Lying Aversion and the Size of the Lie*

Uri Gneezy, Agne Kajackaite, and Joel Sobel†

Abstract

This paper studies lying in a simple framework. An agent first randomly picks a number from a known distribution. She can then claim to have observed any number from the set, receiving a monetary payoff based only on her report. Consistent with previous findings, our participants do not maximize monetary payoff by making the maximal claim dishonestly. The paper posits that this behavior is the result of lying costs and discusses different kinds of lying cost. The paper presents a model of lying costs that is used to generate hypotheses regarding behavior in the experiment. In line with the model, we find that the highest fraction of lies is by reporting the maximal outcome. Reputational concerns matter: More participants lie partially when their outcomes cannot be observed by the experimenter than when the experimenter can later verify the actual outcome, and partial lying increases when the highest outcome is ex ante unlikely. In contrast, the fraction of subjects who lie does not depend on how outcomes are labeled.

Journal of Economic Literature Classification Numbers: D03; C90; C72. Keywords: lying; deception; experiments; behavioral economics.

*We thank Johannes Abeler, Martin Dufwenberg, Kiryl Khalmetski, Daniele Nosenzo, Collin Raymond, Dirk Sliwka, and seminar participants at the University of Arizona for comments. This research was conducted with financial support of the German Science Foundation (DFG) through the research unit "Design & Behavior" (FOR 1371) and Behavioral Economics Research Spearhead Grant (University of Amsterdam). Sobel thanks NSF for financial support.

†Rady School of Management, University of California, San Diego, La Jolla, CA 92093, U.S.A. and CREED, 1018 WB Amsterdam, NL. E-mail: ugneezy@ucsd.edu; Los Angeles Behavioral Economics Laboratory, University of Southern California, Los Angeles, CA 90098, U.S.A. Email: kajackai@usc.edu; Department of Economics, University of California, San Diego, La Jolla, CA 92093, U.S.A. E-mail: jsobel@ucsd.edu.

1 Introduction

Exchanging information is a critical part of daily life and, in particular, economic activity. Situations frequently arise in which people can lie about their private information. Although lying is common, there is compelling real-world and laboratory evidence that sometimes people avoid telling some lies that would increase their material payoffs.[1]

Standard economic incentives may induce agents to avoid lies in natural settings. Businesses may avoid making false claims because if caught they would face substantial penalties. Individuals in long-term relationships may resist opportunities to make short-term gains through lying in order to maintain profitable relationships. Nevertheless, the evidence suggests that there are intrinsic costs to telling lies. The paper studies intrinsic costs of lying.

Not all lies are equal. External sanctions strive to make the punishment fit the crime, so that, for example, the penalties for tax evasion will increase with the amount of unreported income. Presumably the intrinsic costs of lying also depend on the "size of the lie." The literature contains informal discussions of lying costs that identify different ways to measure the size of lies. For example, Mazar, Amir, and Ariely [15] and Fischbacher and Föllmi-Heusi [8] suggest that the marginal cost of a lie is increasing in the magnitude of a lie, leading to the prediction that individuals might lie a little bit, but not take full advantage of strategic opportunities.

We describe three different ways to measure the size of the lie: The payoff dimension (how much more one earns by lying); the outcome dimension (how far the outcome is from the actual result); and the likelihood of the lie (how likely is the outcome to actually happen). The paper presents a model of these costs, derives some implications of the model, and tests these predictions with laboratory data.

To better understand the difference between the three dimensions of lying costs, consider a game in which a person rolls a six-sided die and reports the outcome, being paid a dollar amount equal to the number she reports. In this setting, there are two possible reasons why reporting six when the actual outcome is four is a smaller lie than reporting six when the actual outcome is

[1] Examples include: Abeler, Becker, and Falk [1], Abeler, Nosenzo, and Raymond [2]), Cohn, Fehr, and Maréchal [3], Dreber and Johannesson [4], Erat and Gneezy [6], Fischbacher and Föllmi-Heusi [8], Gneezy [9], Lundquist, Ellingsen, Gribbe, and Johannesson [14], Mazar, Amir, and Ariely [15], Sutter [17], and Shalvi, Dana, Handgraaf, and De Dreu [16].

two. The first reason is reporting six when the true outcome is four leads to a payoff increase of two, which is smaller than the increase of payoff obtained by reporting six when the true outcome is two. This is the difference in the payoff dimension. The second reason is that the reported number six is closer to four than it is to two. This is the difference in the outcome dimension. In the example, it is not possible to distinguish the payoff dimension from the outcome dimension. In other cases, it is possible to distinguish outcome-based costs and payoff-based costs. For example, consider a game in which reporting a five results in a positive payoff, while reporting any other number results in a zero payment. In this game, reporting a five is a smaller lie on the outcome dimension when the outcome is four than when the outcome is two. However, from a payoff perspective, the two cases are the same.[2] To illustrate the third dimension of the size of the lie, consider a game in which the participant rolls an n-sided die and receives a positive payoff by reporting a five (and zero otherwise). In this case, reporting a five is a larger lie on the likelihood dimension the larger is n. That is, bigger lies are statements that are less likely to be true ex ante. In Section 2 we introduce a basic model that generates predictions regarding how sensitivity to the size of the lie on each of these three dimensions will affect behavior. We assume that utility is the sum of three terms: the direct monetary payoff, a term that depends directly on the true state and the report (and indirectly on the monetary payoffs associated with these reports), and a term that depends on the probability that an observer would believe the report is honest. The second term captures the outcome and payoff dimensions. The third term captures the likelihood dimension.

Without the likelihood dimension, the theoretical model is a straightforward decision problem. Adding the likelihood dimension complicates the analysis because it adds a strategic aspect.[3]

[2] The experiment by Lundquist, Ellingsen, Gribbe and Johannesson [14] is a good way to understand the outcome dimension. Their participants play a deception game (Gneezy [9]) in which first the Sender takes a test and then sends a message regarding the results of it to the receiver. The Sender receives a fixed positive payoff if the receiver believes that she passed a certain threshold in her test. Since there are only two possible payoffs (zero if not passing, and a fixed payment if passing), the size of the lie on the payoff dimension is constant. However, the size of the lie can be determined based on how close the sender's performance was to the actual threshold, which is how Lundquist et al. define the size of the lie: "We test whether the aversion to lying depends on the size of the lie (i.e. that the aversion to lying is stronger the further you deviate from the truth) . . . "

[3] Three recent papers, Abeler, Nosenzo, and Raymond [2], Dufwenberg and Dufwen-

In Section 2 we introduce the formal model. Section 3 contains the theoretical analysis. We provide conditions under which the model makes a unique prediction and describe some of the properties of equilibrium. We provide conditions on the lying cost function under which equilibrium involves a cutoff value. If the agent draws an outcome above the cutoff, she never lies. If she draws an outcome below the cutoff, she may lie. If she does, she makes a claim above the cutoff. Reputations for honesty are decreasing in the claim. There are always dishonest claims of the maximal value. There are partial lies in equilibrium if reputation is sufficiently valuable. Our most novel findings are qualitative results that capture the intuition that making the maximal outcome less likely ex ante increases the frequency of partial lies. Section 4 summarizes the theoretical findings in a way that motivates the experimental portion of the paper.

