has explored how monetary incentives affect truth-telling and lying behavior. We extend this literature to consider how two non-monetary incentives– a loaded environment and commitment through a truth-telling oath–affect truth-telling and lying behavior. For a loaded environment, we revise the standard lying experiment by making it explicit and clear to the person that "a lie is a lie". We then combine the lying experiment with a solemn oath procedure, by which subjects commit themselves to tell the truth before entering the laboratory. Both non-monetary incentive devices affect a person's willingness to tell the truth: subjects lie slightly less frequently in the loaded environment, and drastically less after they signed the solemn oath. Interestingly, the loaded environment and oath have distinct effects–the oath changes the incentive to lie only when truthfulness is made meaningful through the loaded environment.

Keywords: Deception, lies, truth-telling oath, experiments.

JEL Classification: C92, D03, D63.

1 Introduction

Situations in which a person (or a group of persons) holds private information that can be held back or manipulated to generate gains are inevitable in human interactions. Examples abound in economics, law, political sciences, and psychology. In economics, these situations involve strategic communication and are associated with two strands of literature Kartik (2009). First, games of persuasion, or verifiable disclosures, are situations in which the information is verifiable and an agent can only under-report (Milgrom, 1981; Grossman, 1981). The implicit assumption in

*We thank Robert-Vincent Joulé, Rosemarie Nagel and Marie-Claire Villeval for their comments. Jacquemet gratefully acknowledges the *Institut Universitaire de France*.

†Université de Lorraine (BETA) and Paris School of Economics. 3 Place Carnot, 54035 Nancy, France. Nicolas.Jacquemet@univ-lorraine.fr

‡GREQAM-CNRS, Centre de la Vieille Charité, 13236 Marseille Cedex 02, France. stephane.luchini@univmed.fr

§Université Montpellier 1, UMR5474 LAMETA, F-34000 Montpellier, France. rosaz@lameta.univ-montp1.fr

¶Department of Economics and Finance, University of Wyoming, Laramie, WY 82071-3985, United States. JRamses@uwyo.edu

this literature is that lying is infinitely costly. In contrast, "cheap talk" games reflect strategic situations in which information is not verifiable and can be manipulated at no cost (see, *e.g.*, Crawford and Sobel, 1982). A bridge between these two strands, however, has emerged in the work on the "intrinsic costs of lying", which suggests that strategic communication results from a trade-off between the individual cost of lying and material consequences (see, in particular, Boles, Croson, and Murnighan, 2000; Brandts and Charness, 2003; Gneezy, 2005; Sutter, 2009; Fischbacher and Heusi, 2013; Ariely, 2012). Several intrinsic motivations for truth-telling have been put forward in the economic literature – e.g., guilt aversion – the willingness to avoid letting down the expectations of others (Charness and Dufwenberg, 2006) – or pure-lying aversion – the intrinsic willingness to tell the truth.

Recent evidence suggests that people cannot be divided into "types", *i.e.* either economic liars who only care about consequences or ethical liars who only care about the process that leads to the outcome. Rather, the intrinsic motivation for truth-telling evolves during one's life and, of particular interest for this paper, depends on context and framing. (Houser, Vetter, and Winter, 2012) show that truth-telling strongly depends on the sequence of events as people are more willing to lie if they feel unfairly treated in previous interactions. Deception is also more frequent when lies are explicit (through untruthful statements) rather than implicit (through information-revealing actions) – (see, *e.g.*, Kriss, Nagel, and Weber, 2013). In the same spirit, Cappelen, Sorensen, and Tungodden (2013) observe that the effect of different kinds of framing (intuitive reasoning or personal content of the message) changes the propensity of lying to the same extent as changes in payoff do. If context matters, the open question is what kind of institutional devices one can design to foster truth-telling.

In situations in which truth-telling is difficult to verify, social institutions have relied and still rely on oaths or codes of conducts to which people have to subscribe. In Antiquity, oaths played an important role in commerce. For instance, a merchant who claimed to have been robbed during his journey was asked to take an oath to avoid retaliation from the person or persons who were supposed to receive those goods (Silver, 1995). Oaths in that sense could be seen as a mechanism to decrease transactions costs and to guarantee a sufficient control which would have been prohibitive or, even impossible to implement. In modern days, oaths are still widespread, *e.g.*, in courtroom before testimony, in codes of ethics presented to college students, in declarations before taking office. Empirical evidence in social psychology, especially in the context of commitment theory, also shows that oath procedures enhance the willingness to tell the truth. Taking an oath works similar to foot-in-the-door experiments, it is a preliminary task that makes people more likely to behave truthfully in subsequent decision problems. The truth-telling oath as a commitment device have proved to be useful in triggering economic commitment in preference elicitation, it increases truth telling in hypothetical settings with no monetary incentives (Jacquemet, Joule, Luchini, and Shogren, 2013; Jacquemet, James, Luchini, and Shogren, 2015)[1] and coordination games with

[1] Also see the follow-up studies by Carlsson, Kataria, Krupnick, Lampi, Lofgren, Qin, Sterner, and Chung (2013);

cheap talk communication (Jacquemet, Luchini, Shogren, and Zylbersztejn, 2015; Kataria and Winter, 2013).

Herein we generalize these results by exploring whether people who voluntarily sign a solemn truth-telling oath are more committed to sincere behavior when offered the chance to lie. We study the oath in two different contexts: a neutral context replicating the typical experiment in the literature, and a context in which subjects are reminded that a lie is a lie. To that end, this treatment introduces an additional non monetary device: moral reminders of ethical standards, by explicitly labeling untruthful communication as a "lie" and truthful communication as "truth". Evidence suggests that recalling moral standards when subjects take decisions can significantly downplay dishonest behavior (see,e.g. Ariely, 2012, for a discussion on this issue). The explanation of this observation is that people in situations in which dishonesty pays have on the one hand a desire for monetary gain, but on the other hand they also wish to preserve a positive self-view, i.e. not to be a "liar" – a self-concept maintenance (Mazar, Amir, and Ariely, 2008). Providing moral reminders to people give them less room to rationalize dishonesty and make them choose more in line with truth-telling behavior.

We study the two non-monetary devices in a 2*2 design that combines the oath and a loaded environment. Compared to typical promises experiments, subjects are committed endogenously to tell the truth by individually deciding whether to make a promise or not. With the oath-procedure, subjects are freely made to commit to tell the truth, that is, they are committed exogenously to truth-telling.[2] It allows to measure the causal effect of commitment on truth-telling motivation unlike in promises experiments where the causal relationship between promises and cooperation can be both ways. The oath also implements commitment to truth telling before the game takes place, rather than in-game promises. Last, it is a real-world and easy to implement procedure, which provides guidance into the design of institutional devices aimed to foster truth-telling. We use Erat and Gneezy's (2012) white lies experimental game. This game provides incentivized observations of truth-telling through the choice of a signal sent to a receiver, which can be either consistent or inconsistent with the sender's private information. We control for decision time in all conditions to assess the extent of cognitive reasoning in decision making (Kahneman, 2003; Rubinstein, 2007).

Our results provide a clear answer to our initial question: Does a solemn oath enhance sincere behavior in truth-telling situations? The answer is "yes" if the social context makes people be explicitly aware that their insincere behavior is explicitly named as lying: they lie half less frequently relative to those who were not asked to sign a truth-telling oath. But the answer is "no" if people are not explicitly told they would be considered "liars". Without an explicit reference to the social context of the decision, the oath had no effect on the frequency of lies in our game. The oath did, however, affect decision times when a lie is involved even in the neutral environment. After

Stevens, Tabatabaei, and Lass (2013); de Magistris and Pascucci (2014).

[2]As explained in more detail below, the oath is freely taken and signed but is designed in such a way that virtually all subjects comply.

signing the oath, the average person in the neutral environment took more time to decide when they lie to the other player, which indicates that the oath makes the person ponder their choice to lie a little longer. Subjects under oath in the neutral environment seem to be aware that they were lying.

Our results complement the growing literature on lying behavior in strategic communication and the ways to undermine it. Gneezy, Meier, and Rey-Biel (2011) who illustrate how non-monetary incentives can work to contradict monetary incentives in several situations. Our results suggest that the social context can make decisions to lie more difficult. We observe that a loaded context induces a small reduction in the share of liars. In that context, commitment through the oath strongly fosters subjects' willingness to tell the truth. Our preferred interpretation of these results is that social context makes it harder to rationalize untruthful decisions as compared to a neutral context. The oath strongly reinforces the effect as it makes self-maintenance more salient in the decision making process.

2 Commitment and the social context: insight from economics and social psychology

The truth-telling oath has deep roots in the experimental literature related to the social psychology theory of commitment. In social psychology, commitment is defined as a "binding of the individual to behavioral acts". Commitment is obtained through preparatory actions, which are purposefully designed to induce a predictable change in subsequent decisions. In a seminal field experiment, Harris (1972) asked people for the time before asking them for a dime. People were 4 times more likely to give him a dime if they were first asked to give away the "time" for free – 44% time-dime vs 11% no time-dime. In this experiment, asking for the time was a preliminary, costless, altruistic deed that helped commit a person to a similar and costly second target action. This method allows people to comply freely to what is expected of them. Social psychologists have explored several such "compliance without pressure" institutions to induce commitment from people (see the review in Joule, Girandola, and Bernard, 2007; Cialdini and Sagarin, 2005).

The behavioral mechanism that these procedures have in common is that compliance with demanding requests (*e.g.*, tell the truth) is improved significantly by using a process that first gets a person to commit to being the type of person that does a certain action (*e.g.*, signing an oath), as long as the course of action remains consistent and compliance to the preparatory action is free. Preparatory actions need to be made in specific commitment conditions to be effective. This means that the initial request does not need to be explicit to trigger the expected behavior but it must be consistent with the target decision to produce a change in behavior – *i.e.*, commitment have to be meaningful to decision makers in the target situation. In a series of experiments, Beauvois and Joule (2002); Joule, Py, and Bernard (2004) show for instance how it is possible to induce people to behave in a more honest way than they spontaneously would in normal circumstances

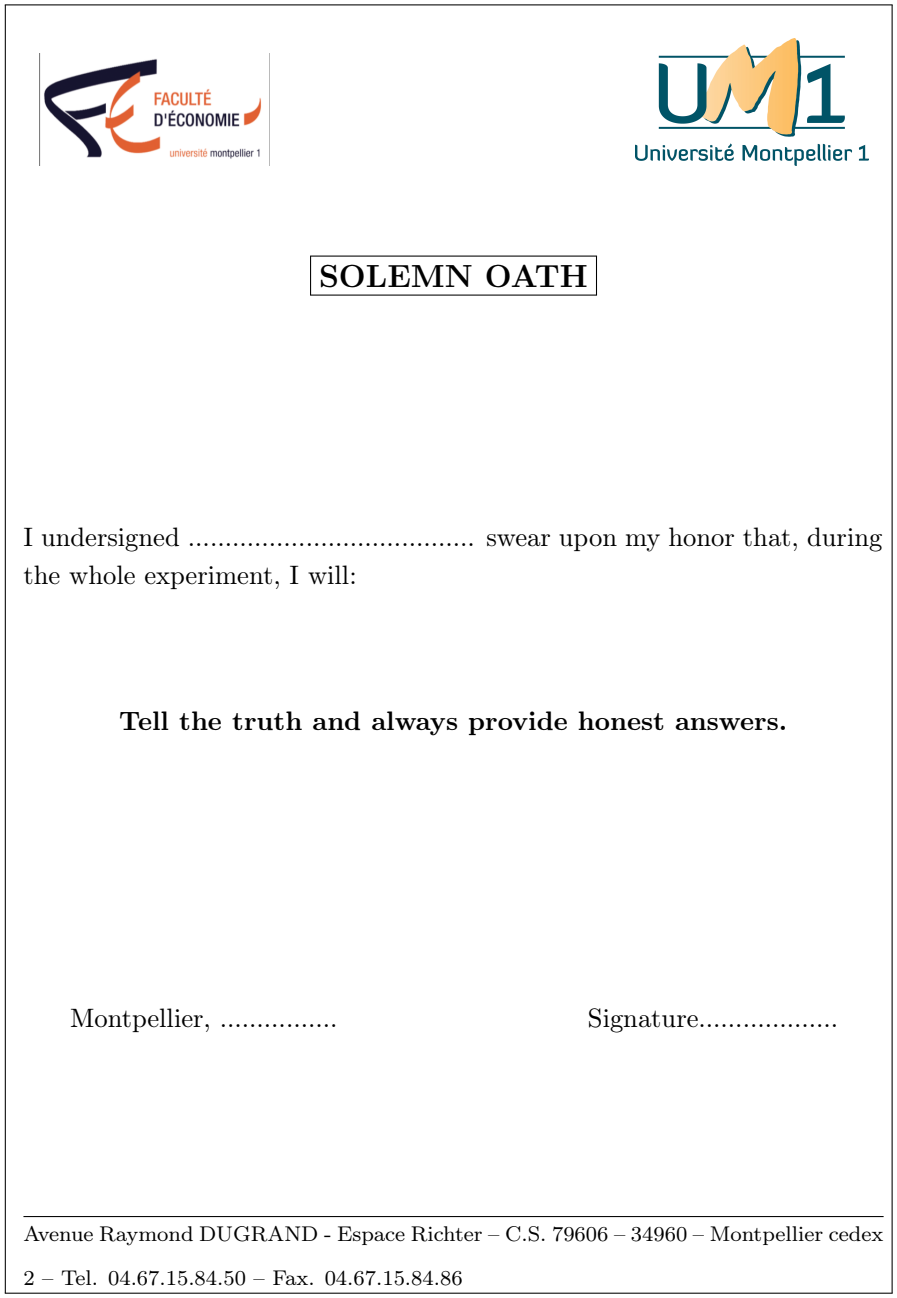
just by asking them to do small favors. In one of these experiments, the subject was asked for a piece of information by the experimenter (small favor). Shortly after, another experimenter lost a bank-bill in front of the same subject: as compared to the control condition without any prior interaction, subjects who have been asked for a small favour signaled significantly more often to the experimenter that she has lost his money. In our setting, this means that a truth-telling oath cannot change honesty if decisions to be taken are not perceived as related to honesty/lying by the decision maker.

Under these conditions, commitment induces strong and lasting changes in behavior (as shown in field experiments by, *e.g.*, Freedman and Fraser, 1966; Pallack, Cook, and Sullivan, 1980; Wang and Katsev, 1990; Katzev and Wang, 1994; Joule, Girandola, and Bernard, 2007; Guéguen, Joule, Halimi-Falkowicz, Pascual, Fischer-Lokou, and Dufourcq-Brana, 2013). Social psychologists have observed that commitment is stronger if it has been made freely, publicly expressed, or if it is unambiguous and costly for the subject (in time, energy, money, ...), or has consequences for the subject, or all conditions hold. Building on this literature, our commitment device is a truth-telling oath first designed by Jacquemet, Joule, Luchini, and Shogren (2013). Subjects are presented with the form displayed in Figure 1 and the procedures (see Section 3 for a detailed description) guarantee that is it taken freely and signed. Based on the social psychology of commitment, and the results from previous works, we expect the oath to induce people to be consistent with their initial commitment to tell the truth in subsequent decisions in the experimental games.

One may however argue that an oath is not per se a commitment device. Rather the oath affects behavior because it contains "moral reminders" and provides cues about the social context in which subjects are taking decisions (Haley and Fessler, 2005; Bateson, Nettle, and Roberts, 2006; Rigdon, Ishii, Watabe, and Kitayama, 2009). Recent experimental findings suggest that social cues are not only a substitute for learning/experience (Cooper and Kagel, 2003), but also affect subjects' reasoning process (Cooper and Kagel, 2009). Self-awareness theory (Silvia and Duval, 2001) and self-concept maintenance theory (Mazar, Amir, and Ariely, 2008) put forward the idea that social cues in the form of reminders about what moral behavior is in the situation as stake make it harder for subjects to maintain a positive self-image while engaging in dishonest behavior. In such a context, some have argued that the intrinsic rewards from honesty (and the intrinsic penalty from dishonesty) are likely to be more significant because people are induced to think deeper about honesty and dishonesty in the course of their economic decisions. The context helps people think about moral standards in solving the trade-off between profiting from cheating and allowing oneself to believe he or she is a honest person. As stressed by Bénabou and Tirole (2011) this strengthens "the repudiation of intended untruthful actions from the point of view of the virtual judgments of imagined spectators" – fostering the feeling of guilt experienced by the decision maker.

To disentangle these two dimensions – to what extent the oath commits subject to truth-telling and whether or not commitment works due to its explicit reference to truth-telling – we consider

Figure 1: Oath form used in the experiment



FACULTÉ D'ÉCONOMIE
université montpellier 1

Université Montpellier 1

SOLEMN OATH

I undersigned ..... swear upon my honor that, during the whole experiment, I will:

Tell the truth and always provide honest answers.

Montpellier, ... Signature.....