Section 5 describes the experimental design. Our experiments manipulate three things. First, we compare a game in which the subjects' outcomes are observed to one in which they are not observed. Second, we vary the way in which the outcomes are labeled, in order to understand the outcome dimension of costs. Finally, we vary the prior distribution in order to understand the effect of reputation on lying behavior. Section 6 describes the experimental results. We find that people lie and a large fraction of people who do lie make the maximum lie. People appear to be sensitive to reputation in two ways. They make more partial lies when no one observes their outcomes. This finding is consistent with reputation effects, because all lies lead to the worst possible reputation in the observed game, while reputation is typically decreasing in the size of the lie in the non-observed game. They also make more partial lies when the ex ante probability of the highest state decreases. Finally, we show that the fraction of reports that are dishonest does not depend on how outcomes are labeled, but the labels do influence the frequency at which subjects make partial lies.

2 Model

An agent's private information or type consists of a pair (i, t), where $i = 1, \ldots, N$ and $t \geq 0$. The agent's information (i, t) consists of the value she observes i (this can be thought of as the outcome of a roll of a die) and a

berg [5] and Khalmetski and Sliwka [12], introduce models that capture the likelihood dimension. We discuss these models after we present our formal results in Section 3.

preference parameter, t, which will be the fixed cost of lying. The quantities i and t are independently distributed. $p_i > 0$ is the probability that the agent receives i. $F(\cdot)$ is the cumulative distribution function of t. We assume $F(0) = 0$ so that $t > 0$ with probability 1. We also assume that $F(t) < 1$ for all t ($\lim_{t\to\infty} F(t) = 1$). Combined, these conditions guarantee that almost all agents find lying costly and that some agents find lying so costly that they will never lie. The agent makes a claim, k, which is also assumed to be a number between 1 and N. We call an agent honest if she reports k when her type is of the form (k, t); she is dishonest (and her report is a lie) otherwise. The agent receives a monetary reward of v_k if she reports k; if $i < j$, then $v_i < v_j$. The observer hears the agent's report and assesses the honesty of the agent using Bayes's Rule. Specifically, suppose that the probability that a strategic agent reports k given (i, t) is $s(k \mid i, t)$,[4] then the probability the observer places on the agent being honest is

$$\rho_k(s) = \frac{h_k}{h_k + r_k}, \tag{1}$$

where

$$h_k = \sum_{k=1}^{N} \int_0^{\infty} s(k \mid k, t)dF(t)p_k \tag{2}$$

is the probability that k is reported honestly and

$$r_k = \sum_{i\neq k}^{N} \int_0^{\infty} s(k \mid i, t)dF(t)p_i \tag{3}$$

is the probability that k is reported dishonestly. It must be that $r_k \geq 0$ for all k. If $h_k + r_k = 0$, then ρ_k is not defined. Our assumptions on the distribution of lying costs guarantee that $h_k > 0$. We assume that the agent's preferences depend on the probability that she is perceived to be honest ($\rho(\cdot)$), the true draw and the reported draw. Under these circumstances, a possible representation for the agent's preferences is

$$v_i - C(i, j, t) + \alpha\rho_j(s). \tag{4}$$

We assume that

$$C(i, j, t) = \begin{cases} 0 & \text{if } i = j \\ t + c(i, j) & \text{if } i \neq j \end{cases}.$$

[4] $s(k \mid i, t) \geq 0$ and $\sum_{k=1}^{N} s(k \mid i, t) = 1$ for all i and t.

We also assume that $c(\cdot)$ is nonnegative, $c(i,i) = 0$, weakly increasing in $\mid i - j \mid$ and $c(i,j) + c(j,k) \geq c(i,k)$. Finally we assume that $c(N-1,N) < v_N - v_{N-1}$. These conditions include as a special case a model in which there is a categorical cost of lying ($C(i,j) = 0$ if $i = j$ and otherwise the lying cost is positive and independent of i and j). We use condition $c(N-1,N) < v_N - v_{N-1}$ to avoid the uninteresting case in which the cost of lying is so great that no one wishes to make the highest report dishonestly.

The representation (4) is special. It leaves out factors that may be important. Distributional concerns may play a role, but there are no other active agents in our experiments. Other studies have investigated the role of norms, emotions, image, and promises in truth-telling behavior. The reputation term is a reduced form that captures some aspects of these concerns. Additive separability, homogeneous preferences over monetary payments, and risk neutrality may be important restrictions.[5] The assumption that $c(i,j) + c(j,k) \geq c(i,k)$ simplifies our analysis because it guarantees that if anyone dishonestly reports k, then no one who observes k will be dishonest. Assuming that t enters the cost function separably means that, conditional on wanting to lie, the preferences of type (i,t) do not depend on t. The notation suppresses the possible dependence of costs on the monetary payoffs v_i, we imagine that if v_j changes, the cost of claiming j dishonestly would change. The cost function $c(i,j)$ is defined for $i > j$, but we will show that in equilibrium no agent will report less than what she observes.

We analyze an equilibrium in this setting, which consists of strategies $s(k \mid i,t)$ such that

(a) $s(k \mid i,t) \geq 0$ for all k,i,t and $\sum_{j=1}^{N} s(j \mid i,t) = 1$ for all i and t.

(b) $s(k \mid i,t) > 0$ only if k maximizes (4) (with respect to j).

(c) $\rho_j(s)$ is computed using (1) and (3).

There exists t^* large enough such that all types of the form (i,t) with $t > t^*$ would prefer to report i instead of $j \neq i$ independent of ρ. The assumption $F(t) < 1$ for all t guarantees that there is a positive probability of honest agents and the denominator in (1) is strictly positive.

This is a finite game and consequently existence of equilibrium follows from standard arguments. In the next section we describe properties of the equilibrium.

[5] One can view v_k is measured in utils, so risk neutrality is not restrictive given separability and homogeneity.

3 Analysis

This section describes properties of equilibria. The first main property is that the reputations generated in equilibrium must be unique. We also show how equilibrium behavior depends on the prior distribution.

We denote strategies by s (or s', s''), the associated reputations by ρ (or ρ', ρ''), and the utility without lying costs of a report by W (or W', W'') so that $W_k = v_k + \alpha\rho_k$. The first result identifies a structural property of equilibrium strategies:

Lemma 1 *Suppose i, j, and k are distinct, if there exists t' such that $s(k \mid j, t') > 0$ for $k \neq j$, then $s(j \mid i, t) = 0$ for all $t > 0$.*

In words, Lemma 1 states that if some agent finds it valuable to lie when the true state is k, then no agent will dishonestly report k.

Proof. If $s(k \mid j, t') > 0$, then

$$W_k - C(j, k, t') \geq W_j. \tag{5}$$

To prove the lemma it suffices to show that for all t,

$$W_k - C(i, k, t) > W_j - C(i, j, t). \tag{6}$$

By inequality (5), $W_k - W_j \geq C(j, k, t')$. However,

$$C(j, k, t') > c(j, k) \geq c(i, k) - c(i, j) = C(i, k, t) - C(i, j, t), \tag{7}$$

where the strict inequality follows from the definition of $C(\cdot)$, and the equation follows (for all t) by the definition of $C(\cdot)$.The weak inequality holds because $c(i, j) + c(j, k) \geq c(i, k)$. Inequality (6) follows immediately from inequalities (5) and (7). ■

Lemma 1 implies that there is no state j with the property that some types would lie if j is the true state and other states would dishonestly report j. The property that guarantees this result is that if $k > j > i$, then $C(j, k, t') > C(j, k, t) - C(i, j, t)$, which (as shown in the proof of the lemma) follows from the maintained assumptions on $C(\cdot)$. Two features of the cost function lead to the proposition. First, when $t > 0$, there is a fixed cost of lying. This suggests that if a type is willing to pay the fixed cost (to report k instead of j), then another type that has already decided to lie would not find it optimal to report j. The second feature of the cost function is that $c(i,j)+c(j,k) \geq c(i,k)$ so that marginal costs of increasing the size (in outcome space) of a lie is non-increasing.