Avenue Raymond DUGRAND - Espace Richter – C.S. 79606 – 34960 – Montpellier cedex 2 – Tel. 04.67.15.84.50 – Fax. 04.67.15.84.86

a second treatment variable that makes explicit the opportunities to lie in the game: decisions are "loaded" according to what they actually are, truth-telling or lies.

Table 1: Design of the baseline experiment

Label	Lie	Payoff configurations: Receiver guesses correct	Payoff configurations: Receiver guesses wrong	Sequences of plays (2 sessions each): Order 1	Order 2	Order 3
T[-5; 10]	Altruistic	(20;20)	(15;30)	1	2	3
T[10; 10]	Pareto	(20;20)	(30;30)	2	3	1
T[1; -5]	Selfish	(20;20)	(21;15)	3	1	4
T[-10;-10]	Reverse	(30;30)	(20;20)	4	4	2

Note. The four payoff configurations presented in row are implemented together in all experimental sessions in varying order, as described in the right-hand side column. Option X describes the payoff implemented if the receiver choice matches the dice roll– $(sender; receiver)$ – ; otherwise subjects earn the payoff described in Option Y.

3 Experimental design

Our baseline condition replicates the sender-receiver game originally proposed by Erat and Gneezy (2012). We study two new treatment variables: commitment through a truth-telling oath and loaded environment that provides subjects with moral reminders of ethical standards

3.1 Design of the baseline experiment

Two players, a sender and a receiver (in the instructions, the sender is player A and the receiver is B), are randomly matched. The computer draws a 6-sided dice, and only the sender is informed about the result. The sender is then asked to send a message to the receiver, which has to be one of the six available messages: "*The outcome of the roll of dice was* x*"*, in which x is any integer between 1 and 6. The sender faces no constraint on whether he or she sends the true outcome to the receiver. One message out of six (the one that matches the actual roll) is an experimental measure of truth-telling, the other five are observations of lies.

Once the message has been received, the receiver chooses a number in the set $[1, 2, 3, 4, 5, 6]$, which determines the payment of both subjects. There are two payment options, X and Y. Only the sender knows the actual payoffs generated by each option. If the number chosen by the receiver matches the dice roll, both subjects are paid based on option X; otherwise, Y is implemented. This is common knowledge to all subjects.

Erat and Gneezy (2009) defined 9 different payoff configurations to investigate the reaction function of the willingness to lie given different monetary incentives.[3] To assess the robustness and sensitivity of our own treatments to monetary incentives, we focus on a subset of 4 of the original

[3]The number of payoff configurations refers here to the working paper version, available for instance at: `http://www.stanford.edu/group/SITE/archive/SITE_2009/segment_7/segment_7_papers/gneezy.pdf`. Among the four payoff configurations we've chosen, 2 are presented in the published version of the paper (Erat and Gneezy, 2012), 2 are not. See Table 2 for a comparison between payoff configurations and across experiments.

payoff configurations, each corresponding to a particular type of lie (see Table 1). The label of each payoff configuration refers to the gain/loss of each player (the sender for the first number, the receiver for the second) when the receiver chooses a number other than the actual outcome of the roll of the dice, as compared to their payoff when the receiver chooses the actual roll. In T[-5;10], if the sender chooses to lie and the receiver chooses a number that follows sender's message, the lie is called *altruistic* since the receiver gains 10 from the lie but the sender loses 5. Following the same logic, T[1;-5] characterizes a *selfish* lie, whereas T[10;10] is a *Pareto* lie since both the sender and the receiver benefits from the sender sending a message that is not the real outcome of the dice. The fourth payoff configuration acts as the control configuration since both players lose from the sender's lie.

We introduce two main changes in the original design. First, while Erat and Gneezy (2009, 2012) paid only one pair of subjects out of 10 pairs, we pay each and every subjects based on actual decisions. Second, we implement all payoff configurations one after the other in each session, as opposed to between subjects configurations in Erat and Gneezy (2009) and the main treatment in Erat and Gneezy (2012) – this between subjects implementation has also been used, however, as a robustness treatment in Erat and Gneezy (2012, Section 4).[4] We use a perfect stranger design to rule out any incentive to invest in partner-specific reputation by rematching subjects across payoff configurations – *i.e.*, each subject meets any other at most once in the whole session, and this is common knowledge among subjects. Such a within-subjects design reduces the noise due to individual heterogeneity in the observed treatment effects, even though its' well-known drawback is to confound order and payoff configuration effects. We control for this dimension by implementing three different orders, chosen in such a way that each pay-off configuration appears first in each order (except for the reverse pay-off, which is only used as a control configuration) : (1) altruistic / Pareto / selfish / reverse , (2) selfish / altruistic / Pareto / reverse and (3) Pareto / reverse / altruistic / selfish. Individual earnings from the experiment are computed based on one randomly drawn decision out of the four, so as to avoid changes in expected wealth across payoff configurations. This core experimental game is played under two treatment variables, implemented according to a factorial design.

Except for these two changes, the wording and framing of the baseline experimental instructions closely follow Erat and Gneezy (2012). We label this condition the *neutral* treatment in that the game is described without using words such as "lie, truth, or (dis)honesty". Neutral wording is the natural way of testing the effect of monetary incentives on the willingness to lie while neutralizing the effect of the decision context. The neutral presentation of the game reads as follows:

[4]Our main reason for choosing the between subject design although it has been less frequently implemented in the literature is that we use Erat and Gneezy (2012) as a benchmark for our study and want to put most of the identification power of the experiment on our treatment variables. Since all our results are derived from comparisons with the baseline, this changes are controlled for in our analysis of the data.

Player A is first informed of the payoffs for two possible options: option X and option Y. These payoffs change from one period to another. Player A must then roll a six sided dice, and is the only participant to know the outcome of this roll. Once informed of the dice rolls' outcome, ***player A must choose a message to send to player B****.*

3.2 Treatment variable 1: loaded environment

Our first treatment variable creates a "loaded environment" in which we provide moral reminders. This is done by explicitly labeling untruthful communication as a "lie" and truthful communication as "truth" in the instructions. To keep the message simple, we add the following paragraph to the instructions:

> *Player A has a choice between: (1) being honest and telling the truth about the dice roll's outcome; or (2) lying and sending a message that is different from the real outcome. Please note that player B is only shown player A's message: If player A choose to lie (choice n^o 2), player B has no way of knowing it.*

This change in the wording of the experimental instructions is the only difference between the loaded treatment and the baseline neutral treatment.

3.3 Treatment variable 2: commitment through a truth-telling oath

In the oath treatment, subjects are offered to sign a truth-telling oath before entering the lab. Since only senders are about to be in a truth-telling situation, only half the subjects are offered the oath procedure – subjects are unaware at that time of the subsequent assignments to their roles, only the experiment is. Each subject enters alone and is directed to a monitor at the front of the laboratory. Each subject that will play as sender in the experiment is then offered a form to sign entitled "solemn oath" (see Figure 1). The monitor explicitly points out to the subject before she reads the form that she is free to sign the oath or not, and that participation and earnings are not conditional on signing the oath. Importantly, subjects are not informed about the topic of the experiment when asked to take the oath. The subject reads the form, which asks whether she agrees "to swear upon my honor that, during the whole experiment, I will **tell the truth and always provide honest answers**" (in bold in the original form). The word "oath" is written on the form and read by the subject, but never said aloud. Regardless of whether the subject signs the oath, he is thanked and invited to enter the lab. The exact wording used by the monitors to offer the oath to respondents was scripted to standardize the phrasing of the oath. One monitor stayed in the lab until all subjects had been presented with the oath, to avoid communication prior to the experiment. Subjects waiting their turn could neither see nor hear what was happening at the oath-desk. Again, the oath procedure before entrance in the laboratory is the only difference with the baseline experiment.

3.4 Experimental procedures

Upon arrival, subjects randomly draw a tag from a bag assigning them to a role: sender or receiver. This draw is performed in a separate room before entering to the laboratory. After drawing their

tag, receivers are asked to enter the lab and seat in front of the computer indicated on their tag. Only the senders are asked to sign a solemn oath (in the oath treatments) and then seat in the lab.

Once all subjects are seated in the room, written instructions are distributed and read aloud. Subjects then observe their type displayed on the screen, A (sender) or B (receiver), which remains the same for the whole experiment. The experiment then consists of four periods. The only difference between the periods is the payoff distribution used. At the beginning of each period, the subjects are paired (one player A, one player B in each pair) and they are informed that they will never play twice with the same partner. The experiment ends with a demographic questionnaire. One of the four decision round is randomly drawn and each player receives the amount in Euros corresponding to her gains in that round, plus a show-up fee equal to 5 Euros.