Lemma 1 has two consequences. The first consequence (Proposition 1) is that if there exists a type (k,t) that lies, then the reputation for reporting k is 1. The second consequence is that no one ever makes a claim that is less than the truth. The first consequence follows directly from the lemma since the lemma states that if there exists a type (k,t) lies, then no one dishonestly claims k. Hence the report k must be honest. The second consequence (Proposition 2) follows because if $i < k$, then $v_k > v_i$. If type k reports i, then $\rho_k = 1$. So $W_k > W_i$ and type (k,t) would be better off reporting honestly than reporting i.

Proposition 1 *If $s(k \mid k,t) < 1$ for some t, then $\rho_k = 1$.*

Proof. Lemma 1 implies that if $s(k \mid k,t) < 1$, then the probability that another type reports k, r_k, is equal to zero. The result follows from the definition of ρ_k (given in equation (1)) and $h_k \neq 0$. ■

Proposition 2 *If $i < j$, then $s(i \mid j,t) = 0$ for all t.*

Proof. If $s(i \mid j,t) > 0$, then $W_j \geq W_i - C(j,i,t)$. Also, Proposition 1 implies that $\rho_j = 1$. Since $i < j$ implies that $v_i < v_j$, It follows that $W_j > W_i$ for all $i < j$. ■

The model provides a unique equilibrium prediction in the sense that all equilibria give rise to the same set of reputations. The next result implies that the set of claims that are (or are not) made dishonestly does not depend on the equilibrium selected.

Proposition 3 *If s' and s'' are two equilibria, then $\rho'_k = \rho''_k$ for all k.*

Here is an intuition for the result. Suppose that there are two equilibria that give rise to different reputations. Suppose that moving from the first equilibrium to the second, the reputation for reporting k goes down by the most. This means that it is less attractive to report k dishonestly in the second equilibrium. However, if there is a lower probability of dishonest reports of k in the second equilibrium, then the reputation associated with reports of k must be higher in the second equilibrium. (This observation requires that that the number of honest reports of k does not go down, which follows from Lemma 1.) Consequently all reputations must be higher in the second equilibrium. One can use the same argument to show that all reputations are higher in the first equilibrium. Consequently the proposition must hold.

Proof. Let M be the set of minimizers of $\rho'_j - \rho''_j$. If

$$W'_k - C(i,k,t) \geq W'_j - C(i,j,t),$$

then $W''_k - C(i,k,t) + \alpha\left((\rho''_j - \rho'_j) - (\rho''_k - \rho'_k)\right) \geq W''_j - C(i,j,t)$, hence if $k \in M$, then $W''_k - C(i,k,t) \geq W''_j - C(i,j,t)$ with strict inequality unless $j \in M$. It follows that if $s'(k \mid i,t) > 0$ for $i \neq k$, $k \in M$, then $s''(j \mid i,t) = 0$ for all $j \notin M, j \neq i$. Hence $\sum_{k \in M} r'_k \leq \sum_{k \in M} r''_k$ and for at least one $k \in M, r'_k \leq r''_k$. By Lemma 1 it follows that $\rho'_k \geq \rho''_k$ and hence $\rho'_i \geq \rho''_i$ for all i. Since we can use the same argument reversing the roles of the two equilibria, it follows that $\rho'_i = \rho''_i$ for all i. ■

Equilibrium may not be unique. Consider the special case in which $c(i,j) = 0$. In this case, all agents have identical preferences over lies. If equilibrium involves $\rho_k < 1$ for more than one value of k, then there will typically be different signaling strategies compatible with equilibrium. Nevertheless, Proposition 3 guarantees that the conditional probability of a lie given the observed value and the conditional probability of a report being honest is uniquely determined in equilibrium.

Let s be an equilibrium and let

$$L(s) = \{k : \text{ there exists } i \neq k \textit{ and } t, \text{ such that } s(k \mid i,t) > 0\}.$$

$L(\cdot)$ is the set of claims that are made dishonestly with positive probability in equilibrium. Note that $\rho_k(s) < 1$ if $k \in L(s)$ and, by Lemma 1, $\rho_k(s) = 1$ if $k \notin L(s)$, so $L(s) = \{k : \rho_k(s) < 1\}$.

Proposition 4 *The highest claim is made dishonestly with positive probability.*

Proof. Recall that N is the highest value. If $N \notin L(s)$, then $\rho_N(s) = 1$ and $W_N - W_{N-1} \geq v_N - v_{N-1}$. Since $c(N-1,N) < v_N - v_{N-1}$ by assumption, for t sufficiently small $W_N - C(N-1,N,t) > W_{N-1}$, so a $(N-1,t)$ agent would prefer to dishonestly report N than to tell the truth. Since Proposition 1 implies that a dishonest agent will never underreport, the result follows. ■

Proposition 4 demonstrates that agents will always tell maximal lies. When reputation matters, there will be partial lies in equilibrium. That is, N will not be the only claim reported dishonestly under intuitive conditions.

Proposition 5 *Suppose that $N > 2$ and $v_{N-1} - v_{N-2} > c(N-2, N-1)$. If either*

1. *$\alpha > v_N - v_{N-1}$ and p_N is sufficiently small or*
2. *α is sufficiently high,*

then N is not the only claim made dishonestly with positive probability.

Proof. We show that $L(s) = \{N\}$ is not possible if the conditions in the proposition hold. If $L(s) = \{N\}$, then $\rho_k(s) = 1$ for $k < N$ and $(N-2, t)$ must report either $N-2$ or N. Consequently,

$$\max\{v_{N-2}+\alpha, v_N+\alpha\rho_N-C(N-2, N, t)\} \geq v_{N-1}+\alpha-C(N-2, N-1, t). \quad (8)$$

Since $v_{N-1} - v_{N-2} > c(N-2, N-1)$, there exists $t^* > 0$ such that if $t < t^*$,

$$v_{N-1} + \alpha - C(N-2, N-1, t) > v_{N-2} + \alpha. \quad (9)$$

It follows from (8) that

$$v_N + \alpha\rho_N - C(N-2, N, t) \geq v_{N-1} + \alpha - C(N-2, N-1, t) \quad (10)$$

and that the probability that $N-2$ reports N is at least $F(t^*) > 0$ for all α and p_N. Consequently, there exists $b < 1$ such that $\rho_N < b$. We now have a contradiction: (10) cannot hold if α approaches infinity; if p_N approaches 0, then ρ_N must converge to 0, and (10) contradicts $\alpha > v_N - v_{N-1}$. ■

Propositions 4 and 5 require conditions on $v_k - v_{k-1}$. Without assumptions on the rate of increase of the rewards, some claims may not be worth lying for. For the next result we impose a stronger condition: $v_k - v_{k-1} > c(k-1, k)$ for all k. In many experimental designs, $v_k - v_{k-1}$ is a positive constant. In several theoretical models of lying costs (for example, those of Dufwenberg and Dufwenberg [5] and Khalmetski and Sliwka [12]), $c(k-1, k)$ is assumed to be zero. Clearly, if $v_k > v_{k-1}$ and $c(k-1, k) = 0$, then $v_k - v_{k-1} > c(k-1, k)$. This assumption adds intuitive structure to the equilibrium.

Proposition 6 *Suppose $v_k - v_{k-1} > c(k-1, k)$ for all k. There exists $n^* < N$ such that $L(s) = \{k : k > n^*\}$.*

$L(s) = \{k : k > n^*\}$ is equivalent to $\rho_k < 1$ for $k > n^*$ and $\rho_k = 1$ for $k \leq n^*$.

Proof. Let $n^* = \max\{k : \rho_k = 1\}$. It follows from Proposition 2 that $\rho_1 = 1$ so n^* is well defined. We must show that $\rho_k = 1$ for $k \leq n^*$. We know that $\rho_{n^*} = 1$. Assume that $\rho_k = 1$ for $k = k^*, \ldots n^*$ and $k^* > 1$. The condition $v_k - v_{k-1} > c(k-1, k)$ implies that for t sufficiently small $(k^* - 1, t)$ strictly prefers to report k^* to $k^* - 1$. Hence, by Proposition 1, $\rho_{k^*-1} = 1$, which establishes the result. ■

Proposition 6 states that there exists a cutoff observation. If the outcome is above this cutoff, then agents never lie. If the outcome is below the cutoff, then they lie with positive probability and dishonest claims are above the cutoff.

If agents care about their reputation for honesty, then the prior distribution over outcomes should influence behavior in a systematic way. In particular, if the prior distribution shifts mass from the most likely profitable outcome N to lower outcomes, telling the biggest lie should become less attractive. The next result formalizes this intuition. We consider a simple shift of probabilities: The distribution $p'' = (p''_1, \ldots, p''_N)$ is a **proportional shift from** N of $p' = (p'_1, \ldots, p'_N)$ if there is $\lambda \in (0, 1)$ such that $p_N = \lambda p'_N$ and $p'_i = (1 - \lambda p'_N)p_i/(1 - p'_N)$ for $i < N$.

The next result is our key comparative statics result. Compare a situation in which the states are ex ante equally likely to one in which the state giving the highest reward is extremely unlikely. In the second case, if all dishonest agents make the highest claim, it would lead to a low reputation. If reputation is sufficiently valuable, then dishonest agents would prefer to make a smaller claim, losing reward, but gaining reputation.

Proposition 7 *Suppose $v_k - v_{k-1} > c(k-1, k)$ for all k. Let p'' be a proportional shift from N of p'. Let s' (s'') be an equilibrium associated with a prior probability distribution p' (p''). For each i, $\rho'_i \geq \rho''_i$, $L(s') \subset L(s'')$, and for all $k < N$, the probability of a dishonest report of k is at least as great under p'' as under p'.*

Proposition 7 makes three conclusions. The first conclusion is that shifting prior probability to less valuable outcomes lowers the probability that any claim is viewed as honest. Hence a proportional shift from N lowers the utility of the agent. It is intuitive that the shift should lower ρ_N. If p_N decreases, then $\rho'_N < \rho''_N$ suggests that it is more attractive to report N under p'' than under p', but this would imply that $\rho'_N \geq \rho''_N$. The second conclusion is that more claims are made dishonestly under s'' than under s'. This means that shifting probability from N makes the subject willing to dishonestly report lower claims. Since there is a cutoff observation (by Proposition 6), this means that shifting probability from N lowers the lowest value that is reported dishonestly. Loosely, the reputation loss associated with higher claims could be great enough to convince the subject to make a more modest lie. The third claim is that the probability of partial lies (reports that are both dishonest and less than N) increases when p_N decreases. One reason for this change is non strategic. If p'' is a proportional shift of p', then there is more prior probability on low outcomes. Hence there are more situations under which it is attractive to lie. The conclusion depends on more than this observation. The first conclusion implies that there will be lower reputations after the proportional shift. This decreases the incentive to lie (and acts against the third conclusion). Further, some of the lies are maximal lies. This means that it could be possible that there are more lies under p'', but not more partial lies.

One might conjecture that a proportional shift from N would lead to a reduction in the fraction of maximal lies. We cannot establish this property. Because the proportional shift creates more possible lies (since a larger fraction of outcomes is less than N) it is possible that the ex ante fraction of maximal lies increases even as the payoff of these lies decreases. Hence there is an asymmetry between claims of N and claims of $k < N$. The reason for this is that if a smaller fraction of subjects dishonestly report $k < N$ under p'' than under p', it must be that $\rho''_k > \rho'_k$, since $p''_k > p'_k$, while it is possible for $\rho'_N \geq \rho''_N$ even if a smaller fraction of subjects dishonestly report N under p'' than under p' because $p'_N > p''_N$.

Proof. Let l'_{ik} (l''_{ik}) be the conditional probability that an agent who observes i dishonestly reports k under p' (p''). Hence $r'_k = \sum_i p'_i l'_{ik}$ and $r''_k = \sum_i p''_i l''_{ik} = (1 - \lambda p'_N) \sum_i p'_i l''_{ik} / (1 - p'_N)$ by the definition of p'' (note that since $l_{NN} = 0$, there is no term involving $p_N l_{NN}$).

If k is reported dishonestly, then no one who observes k lies, so $h_k = p_k$ by Lemma 1. It follows that if k is claimed dishonestly under s', $\rho'_k \geq \rho''_k$

if $r'_k/p'_k \leq r''_k/p''_k$ and if k is claimed dishonestly under both s' and s'' then $\rho'_k \geq \rho''_k$ if and only if $r'_k/p'_k \leq r''_k/p''_k$

Hence if $k \in L(s'), k < N$, then $\rho'_k \geq \rho''_k$ if $\sum_i p'_i l'_{ik} \leq \sum_i p'_i l''_{ik}$,

$$\text{if } k \in L(s') \cup L(s''), k < N, \text{ then } \rho'_k \geq \rho''_k \text{ if and only if } \sum_i p'_i l'_{ik} \leq \sum_i p'_i l''_{ik} \tag{11}$$

and

$$\rho'_N \geq \rho''_N \text{ if and only if } \sum_i p'_i l'_{iN} \leq \frac{(1-\lambda p'_N)\sum_i p'_i l''_{iN}}{\lambda(1-p'_N)}. \tag{12}$$

Since $\lambda(1-p'_N) \leq 1 - \lambda p'_N$, it follows from inequalities (11) and (12) that if $s'(k \mid i,t) > 0$ for some (i,t), then $\rho'_k \geq \rho''_k$ if

$$\sum_i p'_i l'_{ik} \leq \sum_i p'_i l''_{ik}. \tag{13}$$

Let M be the set of minimizers of $\rho'_j - \rho''_j$. If

$$W'_k - C(i,k,t) \geq W'_j - C(i,j,t),$$

then $W''_k - C(i,k,t) + \alpha\left((\rho''_j - \rho'_j) - (\rho''_k - \rho'_k)\right) \geq W''_j - C(i,j,t)$, hence if $k \in M$, then $W''_k - C(i,k,t) \geq W''_j - C(i,j,t)$ with strict inequality unless $j \in M$. It follows that if $s'(k \mid i,t) > 0$ for $i \neq k$, $k \in M$, then $s''(j \mid i,t) = 0$ for all $j \notin M, j \neq i$.