We combine our two treatment variables according to a 2*2 factorial design, resulting in four treatments: baseline (no oath, neutral framing), loaded frame without and under oath and last, oath in the baseline. All sessions took place in the experimental lab of LAMETA[5], France between March and June 2012. For each treatment, we ran six 20-subjects sessions (two for each order defined in Table 1): this generates 480 observations of sender's decisions from 120 different individuals. Overall 480 subjects participated to the experiment, 220 males and 260 females. 91.67% were students, among which 38.86% were likely to have some background in game theory due to their field of study.[6] Participants' average age is about 24 years old. No subject participated in more than one experimental session. Each session lasted about 50 minutes, with an average payoff of 11.70€ (including a 5€ show-up fee).

4 Results

For the sake of statistical analysis, we define lying as any decision in which player A deceives player B by sending a message which does not match the roll of the dice. We first benchmark the internal validity of our study by comparing the results to Erat and Gneezy (2009, 2012). We then move to our main treatment effects: loaded environment and commitment.

Table 2 provides a benchmark comparison of observed behavior in our baseline treatment (neutral, no oath) with previous evidence. Pooling all payoff configurations, we observe that a significant fraction of our subjects, 34.2%, are willing to lie. Lying is the more likely when *(i)* it is in the subject's monetary interest to lie (selfish lie = 41.7%) and *(ii)* when the other player also gains from the lie (Pareto lie = 68.3%). These results are in line with those obtained in both a classroom setting and over the internet by Erat and Gneezy (2009, 2012): the order in terms of intensity of lying is preserved in all instances: lies are more likely in the Pareto payoff

[5]Laboratoire Montpellierain d'Economie Théorique et Appliquée, `http://leem.lameta.univ-montp1.fr/`. The recruitment uses Orsee (Greiner, 2004).

[6]Disciplines such as economics, engineering, management, political science, psychology, mathematics applied in social science, mathematics, computer science, sociology.

Table 2: Comparison with Erat and Gneezy experimental results

(*Setting*)		Erat and Gneezy (2009) *(Internet)*	Erat and Gneezy (2012) *(Classroom)*		This paper *(Laboratory)*
Matching scheme		*Between*	*Between*	*Within*	*Within*
Payoffs	Type				
T[-1;10]	Altruistic	33%	33%	43%	-
T[-5;10]	Altruistic	29%	–	–	13%
T[1;10]	Pareto	46%	49%	66%	-
T[10;10]	Pareto	61%	65%	76%	68%
T[1; 0]	Pareto	41%	–	–	–
T[1;-1]	Selfish	32%	–	–	–
T[1;-5]	Selfish	20%	37%	52%	42%
T[10; 0]	Pareto	47%	52%	-	-
Reverse	Spite	6%	–	–	13%

Note. For each payoff configuration in row, the Table reports the share of senders who decide to lie in each experiment presented in column. The last column reports the results from the baseline treatment– neutral, no oath.

configuration, lower in the selfish condition, and (slightly) lower in the altruistic condition. The main qualitative difference is that our study produces less altruistic lies than previous experiments. In Erat and Gneezy (2009, 2012), 33% of subjects engage in altruistic lying behavior in between subjects configurations (subjects face only one payoff configuration) and 43% in the within subjects configuration (subjects face four payoff configurations). These results are obtained for an altruistic lie configuration with a low cost-benefit ratio, as it costs 1 to the sender for the receiver to gain 10. Our altruistic payoff configuration has only been implemented in Erat and Gneezy (2009) in a between subjects context and over the Internet, in which it results in a 29% share of liars – a figure lower than between subjects implementation, but still higher than the one we observe.

The left-hand side of Table 3 provides the results of a panel Logit regression with random effects that controls for payoff configuration, order effects and demographics. The dependent variable is whether the subject lies and three demographics are considered: gender, living as a couple and being religious. The reference person is an atheist single female observed in the selfish-lie payoff configuration. First, conditional statistics confirm that lying is significantly and less likely for altruistic and spite lies relative to selfish lie. Second, subjects are more likely to make Pareto lies. According to odds ratios, the probability of making a Pareto lie is three times greater than making a selfish lie. The probability of making an altruistic and a spite lie are 20% lower than a selfish lie. All three dummy variables are significant at a 1% threshold. Third, order effects are present in baseline: lying is fostered by facing a Pareto or an Altruistic lie first in the sequence. Only one demographic is significant: living in a couple.

Table 3: Random effect panel Logit regressions: baseline *versus* loaded treatment

	Baseline treatment (n = 240, ♯ subjects = 60)			**Baseline & Loaded treatments (n = 480, ♯ subjects = 120)**		
	Estimated Parameter	Odds-ratio	p-value	Estimated Parameter	Odds-ratio	p-value
Constant term	-.587	-	.339	-.294	-	.521
Payoff dummies						
Altruistic lie	-1.615	0.198	0.001	-1.910	0.148	0.000
Pareto lie	1.188	3.283	0.003	1.137	3.120	0.000
Spite lie	-1.615	0.198	0.001	-2.134	0.118	0.000
Order effect dummies						
Pareto first	1.084	2.957	0.012	0.882	2.417	0.005
Altruistic first	0.843	2.323	0.040	0.323	1.382	0.281
Demographics						
Male	-0.438	0.645	0.200	-0.663	0.530	0.013
Couple	0.765	0.465	0.035	-0.556	0.573	0.034
Religious	-0.001	0.998	0.987	-0.018	1.018	0.760
Treatment effect						
Loaded treatment	-	-	-	-0.486	0.615	0.047

Note. The endogenous variable is whether sender i sends a message which does not match the actual roll of the dice in the t^{th} game. Dummy variables control for order effect, payoff configurations, and demographics (the reference person is an atheist single female in selfish lie payoff configuration). The left-hand side regression is performed on data from neutral treatment (no oath) only, the right-hand side regression is performed on pooled data from both the neutral and the loaded framing (no oath) treatments.

4.1 Truth-telling in a loaded environment

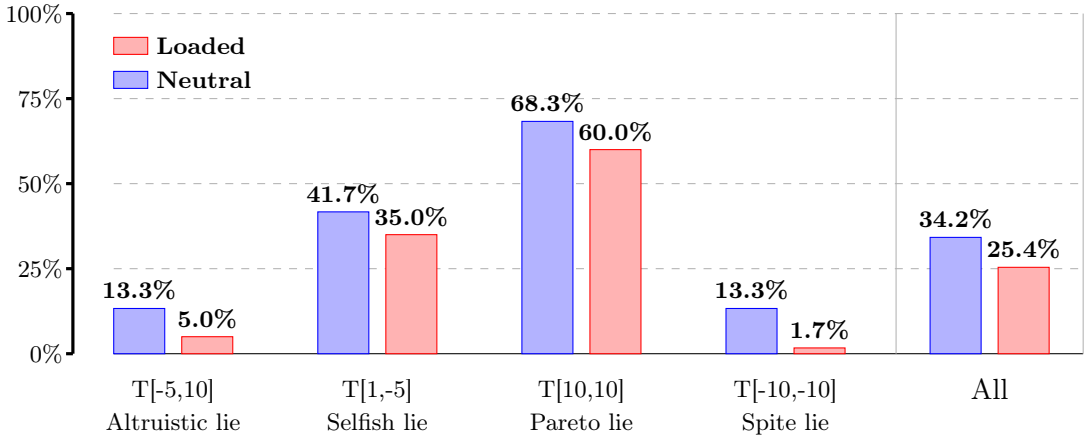
Figure 2 illustrates the effect of our first treatment variable through the fraction of subjects who lie by payoff configuration, in both the baseline and the loaded treatments. Overall, lying occurs in 25.4% of observed decisions, which is a 25% decrease in lying behavior in comparison to the baseline treatment. We test for the effect of loading the game by computing a bootstrap unconditional proportion test that accounts for within correlation.[7] According to this testing procedure, the decrease in lying behavior is significant, with $p = 0.022$. This allows us to state our first result:

Result 1 *Lying is less likely when subjects send signals in the loaded environment.*

Conditional regressions confirm the overall decrease in lying behavior. On the right-hand side of Table 3, we report the results from a panel Logit regression that pools the data from baseline and loaded treatments. A dummy variable that equals 1 if the subject was in the loaded treatment and 0 otherwise is introduced in the regression. The parameter associated with this variable is

[7]The correlation between decisions comes from the within subject implementation of the payoff configurations – uncorrected p-values mistakenly consider each individual decision as independent. The bootstrap procedure is implemented by bootstrapping subjects and their 4 behaviors in the sample rather than observations. We also allow for asymmetry in the empirical distribution of lying behavior (Davidson and MacKinnon, 2006).

Figure 2: Lying behavior in neutral and loaded environments



negative and significant at a 5% threshold. The odds ratio indicates that the probability to lie in baseline decreased by 61.5% when subjects face loaded instructions with explicit wording.

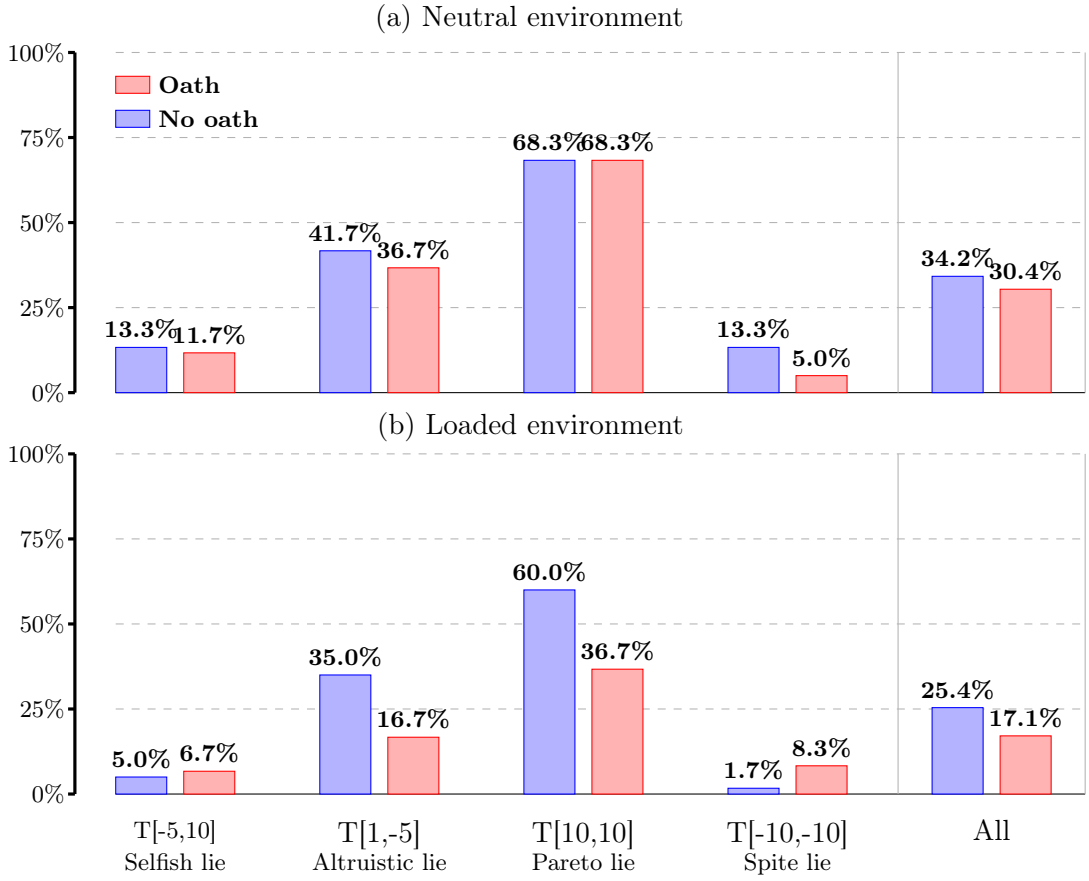
4.2 Truth-telling under oath

Figure 3 contrasts observed behavior in both the neutral and loaded environment according to whether subjects are offered to sign a truth-telling oath or not. The upper part of the figure shows the effect of the oath when the environment is neutral. The oath only marginally decreases lying behavior in the altruistic, selfish and spite payoff conditions, whereas it has no effect on Pareto lies. The bootstrap proportion test cannot reject the null of no difference in overall lying with and without the oath ($p = .631$) in the neutral treatment. Conditional regressions that pool data without and with oath are presented in Table 4. As in the regressions presented earlier, we control for demographics, payoff configuration and order effects. A dummy variable that equals 1 if the subjects was asked to take an oath prior to entering the lab and 0 otherwise is introduced. The results presented on the left-hand side, pooling data from the neutral environment and the neutral environment with oath, confirm that lying behavior is the same with and without oath in this context ($p = .263$).

Result 2 *Having subjects sign a truth-telling oath before participating in the neutrally framed lying game leaves truth-telling behavior unchanged.*

We now turn to the comparison of loaded treatments with and without oath. Overall, lying amounts for 17.1% of all messages sent when the oath is combined with a loaded environment. This corresponds to a 33% decrease as compared to lying behavior when only loaded environment is carried out. The decrease is significant according to a bootstrap proportion test ($p = .008$). In particular, we observe that Selfish lies decrease by 52.2% and Pareto lies decrease by 38.3%.

Figure 3: Changes in lying behavior due signing a truth-telling oath



Altruistic lies are slightly increased from 5.0% to 6.7%. Spite lies are more frequent than in heavy baseline but remain at a low level.

Result 3 *The oath decreases lying when lies are made explicit in the loaded environment.*

The Logit regression presented in the right-hand side of Table 4 pools data from the loaded environment and the loaded environment with oath and again controls for the lie type, order effects and demographics and includes a dummy variable that equals 1 if the subjects was asked to take an oath prior to entering the lab, 0 otherwise. The estimated parameter associated with this variable is significant at a 5% threshold with $p = .022$.

4.3 Treatment effects at the individual level

We further explore the effect of the treatment on lying behavior by looking more closely to the within individual dimension of our data. First, we compute a truth ratio for each subject, *i.e.*

Table 4: Random effect panel Logit regressions: subjects under oath

	Neutral & neutral oath treatments (n = 480, ♯ subjects = 120)			**Loaded & Loaded oath treatments (n = 480, ♯ subjects = 120)**		
	Estimated Parameter	Odds-ratio	p-value	Estimated Parameter	Odds-ratio	p-value
Constant term	-0.429	-	0.352	-1.136	-	0.034
Payoff dummies						
Altruistic lie	-1.594	0.202	0.000	-1.882	0.152	0.000
Pareto lie	1.301	3.675	0.000	1.140	3.127	0.000
Spite lie	-1.954	0.141	0.000	-2.050	0.128	0.000
Order effect dummies						
Pareto first	0.495	1.641	0.116	0.290	1.336	0.423
Altruistic first	0.538	1.713	0.088	-0.255	0.774	0.494
Demographics						
Male	-0.067	0.935	0.799	-0.236	0.789	0.451
Couple	-0.454	0.634	0.118	-0.076	0.926	0.811
Religious	0.017	0.982	0.771	0.083	1.086	0.253
Treatment effects						
Oath	-0.291	0.747	0.263	-0.707	0.492	0.022

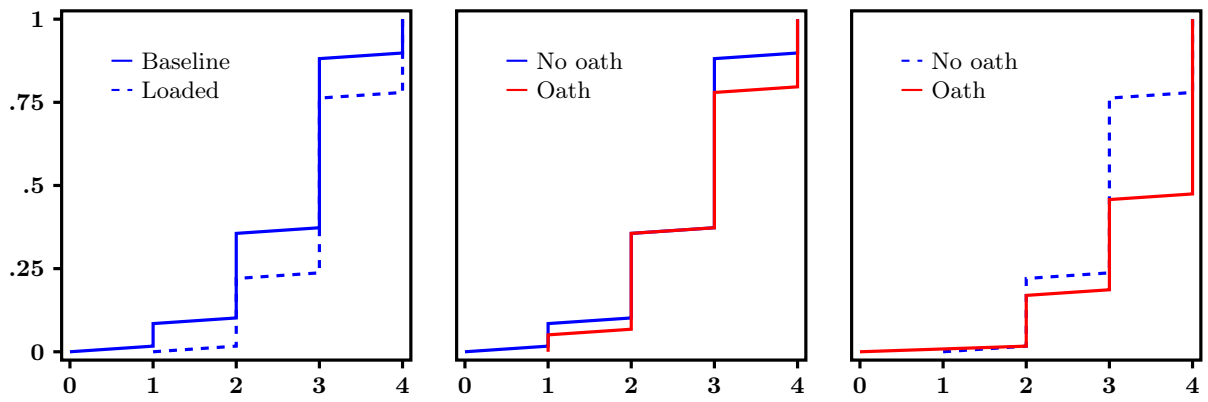
Note. The endogenous variable is whether sender i sends a message which does not match the actual roll of the dice in the t^{th} game. Dummy variables control for order effect, payoff configurations and demographics (the reference person is an atheist single female in selfish lie payoff configuration). The left-hand side regression is performed on pooled data from neutral treatments, the right-hand side regression is performed on pooled data from loaded treatments.

how many times out of the four observed decisions a subject tells the truth. Figure 4 provides two-by-two comparisons of the Empirical Distribution Functions (EDF) of the truth ratio for our treatments. Figure 4.a plots the EDF of the baseline along with the loaded framing (both with no oath). The EDF of loaded treatment first order dominates the EDF of the baseline at a 10% threshold, $p = .097$.[8] The two subsequent figures contrast behavior under oath with the no oath situation in the neutral and loaded treatment. In the neutral treatment (Figure 4.b), the only visible difference between the two curves is a small increase of truth ratios that equal 4, *i.e.*, subjects who always tell the truth. This increases is however not significant ($p = 0.269$). Consequently, the EDF from the oath(-neutral) treatment does not first order dominate the EDF from the baseline treatment ($p = 0.549$). In the loaded environment (Figure 4.c), by contrast, the oath unambiguously first order dominates the EDF of the no oath loaded environment – the difference is highly significant, with $p < .001$.