Hence $\sum_{k\in M} l'_{ik} \leq \sum_{k \in M} l''_{ik}$ and $\sum_i \sum_{k\in M} p_i l'_{ik} \leq \sum_i \sum_{k\in M} p_i l''_{ik}$ and for at least one $k \in M$,

$$\sum_i p_i l'_{ik} \leq \sum_i p'_i l''_{ik}. \tag{14}$$

If follows from inequalities (13) and (14) that $\rho'_k \geq \rho''_k$ for some $k \in M$. By the definition of M it must be that $\rho'_j \geq \rho''_j$ for all j. This establishes the first part of the proposition. The inclusion $L(s') \subset L(s'')$ follows from the definition of L and Lemma 1. Since $p'_k \leq p''_k$ for $k < N$, inequality (11) implies that there at least as many partial lies to k under p'' as under p' for all k. ■

Lemma 2 *Suppose $v_k - v_{k-1} > c(k-1,k)$ for all k. Reputation is weakly decreasing in report. That is, if $j' > j$, then $\rho_{j'} \leq \rho_j$.*

Proof. It follows from Proposition 6 that there exists n^* such that $\rho_j = 1$ if and only if $j \leq n^*$. We claim that if $j \geq n^*$, then $\rho_{j+1} < \rho_j$. This claim is sufficient to establish the proposition. The claim is true when $j = n^*$. When $j > n^*$, there must be a type (i, t) that dishonestly reports j. Hence j solves $\max_k W_k - C(i, k, t)$ and, in particular,

$$W_j - C(i, j, t) \geq W_{j+1} - C(i, j + 1, t), \tag{15}$$

which implies that

$$\alpha(\rho_j - \rho_{j+1}) \geq v_{j+1} - v_j + c(i, j) - c(i, j+1) > c(j, j+1) + c(i, j) - c(i, j+1) \geq 0,$$

where the first inequality follows from (15), the second inequality by $v_j - v_{j-1} > c(j - 1, j)$, and the third by the maintained assumption on $c(\cdot)$. It follows that $\rho_j > \rho_{j+1}$, which establishes the proposition. ■

Proposition 7 describes what happens if one shifts probability from the most attractive states to lower states. It demonstrates that such a shift increases the fraction of partial lies. One could ask whether this qualitative feature is a consequence of the highest state having low absolute probability (p_N small) or low relative probability (p_i/p_N large for $i < N$). Proposition 5 suggests that reductions in absolute probability of the most attractive outcome lead to increases in the fraction of partial lies. The next proposition confirms this. For the proposition, we compare two environments. In one, the outcome is uniformly distributed over $1, 2, \ldots, N$. In the second, the outcome is uniformly distributed over $.5, 1, \ldots, N - .5, N$. Hence one moves from the first environment to the second by "splitting" outcomes.[6]

Proposition 8 *Suppose $v_k - v_{k-.5} > c(k - .5, k)$ for all $k = 1, 1.5, \ldots, N$. Let p' be a uniform distribution on $\{1, 2, \ldots, N\}$ and Let p'' be a uniform distribution on $\{.5, 1, \ldots, N\}$. Let s' (s'') be an equilibrium associated with a prior probability distribution p' (p''). For each $i = 1, \ldots, N$, $\rho'_i \geq \rho''_i$ and $L(s') \subset L(s'')$. The probability of a dishonest report of $k < N$ is at least as great under p'' as under p'.*

The conclusion that $L(s') \subset L(s'')$ means that splitting states lowers the lowest value that is reported dishonestly. For example, if when $N = 10$ we

[6]The proposition uses a specific notion of splitting that is consistent with our experimental design.

observe dishonest reports of 8, 9, and 10, then when states are split we would expect to see dishonest reports of 7.5, 8, 8.5, 9, 9.5, and 10.

If the only dishonest claim made is $\{N\}$ (that is, $L(s'') = \{N\}$), then it is possible that more agents will make this claim dishonestly under p'' because the probability of an observation less than N is greater under p'' than under p'. This could happen if α is so low that nearly all agents make the maximum lie. In general, we cannot rule out the possibility that maximal lies are more common under p'' for the same reason we could not do so in Proposition 7: Under p'', outcomes less than N are more common and hence there is a higher ex ante probability of lying (and, perhaps, maximal lying).

Proposition 8 states that splitting states lowers the threshold below which subjects lie. A careful examination of the argument demonstrates that this can be thought of as a consequence of "rounding." That is, $\rho'_i \leq \rho''_{i-1}$ and if $L(s') = \{k : k \geq n^*\}$ then $L(s'') \subset \{k : k \geq n^* - 1\}$.

Proof. Let l'_{ik} (l''_{ik}) be the conditional probability that an agent who observes i dishonestly reports k under p' (p''). Hence $r'_k = \sum_i l'_{ik}/N$ and $r''_k = \sum_i l''_{ik}/2N$.

If k is reported dishonestly no one who observes k lies, so $h_k = p_k$ by Lemma 1. It follows that if k is claimed dishonestly under s', $\rho'_k \geq \rho''_k$ if $r'_k \leq 2r''_k$ and if k is claimed dishonestly under both s' and s'' then $\rho'_k \geq \rho''_k$ if and only if $r'_k \leq 2r''_k$. Hence if $k \in L(s')$, then $\rho'_k \geq \rho''_k$ if $\sum_i l'_{ik} \leq \sum_i l''_{ik}$,

$$\text{if } k \in L(s') \cup L(s''), \text{ then } \rho'_k \geq \rho''_k \text{ if and only if } \sum_i l'_{ik} \leq \sum_i l''_{ik} \tag{16}$$

It follows from inequality (16) that if $s'(k \mid i, t) > 0$ for some (i, t), then $\rho'_k \geq \rho''_k$ if

$$\sum_i l'_{ik} \leq \sum_i l''_{ik}. \tag{17}$$

Let M be the set of minimizers of $\rho'_j - \rho''_j$. If

$$W'_k - C(i, k, t) \geq W'_j - C(i, j, t),$$

then $W''_k - C(i, k, t) + \alpha\left((\rho''_j - \rho'_j) - (\rho''_k - \rho'_k)\right) \geq W''_j - C(i, j, t)$, hence if $k \in M$, then $W''_k - C(i, k, t) \geq W''_j - C(i, j, t)$ with strict inequality unless $j \in M$. It follows that if $s'(k \mid i, t) > 0$ for $i \neq k$, $k \in M$, then $s''(j \mid i, t) = 0$

for all $j \notin M, j \neq i$. Hence $\sum_{k \in M} l'_{ik} \leq \sum_{k \in M} l''_{ik}$ and therefore

$$\sum_{i=1}^{N} \sum_{k \in M} l'_{ik} \leq \sum_{i=1}^{N} \sum_{k \in M} p_i l''_{ik} \tag{18}$$

and for at least one $k \in M$,

$$\sum_{i=1}^{N} l'_{ik} \leq \sum_{i=1}^{N} l''_{ik}. \tag{19}$$

If follows from inequalities (17) and (19) that $\rho'_k \geq \rho''_k$ for some $k \in M$. By the definition of M it must be that $\rho'_j \geq \rho''_j$ for all $j = 1, 2, \ldots, N$. Consequently, inequality (16) implies that inequality (19) holds for all k. This establishes the first part of the proposition. The inclusion $L(s') \subset L(s'')$ follows from the definition of L and Lemma 1. It remains to show that for $i < N$, $\sum_i l'_{ik} \geq 2 \sum l''_{ik}$. It follows from Lemma 2 that if $j < k$, then $\sum l''_{ij} \geq \sum l''_{ik}$. In particular, this inequality holds when $j = k - .5$. Using inequality (19), we conclude that

$$\sum_{i=1}^{N} l'_{ik} \leq \sum_{i=1}^{N} l''_{ik} \leq 2 \sum_{i} \sum_{k} l''_{ik},$$

which is the desired result. ■

Let us further specialize the model and assume that $v_i - v_{i-1} \equiv \nu$, where ν is a positive constant and that $c(i,j) = d(j-i)$ depends on the difference between the reported state and the true state.