In Table 5, we reorganize observed behavior on an individual basis according to the lying pattern over all four treatments. For each treatment, the table provides the number of subjects the decisions of whom matches each of the 16 potential combinations of lies. In the baseline

[8]This result comes from a bootstrap version of the univariate Kolmogorov-Smirnov test. This modified test provides correct coverage even when the distributions being compared are not entirely continuous and, unlike the traditional Kolmogorov-Smirnov test, allows for ties (see Abadie, 2002; Sekhon, 2011).

Figure 4: Truth ratio by treatment



(a) No oath (b) Neutral (c) Loaded environment

treatment, we observe three main lying patterns: *(i)* No lie, 7 subjects (11.7%), *(ii)* Pareto lie only, 25 subjects (41.7%) and *(iii)* Selfish and Pareto lie, 14 subjects (23.3%). In the loaded environment, the number of subjects who engage in "nothing-but-the-truth" behavior, *i.e.* who never lie, happens to double. They are now 23.3% in comparison to 11.7% in the baseline. A bootstrap proportion test indicates that this change is however not significant, with $p = 0.189$. We observe smaller differences for subjects who choose to make a Pareto lie only or a selfish and Pareto lie. The main lesson from this comparison is that the overall decrease in lying observed in Result 1 cannot be attributed to a decrease in only one specific type of lie.

The comparison between oath and no oath treatments when instructions are neutral confirm that Result 2 applies to all lying patterns. The lying patterns are similar under both conditions, the dominant patterns under oath being no lie at all, only Pareto lie and selfish and Pareto lies only. Again, this is in sharp contrast with the variation in lying patterns when lies are made explicit in the instructions. When this is combined with an oath, we now observe that 32 subjects engage in truthful behavior for all four payoff configurations. Nothing-but-the-truth behavior equals 53.3% of the sample. This increase is significant at a 1% level in comparison to loaded environment without oath – a bootstrap proportion test yields a p-value $p = 0.003$. The decrease in Pareto lie only is significant with $p = 0.039$. We also observe less selfish lie only behavior: 8.3% of subjects make selfish and Pareto lies under oath whereas 18.3% do so without it – the decrease is however not significant, $p = .231$.

Now we pool all four treatments to jointly test Results 1, 2 and 3. We estimate a random effect panel Logit regression based on the 960 observations from the four treatments. Along with lie type, order effects and demographics, we introduce three dummies– one for each of the three results. A first dummy, *loaded treatment*, equals 1 (0 otherwise) for all subjects in the loaded environment – providing a robustness test of Result 1 to further conditioning. A second dummy, *oath neutral*, tests Result 2, equals 1 when respondents were in the neutral oath treatment and equals 0 otherwise.

Table 5: Lying patterns by treatment

	Lying patterns															
	← Nothing but the truth															All lies →
Altruist	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
Selfish	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
Pareto	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
Spite	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
	Number of subjects															
Neutral	7	0	25	1	4	1	14	0	2	0	0	0	0	5	0	1
Neutral oath	13	0	19	1	3	0	16	1	3	0	1	1	0	0	2	0
Loaded	14	0	23	0	8	0	11	1	1	0	1	0	1	0	0	0
Loaded oath	32	0	12	2	3	1	5	0	1	0	2	0	0	0	0	1

Note. The upper part of the table describes all possible lying patterns in all four games of the experiment – *i.e.* all combinations of sender's decision in each game. The bottom part describes for each treatment in row the number of subjects for whom the corresponding lying pattern has been observed.

Third, we test Results 3 by introducing a dummy variable, *oath-loaded treatment*, that equals 1 when the subject was in the loaded treatment with oath. Table 6 presents the estimations and confirms that all three results jointly hold when pooling the data: the parameter associated with the dummy variable *oath neutral* is not significant ($p = .420$) while parameters associated with the dummies *loaded environment* and *oath-loaded environment* are both statistically significant at a 5% level ($p = .044$ and $p = .014$).

4.4 Why the oath (sometimes) works: evidence from decision times

Our main findings are twofold. First, we find that if we "load" the decision problem with an explicit mention of its truth-telling content drastically affects a subject's willingness to lie – half of the original lies disappear. Second, a truth-telling oath strongly reinforces the non-monetary incentives to tell the truth – one third more being eliminated under oath. There are good reasons to expect an effect of each of the two non monetary devices on truth-telling behavior based on the literature in social psychology. It is more surprising, and puzzling, that the oath works only in the loaded environment. We now discuss possible reasons for these results given evidence on observed decision times.

Two hypothesis are consistent with the observed asymmetric effect of the oath across experimental environments. First, the effect of the oath in the loaded environment confirms a key insight from social psychology: subjects committed to tell the truth through a truth-telling preparatory act are more willing to comply with it in an explicit lying environment. This suggests it could be that subjects do not realize they are lying in the neutral condition. If so, the truth-telling prescription from the oath is meaningless. In social psychology terms, the behavioral prescription induced by the preparatory action does not match the target behavior – technically, the oath and

Table 6: Joint test - Random effect panel Logit regression

	All treatments (n = 960, ♯ subjects = 240)		
	Estimated Parameter	Odds-ratio	p-value
Constant term	-.520	-	0.168
Payoff dummies			
Altruistic lie	-1.712	.180	0.000
Pareto lie	1.222	3.396	0.000
Spite lie	-2.007	.134	0.000
Order effect dummies			
Pareto first	.405	1.500	0.091
Altruistic first	.141	1.152	0.555
Demographics			
male	-1.731	.840	0.394
couple	-.232	.792	0.279
religious	.032	1.032	0.484
Treatment effects			
Loaded treatment	-.545	.579	0.044
Neutral oath	-.215	.805	0.420
Oath-loaded treatment	-.715	.488	0.014

Note. The endogenous variable is whether sender i sends a message which does not match the actual roll of the dice in the t^{th} game. Dummy variables control for order effect, payoff configurations and demographics (the referent person is an atheist single female in selfish lie payoff configuration). The regression is performed on pooled data from all four treatments.

the game are not at the same 'level of identification" (Joule and Beauvois, 1998; Joule, Girandola, and Bernard, 2007). Our first hypothesis is that although subjects are committed to truth-telling, there is no such thing involved in the decisions to choose a message as subjects do not perceive the mismatch as a lie.

The second explanation, also from social psychology, is illustrated by the famous *Stanford jail experiment* (Zimbardo, 2007). In Stanford, in 1971, students were randomly selected to be either guards or prisoners in what students knew to be a fake jail.[9] After only a few days, Zimbardo and his colleagues stopped the experiment before the end because the guards were adopting dangerous and psychologically harming behavior. One of Zimbardo's conclusion from this failure is that situational forces are crucial to generate detrimental behavior. One important driving force is that a legitimate authority explicitly banishes, or implicitly allows, certain types of behavior.[10] Because

[9]Students were selected for their maturity and stability using personality questionnaires and interviews.

[10]On a side note, Zimbardo was an expert in the Abou Ghraib case, where he greatly insisted on the fact that although torture was prohibited in the military code, the military hierarchy was asking for results in terms of obtaining information from the prisoners – without insisting on how the information was to be obtained. In Ariely (2012)'s terms, there were no "moral reminder" about torture.

the neutral condition does not mention either truth-telling or honesty, subject may consider that lying is implicitly "allowed" – *i.e.* sending a false message to obtain more money is not perceived to go against any norm, at least in the laboratory. In other words, the neutral environment gives more room for subjects to rationalize lying behavior.