Proposition 9 *Suppose $v_k - v_{k-1} > c(k-1, k)$ for all k and $v_i - v_{i-1} \equiv \nu > 0$ for all i. The probability of an honest report is a non-increasing function of the observed value.*

Proof. It suffices to show that if $j > i$ and type (j,t) dishonestly reports k, then type (i,t) prefers to report $k - j + i$ rather than to tell the truth. That is, if

$$W_k - d(k-j) - t \geq v_j \text{ implies that } W_{k-j+i} - d(k-j+i-i) - t \geq v_i. \tag{20}$$

Implication (20) follows because

$$W_k - v_j = v_k - v_j + \alpha r_k = v_{k-j+i} - v_i + \alpha r_k \geq v_{k-j+i} - v_i + \alpha r_{k-j+i} = W_{k-j+i} - v_i,$$

where the first and last equations are definitions, the second inequality follows because $v_k - v_j = (k-j)\nu = v_{k-j+i} - v_i$ and the inequality follows from Lemma 2. ■

Abeler, Nosenzo, and Raymond [2] introduce a variety of models for an environment in which there are two possible outcomes and two possible reports. The models include preferences in which making a dishonest report lowers utility and in which agents' utility is increasing in their reputation for honesty, a force that operates similarly to our likelihood dimension. They find that models that include both of these features helps organize experimental data.

Dufwenberg and Dufwenberg [5] study a model in which the agent's preference depends on the claim and a term similar to our reputation term. The reputation term is a decreasing function of the difference between the claim and an outsider's expectations about the observed value. In our model, the likelihood term is proportional to the probability that the report is honest. In Dufwenberg and Dufwenberg, this term depends on beliefs about the true observation. Hence, the reputation cost of making a dishonest report may depend on the level of dishonesty. Dufwenberg and Dufwenberg study a model in which states are equally likely. Similar to us, they show that partial lies are possible in equilibrium and they arise when the number of states increases (and hence the prior probability of the best state decreases). In contrast to us, their model typically exhibits multiple, qualitatively different equilibria.

Khalmetski and Sliwka [12] analyze a special case of our model in which $c(i,j) = 0$ (they also assume that the prior distribution is uniform). This specialization permits them to provide a more complete description of the unique symmetric equilibrium of the model. While they focus on comparative statics with respect to the value of reputation (α), they identify some of the same important qualitative properties that we do.

In Dufwenberg and Dufwenberg [5] and Khalmetski and Sliwka [12] there will be broad indifference across possible lies and randomization is essential in the construction of equilibrium. If k and k' are claimed dishonestly in equilibrium, then anyone who claims k dishonestly in equilibrium will obtain the same utility by claiming k'. In our most general specification, if $k \neq k'$ type (k,t) may strictly prefer to make a different dishonest report than type (k',t). Because the fixed cost of lying enters additively in $C(\cdot)$, type (k,t) and type (k,t') will have the same preferences over lies for every k (if $t' > t$, then (k,t) might prefer to report honestly when (k,t') prefers to lie). That is, if two agents observe the same k and make different dishonest reports, then the model says that they must be indifferent between these reports.

4 Hypotheses

This section describes testable implications of the theory. The results section below will be based on these hypotheses.

Hypothesis 1 *If some type lies when reporting k, then no type with true value k lies.*

Hypothesis 1 is a consequence of Lemma 1.

Hypothesis 2 *No agent underreports.*

Hypothesis 2 is a consequence of Proposition 2.

Hypothesis 3 *The highest claim is made dishonestly with positive probability.*

Hypothesis 3 is a consequence of Proposition 4.

Hypothesis 4 *Due to the reputational concerns, some agents will lie partially.*

Hypothesis 4 is a consequence of Proposition 2. .

Hypothesis 5 *There exists a threshold of true values, below which there are lies with positive probability, above which there are no lies.*

Hypothesis 5 is a consequence of Proposition 6.

It follows from Hypotheses 1–3 and 5 that we should observe the number of reports of a k to be below the actual number of times the true value is k up to a cutoff for $k \leq k^*$ and the number of reports of k to be above the true number for $k > k^*$.

Hypothesis 6 *An increase in the probability that the true type is less than the highest type increases the number of values reported dishonestly.*

Hypothesis 7 *An increase in the probability that the true type is less than the highest type increases the probability of partial lies.*

Hypotheses 6 and 7 are consequences of Proposition 7.

The next two hypotheses refer to the exercise of "splitting the states" discussed in Proposition 8. That is, we compare equilibria of a situation in which there are N equally likely, equally spaced, outcomes to one in which there are $2N$ equally likely, equally spaced, outcomes (with new outcomes inserted between old ones).

Hypothesis 8 *Splitting the states increases the range of the values that are reported dishonestly.*

Hypothesis 8 states that splitting states lowers the cutoff that determines the lowest claim that would be made dishonestly.

Hypothesis 9 *Splitting the states increases the probability of partial lies.*

Hypotheses 8 and 9 are consequences of Proposition 8.

Hypothesis 10 *The lower the true value, the higher the fraction of dishonest reports.*

Hypothesis 10 is a consequence of Proposition 9.

5 Experimental Design and Procedure

To test the theoretical predictions of our model regarding the dimensions of lying costs, we introduce two types of games, which we call **observed** and **non-observed** games.

5.1 Observed Games

The observed game is a variation of a cheating game, in which we can observe the individual lying behavior. In this game, we ask participants to click, in private, on one of ten boxes on a computer and reveal an outcome. We use three different observed game variations. In the "Number" treatment, the outcomes behind the ten boxes are numbers between one and ten, where each box has a different number. After seeing the number, the participant is asked to report it to an experimenter, knowing that payments will be equal to the number s/he reports in Euros. In this treatment we know how often and to what extent participants lie because we can later observe the actual number each participant saw and compare it with the number s/he reported.

One possibility is that lying costs depend on the distance between what the subject observes and what the subject says. We call this the outcome dimension of lying costs. In the numbers treatment, there is a natural, common, measure of distance in the space of outcomes and reports. For example, if the participant observes a "four" and reports "ten," her report is six units from the truth. One can imagine that lying costs depend on this distance.

Another possibility is that lying costs depend on the amount of payoff gained by the report relative to what the participant would earn if she reports honestly. We call this the payoff dimension. In the numbers treatment payoffs are linked directly to outcomes so one cannot distinguish the outcome dimension from the payoff dimension. The second treatment, "Numbers Mixed," is designed to separate the reported outcome dimension from the payoff dimension. This treatment is similar to the Numbers treatment, but the ten numbers are assigned to the ten payoffs in a random order. Table 1 presents the assignment that we predefined in a random draw.