The two explanations are consistent with the oath affecting behavior in the loaded environment while letting it unchanged in the neutral context. They are observationally identical in the game. One important difference between the two, though, is whether subjects recognize their actions as lies when they decide on a message. According to the first explanation, subjects do not realize that some actions in the game are "lies": the oath is ineffective because it is irrelevant. According to the second explanation, they realize what actions mean in terms of honesty, but they view lying *in that way* as something allowed by the experimenter – making the oath ineffective. This difference has testable consequences on the time subjects take to choose their message. Response times has been promoted as a tool to better understand how people make decision (Rubinstein, 2007, 2012). In line with the System 1 / System 2 hypothesis put forward by Kahneman (2003), "*choices made instinctively, that is, on the basis of an emotional response, require less response time than choices that require the use of cognitive reasoning*" (Rubinstein, 2007, p.1243). Such a pattern produces a strong correlation between mistakes and short decision times in decision problems for which there is an unambiguous solution.[11] In settings close to ours, longer decision times have been observed in social preferences games when subjects face higher social dilemmas – such as, *e.g.*, higher pay-off inequalities in a Dictator Game (Piovesan and Wengstrom, 2009). Similarly, when emotions linked to instinctive behavior are at stake, it takes more time for subjects to go against the emotionally-charged instinctive action (Greene, Sommerville, Nystrom, Darley, and Cohen, 2001).

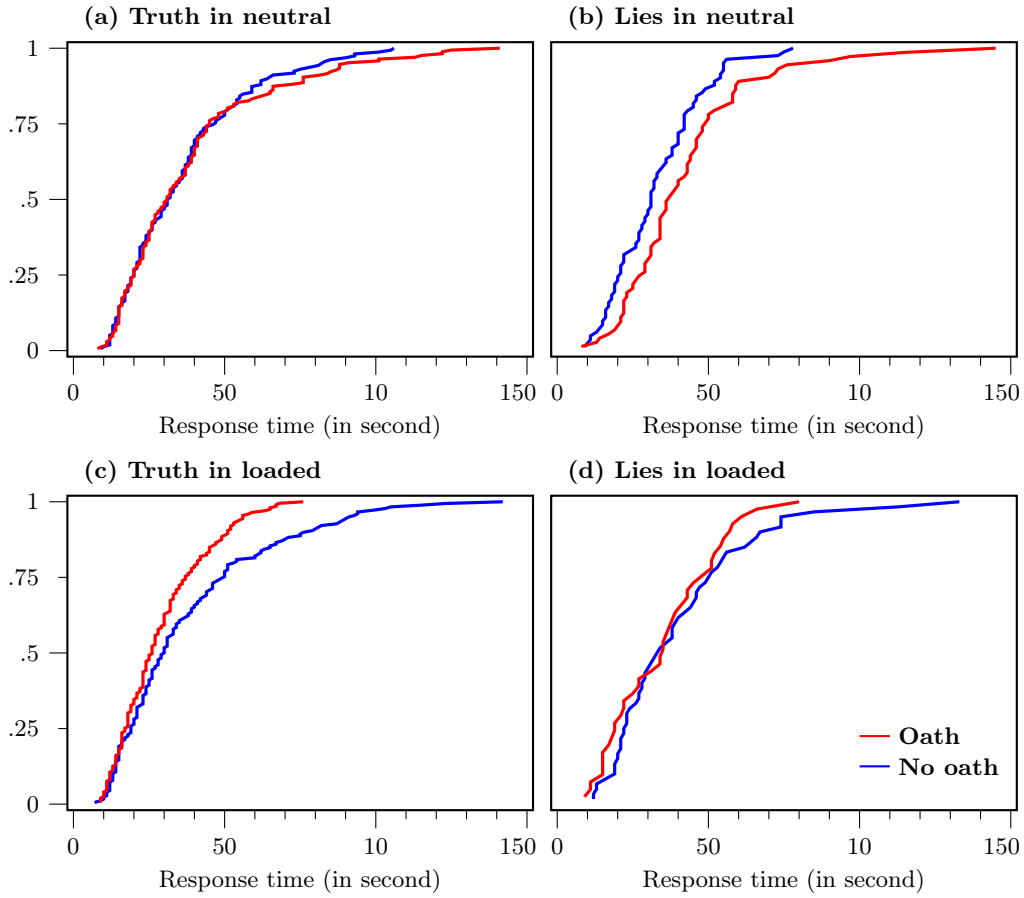
Our working hypothesis is the following: If the not-knowing-you-are-lying explanation holds, response time in the neutral environment should be of comparable magnitude whether subject are under oath or not–the oath should make no difference because subjects would not consider untruthful communication as "lies" in the neutral environment. In contrast, longer response time under oath would favour the second explanation: subjects under oath face a greater social dilemma.

Following a standard definition in the literature, we measure response times as the time elapsed from the moment the choice is presented to the subject on the computer screen until the choice is made. In Figure 5, we report the Empirical Distribution Function of response time by treatment and type of behavior. The main observations are two-fold. First, the top panel reports the behavior of subjects in the neutral condition. While the curves almost perfectly coincide on the left-hand side figure,[12] the right-hand side shows a sharp difference: subjects take significantly more time to decide to send an untruthful message when under oath. The EDF of response time under oath first order dominates that of response time in neutral without oath ($p = .011$). The increase in

[11] In more complex environment, such as decision under risk, short decision time is not correlated with choice anomalies, suggesting that such anomalies are not "mistakes".

[12] We observe no change in response time for honest decisions ($p = .544$).

Figure 5: Empirical Distribution Function of Response times by treatment



median response time amounts to six seconds: the decision to send a lie in the neutral-no oath condition takes 31 seconds whereas it takes 37 seconds under oath – a 20% increase in response time. Under the assumption that higher response times are associated with strengthened social dilemmas, this suggests that the truth-telling prescription of the oath is meaningful for subjects who decide to send an untruthful message in the neutral condition, even though the oath has no behavioral consequence. Subjects have to solve a stronger social dilemma in the game when they previously complied with a truth-telling oath, as compared to no oath, which seems to generate greater difficulties to overcome the oath prescription. This tends to rule out the possibility of a mismatch between the oath behavioral prescription and the behavior emphasized in the game ; and rather favors the second interpretation: the oath is meaningful to subjects, but it is ineffective due to the norms suggested by the neutral condition.

This is reinforced by the pattern of response times in the loaded environment, reported in the bottom panel. While response time of liars are both observationally and statistically ($p = .610$) the same, honest subjects in this loaded condition take less time to answer when under oath. The EDF of response time in loaded environment-no oath first order dominates that of response time

in loaded environment-oath ($p = .002$). Median response time in the loaded environment-no oath condition is 30 seconds while median response time is 26 seconds in the oath condition, inducing a 13.3% decrease. Again, under the assumption of a positive correlation between response time and the intensity of the dilemmas subjects have to solve, this tends to suggest that loading the instructions with explicit wording removes any ambiguity as to what is the norm in the game, and indicates that honesty matters. Such correlations are consistent with the second interpretation: the instructions in the neutral conditions suggests that lying *in that way* is allowed in the lab and, although surprising at first glance, the oath is ineffective on lying. In the loaded treatment, by contrast, the oath eases the decision to tell the truth because subjects face a lower social dilemma in the course of sending a truthful message.

5 Conclusion

Herein we explore the influence of two kinds of non-monetary incentives to tell-the-truth: loaded environment and commitment. Both devices appear influential on the willingness to tell the truth. When the game is described to explicitly refer to truth/lie-telling, we observe one half less lies. In a neutral framing, we do not observe any improvement in truth telling after having subject sign an oath to tell the truth before entering the laboratory. It is only when this commitment device is associated with reminders of ethical standards, that the oath drastically improves truth-telling behavior – with a further one third drop in the share of liars.

These results are in line with recent findings from psychology. First, in accordance with Mazar, Amir, and Ariely (2008) self-concept maintenance theory, we find a strong effect of internal rewards to honesty when people are induced to think in terms of dis/honesty in taking their decisions. To strengthen the interpretation of these results, we provide evidence based on response times, showing that subjects decide more quickly to *(i)* tell the truth under oath in the loaded environment and *(ii)* to lie with no oath in the neutral environment. Under the assumption that higher response times are associated with more intense social dilemmas, these correlations suggest that the oath is ineffective in the neutral condition not because it is misaligned with the game, but rather because the neutral condition suggests that lying does not contradicts the norms implemented in the laboratory. Non-monetary incentives implemented through an oath do have a strong impact on lying behavior, but social context matters for commitment.

References

Abadie, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Model," *Journal of the American Statistical Association*, 97(457), 284–292.

Ariely, D. (2012): *The Honest Truth About Dishonesty: How We Lie to Everyone–Especially Ourselves*. HarperCollins Publishers, New York.