[Insert Table 1 here]

While in the Numbers Mixed treatment the outcome and payoff dimension are separated, there is a natural way to measure the distance between different observed outcomes. In the third observed treatment, "Words," there is no natural ordering of the outcomes independent of payoffs. In this treatment, participants are asked to click on one of ten boxes in private and are told that the outcomes behind the boxes are ten Lithuanian words; each box has a different word. The words have payoffs between one and ten Euros assigned to them, as presented in Table 2.

[Insert Table 2 here]

There is no natural notion of distance in the outcome dimension because none of the participants knows Lithuanian and all the words are six letters words similar to each other.[7] More specifically, participants may distinguish between reporting truthfully or not, but we assume that the "outcome cost" of reporting "stirna" when the outcome is "vilkas" is the same as the outcome cost of reporting "kiskis" when the outcome is "vilkas." More generally, we assume that the outcome cost is zero for honest reports and the same for dishonest reports.

5.2 Non-Observed Games

As discussed in Sections 2 and 4, some participants might have reputational concerns when reporting the number. To signal she is not a liar, the participant may refrain from reporting the highest number. Since only the participant sees actual outcomes, in the non-observed game, third-parties must infer the probability that a report is honest. It is possible that lower reports are viewed as honest with sufficiently high probability that participants will sacrifice material payoffs in order to be perceived more favorably. In contrast, in the observed game, any dishonest report leads to a reputation of zero.

In the "Basic" non-observed treatment, we give the participant a sealed envelope with ten folded pieces of paper that have numbers from one to ten on them. We ask the participant to take out one piece of paper, observe the number she took out, put it back into the envelope, and then report it. As in the observed treatment, payments are equal to the number reported in Euros. However, differently from the observed game, the experimenter can never know the actual outcome. If reputational concerns affect the lying decision, we would expect more participants lying and/or a higher fraction of participants who partially lie in the non-observed than in the observed game.

To test an important prediction of the model, we are also interested to see how decreasing the prior probability of the highest outcome affects the number of values reported dishonestly. In the "Low Probability" non-observed treatment we use a similar procedure as in the Basic non-observed treatment and adjust the prior probabilities of the outcomes occurring. We give the participant a sealed envelope with 100 folded pieces of paper that have num-

[7]The meaning of the words is: wolf, forest, heaven, deer, paradise, north, rabbit, lilac, apple, south.

bers from one to ten on them. We inform the participant there are eleven pieces of paper with the number "one" on it, eleven pieces with the number "two," and so on until "nine." However, there is only one piece of paper with the number "ten" on it. As in the Basic non-observed treatment, the payments are equal to the number she reported in Euros. In line with the model predictions, we expect to observe a higher range of values reported, since drawing a "ten" has a chance of only 1%, as opposed to 10% in the Basic non-observed treatment. Note that if the prior distribution of outcomes influences reports, then outcome/payoff lying costs are insufficient to understand data.

In the final treatment –"100-States" non-observed treatment – we investigate the robustness of the predictions in the low probability treatment. In this treatment, we give participants a sealed envelope with 100 pieces of paper with numbers between one and 100 on them, informing them that there is one piece of paper with each of the numbers. Participants are paid the equivalent in Euros to the number they report divided by ten. While the probability of drawing 100 is the same as in Low Probability treatment, all the other outcomes are equally likely. The question is whether the partial lying that we observe in the low probability treatment is the result of the difference in relative probability of the highest state and lower states or is due to the fact that the absolute probability of the highest state is low.

5.3 Experimental Procedure

We conducted the experiments between April 2015 and April 2016 at the Cologne Laboratory for Economic Research, University of Cologne. We used the experimental software zTree (Fischbacher [7]) and recruited participants via ORSEE (Greiner [10]). Overall, we recruited 916 participants (55.9% female), and none of them participated in more than one session. We collected 102–390 observations per treatment. Participants played only our treatment in the experimental session with a session lasting approximately 30 minutes.

After being randomized into a treatment, participants read the instructions on the computer screen, and were allowed to ask questions privately. Then, depending on the treatment, participants either received the envelopes with numbers or were asked to click on one of the boxes on the computer screen and reveal the number. After observing the number, participants reported the outcomes on a sheet of paper and filled out a post-experiment questionnaire that included questions on gender, age, field of study, and motives behind the decisions. At the end, participants privately received their payoffs in cash and left the laboratory. Table 3 presents all of our treatments and the number of participants in each.

[Insert Table 3 here]

6 Results

In what follows, we first present the data from the observed treatments and then move to testing the hypotheses that are based only on the observed treatments (Hypotheses 1–3, 5, and 10). We then report the results from the non-observed treatments and the tests of the corresponding hypotheses (Hypotheses 4 and 6–9).

6.1 Observed Game

Figure 1 presents the distributions of actual and reported payoffs and is the first indicator that the reported payoffs are higher than the actual payoffs resulting from the outcomes (numbers or words).

[Insert Figure 1 here]

We are among the first in the literature to observe the actual outcomes in cheating games,[8] we can measure lying on the individual level by comparing the distribution of actual payoffs with the distribution of reported payoffs. Using a Wilcoxon Matched-Pairs Signed-Ranks test we find that in all the observed treatments the reported numbers are significantly higher than the actual outcomes with $p < 0.001$.[9] Overall, 26%, 33% and 27% participants lie in the Numbers, Words and Numbers Mixed treatments, respectively. The overall level of lying is not significantly different between the treatments in pairwise comparisons in a Fisher exact test ($p > 0.1$).

The observed game also allows us to analyze the probability of lying conditional on the actual payoff observed, as presented in Figure 2. We can see from the figure that in the Numbers treatment, the lower the actual

[8] Kröll and Rustagi [13] use blue-tooth equipped dice to "observe" true outcomes in a Fischbacher and Föllmi-Heusi [8] game.

[9] All tests in the paper are two sided. We call an effect highly significant, significant or marginally significant if the test generates $p < 0.01$, $p < 0.05$, $p < 0.1$, respectively.

outcome is, the more likely participants are to lie. For example, only 5% of participants who observed a nine overreport their number, whereas 47% of the participants who observed a one did. The results show a significant negative correlation (Spearman's rho = -0.318, $p < 0.001$) between the payoff observed and the probability of lying. In the Words treatment we also observe a negative correlation, but less strong than in the Numbers treatment (Spearman's rho = -0.202, $p = 0.042$). In the Numbers Mixed treatment there is a marginally significant correlation between the actual payoff and the probability to lie (Spearman's rho = -0.170, $p = 0.079$).[10]

[Insert Figure 2 here]

Another important feature of the observed treatment is that it allows us to know what payoff people who lie report. Figure 3 presents this data. We find no correlation between the actual and reported payoff when one lies (Spearman's rho = 0.052, $p = 0.601$ for the Numbers treatment; rho = 0.188, $p = 0.288$ for the Words treatment and rho= -0.021, $p = 0.914$ for the Numbers Mixed treatment).