Bateson, M., D. Nettle, and G. Roberts (2006): "Cues of being watched enhance cooperation in a real-world setting," *Biology Letters*, 2(3), 412–414.

Beauvois, J.-L., and R.-V. Joule (2002): *Petit traité de manipulation à l'usage des honnêtes gens*. Presses Universitaires de Grenoble.

Boles, T. L., R. T. A. Croson, and J. K. Murnighan (2000): "Deception and Retribution in Repeated Ultimatum Bargaining," *Organizational Behavior and Human Decision Processes*, 83(2), 235–259.

Brandts, J., and G. Charness (2003): "Truth or Consequences: An Experiment," *Management Science*, 49(1), 116–130.

Bénabou, R., and J. Tirole (2011): "Identity, Morals, and Taboos: Beliefs as Assets," *Quarterly Journal of Economics*, 126(2), 805–855.

Cappelen, A. W., E. A. Sorensen, and B. Tungodden (2013): "When do we lie?," *Journal of Economic Behavior & Organization*, 93(1), 258–265.

Carlsson, F., M. Kataria, A. Krupnick, E. Lampi, A. Lofgren, P. Qin, T. Sterner, and S. Chung (2013): "The Truth, the Whole Truth, and Nothing but the Truth - A Multiple Country Test of an Oath Script," *Journal of Economic Behavior & Organization*, 89, 105–121.

Charness, G., and M. Dufwenberg (2006): "Promises and Partnership," *Econometrica*, 74(6), 1579–1601.

Cialdini, R., and B. Sagarin (2005): "Interpersonal influence," in *Persuasion: Psychological insights and perspectives*, ed. by T. Brock, and M. Green, pp. 143–169. Sage Press, Newbury Park (CA).

Cooper, D. J., and J. H. Kagel (2003): "The impact of meaningful context on strategic play in signaling games," *Journal of Economic Behavior & Organization*, 50(3), 311–337.

Cooper, D. J., and J. H. Kagel (2009): "The role of context and team play in cross-game learning," *Journal of the European Economic Association*, 7(5), 1101–1139.

Crawford, V. P., and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50(6), 1431–1451.

Davidson, R., and J. G. MacKinnon (2006): "Bootstrap methods in econometrics," in *Palgrave Handbooks of Econometrics*, ed. by T. C. Mills, and K. D. Patterson, vol. 1, chap. 23, pp. 812–838. Palgrave Macmillan, Basingstoke.

de Magistris, T., and S. Pascucci (2014): "Does "solemn oath" mitigate the hypothetical bias in choice experiment? A pilot Study," *Economics Letters*, 123(2), 252–255.

Erat, S., and U. Gneezy (2009): "White Lies," *Rady School of Management Working paper*.

——— (2012): "White Lies," *Management Science*, 58(4), 723–733.

Fischbacher, U., and F. Heusi (2013): "Lies in Disguise. An experimental study on cheating," *Journal of the European Economic Association*, 11(3), 525–547.

Freedman, J. L., and S. Fraser (1966): "Compliance without pressure: the foot-in-the-door technique," *Journal of Personality and Social Psychology*, 4(2), 195–202.

Gneezy, U. (2005): "Deception: The Role of Consequences," *American Economic Review*, 95(1), 384–394.

Gneezy, U., S. Meier, and P. Rey-Biel (2011): "When and Why Incentives (Don't) Work to Modify Behavior," *Journal of Economic Perspectives*, 25(4), 191–210.

Greene, J. D., R. B. Sommerville, L. E. Nystrom, J. M. Darley, and J. D. Cohen (2001): "An fMRI Investigation of Emotional Engagement in Moral Judgment," *Science*, 293(5537), 2105–2108.

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments.," *University of Cologne, Working Paper Series in Economics*, 10, 79–93.

Grossman, S. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24(3), 461–83.

Guéguen, N., R.-V. Joule, S. Halimi-Falkowicz, A. Pascual, J. Fischer-Lokou, and M. Dufourcq-Brana (2013): "I'm Free but I'll Comply With Your Request : Generalization and Multidimensional Effects of "Evoking freedoom" Technique," *Journal of Applied Social Psychology*, 43(1), 116–137.

Haley, K. J., and D. M. T. Fessler (2005): "Nobody's watching? Subtle cues affect generosity in an anonymous economic game," *Evolution and Human Behavior*, 26(3), 245–256.

Harris, M. B. (1972): "The Effects of Performing One Altruistic Act on the Likelihood of Performing Another," *Journal of Social Psychology*, 88(1), 65–73.

Houser, D., S. Vetter, and J. Winter (2012): "Fairness and cheating," *European Economic Review*, 56(8), 1645–1655.

Jacquemet, N., A. James, S. Luchini, and J. Shogren (2015): "Referenda under oath," *GREQAM WP, revised*, 10(15).

Jacquemet, N., R.-V. Joule, S. Luchini, and J. F. Shogren (2013): "Preference Elicitation under Oath," *Journal of Environmental Economics and Management*, 65(1), 110–132.

Jacquemet, N., S. Luchini, J. Shogren, and A. Zylbersztejn (2015): "Coordination with Communication under Oath," *GREQAM WP, Revised*, 11(49).

Joule, R., and J. Beauvois (1998): *La soumission librement consentie*. Presses Universitaires de France, Paris.

Joule, R.-V., F. Girandola, and F. Bernard (2007): "How Can People Be Induced to Willingly Change Their Behavior? The Path from Persuasive Communication to Binding Communication," *Social and Personality Psychology Compass*, 1(1), 493–505.

Joule, R. V., J. Py, and F. Bernard (2004): "Qui dit quoi, à qui, en lui faisant faire quoi? Vers une communication engageante," in *Psychologie Sociale et Communication*, ed. by M. Bromberg, and A. Trognon, pp. 205–218. Dunod, Paris.

Kahneman, D. (2003): "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 93(5), 1449–1475.

Kartik, N. (2009): "Strategic Communication with Lying Costs," *Review of Economic Studies*, 76(4), 1359–1395.

Kataria, M., and F. Winter (2013): "Third Party Assessments in Trust Problems with Conflict of Interest: An Experiment on the Effects of Promises," *Economics Letters*, 120(1), 53–56.

Katzev, R., and T. Wang (1994): "Can commitment change behavior ? A case study of environmental actions," *Journal of Social Behavior and Personality*, 9(1), 13–26.

Kriss, P. H., R. Nagel, and R. A. Weber (2013): "Implicit vs. explicit deception in ultimatum games with incomplete information," *Journal of Economic Behavior & Organization*, 93(1), 337–346.

Mazar, N., O. Amir, and D. Ariely (2008): "The dishonesty of honest people: A theory of self-concept maintenance," *Journal of Marketing Research*, 45(6), 633–644.

Milgrom, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12(2), 380–391.

Pallack, M., D. Cook, and J. Sullivan (1980): "Commitment and energy conservation," in *Applied Social Psychology Annual*, ed. by L. Bickman, pp. 235–253. Beverly Hills, CA: Sage.

Piovesan, M., and E. Wengstrom (2009): "Fast or fair? A study of response times," *Economics Letters*, 105(2), 193–196.

Rigdon, M., K. Ishii, M. Watabe, and S. Kitayama (2009): "Minimal social cues in the dictator game," *Journal of Economic Psychology*, 30(3), 358–367.

Rubinstein, A. (2007): "Instinctive and Cognitive Reasoning: A Study of Response Times," *Economic Journal*, 117(523), 1243–1259.

——— (2012): "Response Time and Decision Making: A "Free" Experimental Study," *Working Paper*.

Sekhon, J. (2011): "Multivariate and Propensity Score Matching Software with Automated Balance Optimization," *Journal of Statistical Software*, 42(7), 1–52.

Silver, M. (1995): *Economic Structures of Antiquity*. Greenwood Press, Westport, CT.

Silvia, P. J., and T. S. Duval (2001): "Objective Self-Awareness Theory: Recent Progress and Enduring Problems," *Personality and Social Psychology Review*, 5(3), 230–241.

Stevens, T., M. Tabatabaei, and D. Lass (2013): "Oaths and hypothetical bias," *Journal of Environmental Management*, 127(1), 135–141.

Sutter, M. (2009): "Deception Through Telling the Truth?! Experimental Evidence From Individuals and Teams," *Economic Journal*, 119(534), 47–60.

Wang, T., and R. Katsev (1990): "Groupe commitment and resource conservation : two field experiments on promoting recycling," *Journal of Applied Social Psychology*, 20(4), 265–275.

ZIMBARDO, P. (2007): *The Lucifer effect. Understanding How Good People Turn Evil.* Random House, New York.