[Insert Figure 3 here]

While the actual observed payoffs are not different between the treatments (i.e., the randomization worked; $p > 0.1$, MWU), the average reported payoff in the Numbers treatment is marginally lower than in the Words treatment (7.02 versus 7.39, respectively; $p = 0.090$, MWU), and is not statistically different from the average reported payoff of 7.03 in the Numbers Mixed treatment ($p = 0.198$, MWU). The difference between the Numbers Mixed and Words treatments is also not statistically significant ($p = 0.205$, MWU). Thus, in the extensive margin the reporting is not significantly different between the observed treatments.[11]

Next we consider the payoffs reported by participants who lie, which is presented in Figures 4 and A1. We observe that in the Numbers treatment, 68% of participants who lie, lie to the full extent by saying ten. This fraction is 91% in the Words treatment and 80% in the Numbers Mixed treatment.

[10]One person under-reported in this treatment (observing ten and reporting four) and two did not click on any boxes and then reported a ten. The two participants who did not click are excluded in Figure 2 and in the Spearman's correlations, since they have no actual payoff. If the two participants are not excluded, Spearman's rho amounts to -0.213 with $p = 0.025$.

[11]The extensive margin corresponds to the fraction of people who lie, while the intensive margin corresponds to the size of the lie for people who choose to do so.

The fraction of participants who lie by reporting ten in the Words treatment is significantly higher than in the Numbers treatment ($p = 0.007$, Fisher exact test). The average payoff reported by participants who lie in the Words treatment, 9.80, is also significantly higher than the Numbers treatment (9.80 versus 9.32, respectively; $p = 0.011$, MWU). Thus, in the intensive margin, we find significant differences between lying behavior in the Words and Numbers treatments.

[Insert Figure 4 here]

The difference between lying in the Words and the Numbers treatments suggests that the outcome dimension affects lying cost, since we observe less partial lying in the Words treatment without a clear notion of partial lying than in the Numbers treatment in which there is an intuitive notion of intermediate lies. That is, some participants perceive reporting an "eight" when observing a "four" a smaller lie than reporting a "ten" in the Number treatment, while in the Words treatment reporting "alyvos" when observing "vilkas" has the same outcome cost as reporting "alyvos" when observing "stirna." However, the role of the outcome dimension on the extent of lying is relatively small, since only 8.45% of the participants (33 out of 390) lie partially in the Numbers treatment and the fraction decreases to 2.94% (3 out of 102) in the Words treatment. In addition, as we show above, there is no significant effect on the extensive margin. The absence of partial lying in the Words treatment suggests that the payoff dimension has no effect on the cost of lying on the intensive margin. It appears that when observing an outcome that results in 4 Euros if reported honestly, the cost of dishonestly reporting something that leads to a payoff of 6 Euros is not significantly lower than the cost of dishonestly reporting something that leads to a payoff of 8 Euros.

In the Numbers Mixed treatment, the fraction of participants who lie is between the Numbers and Words treatments and is not significantly different from the two treatments ($p = 0.257$ and $p = 0.285$, Fisher exact test). The average payoff reported by participants who lie in the Number Mixed treatment is not significantly different from neither Words nor the Numbers treatments (9.80 versus 9.32, respectively; $p = 0.205$ and $p = 0.250$, MWU).

Figures 5a–c show the results for the Numbers Mixed treatment with respect to the outcome dimension instead of the payoff dimension. The results clearly show that the lying behavior in this treatment is not related to the outcome dimension – the decision to lie does not depend on the actual number observed (see Figure 5b; Spearman's rho = -0.157, $p = 0.102$) and when participants lie, they lie mostly by reporting a "two," which results in a payoff of ten (Figures 5a and 5c). That is, when there is a trade-off between the outcome and payoff dimensions, the participants lie according to the payoffs and neglect the outcome dimension. This finding again suggests that there is only a limited role of the outcome dimension on the lying cost.[12]

[Insert Figure 5 here]

Based on the comparisons between the observed treatments, we have our first results on the cost of lying:

Result 1 *The outcome dimension has a limited effect on the intensive margin and no effect on the extensive margin.*

The results from the observed treatments allow us to test Hypothesis 1, which states that if a participant observing k ever lies, then no one ever lies by saying k. As Figure 6 and Table A2 show, the results are generally in line with this prediction. We use the pooled data from the observed game to count for deviations from the prediction with respect to the theory; we define a behavior a "mistake" if it violates Hypothesis 1 (i.e., if given that someone observing k lies, someone lies by reporting a k). Under this definition, mistakes are the minimum of the fraction who lie when observing the number and the fraction who lie by reporting the number. In Figure 6, the mistakes are marked with a dashed line. Lines 1–3 in Table A2 (left side of the Figure 6) contain no mistakes. Line 4 in Table A2 contains one mistake: participants lie after observing a four and there is one participant out of 167 (0.60%) who lies by reporting a four. Lines 5, 6, 7, 8 and 9 contain 2, 5, 6, 13 and 15 mistakes, respectively (1.20%, 2.99%, 3.59%, 7.78% and 8.98% out of 167 participants). Line 10 contains one mistake, since one person out of 73 (1.37 %) observing a ten lies downwards.

[Insert Figure 6 here]

Overall, we observe 43 mistakes for 602 participants (7.14%) in our data. We conclude:

[12] We do not have a test of what happens if we change the payoffs associated with the decisions. For this reason, we cannot estimate the effect of the payoff costs on the extensive margin. Note that it is not trivial to provide such a test, since changing the payoffs would lead to changes in the behavior independent of the intrinsic lying cost (see Kajackaite and Gneezy [11]). We leave this exercise for future research.

Result 2 *The data show that if a participant observing k lies, then only a small fraction of participants lie by saying k.*

Hypothesis 2 states that a participant would not underreport her payoff. This is easy to test in our data. As we reported above, only 1 participant out of 602 underreported in our experiment. Therefore we conclude:

Result 3 *Most participants (99.83%) do not underreport their payoffs.*

Hypothesis 3, which asserts that the highest claim is made dishonestly with a positive probability, is also clearly supported by the data presented in Figure 3 above. In particular, we observe that of the people who lie, 68%, 80% and 91% of participants report the highest possible payoff in the observed treatments:

Result 4 *From the participants who lie, a high fraction (an average of 74.85%) report the highest payoff.*

We find partial support for Hypothesis 5, that there exists a threshold of actual payoffs, below which there are lies with positive probability and above which there are no lies. In particular, in the observed treatments 27.63-46.94% (an average of 34.26%) of participants lie when observing a payoff below nine, but only 9.68% and 1.37% lie after observing nine or ten, respectively (see Figure 2, Figure 6 and Table A2 in the Appendix):

Result 5 *There is a threshold of actual payoffs of nine, below which there is a high fraction of lies (average 34.26%), and above which there are only few lies (average of 5.19%).*

We report support for Hypothesis 10, which states that the lower the true value the higher the fraction of dishonest reports, in Figure 2. As we have observed, there is a significant negative correlation between the payoff observed and the probability of lying in the observed game:

Result 6 *There is a significant negative correlation between the payoff observed and the probability of lying in the observed game: The lower the actual payoff, the higher the fraction of dishonest reports.*

6.2 Non-observed Game

We now move to presenting the results of the non-observed treatments that will be used in testing the next set of hypotheses. Figures 7a, b, c, d present the results from the non-observed treatments. The figures show the distributions of the reported payoffs on the aggregate level and, as in the case of the observed treatments, indicate (this time only statistically) that reported payoffs are higher than expected observed payoffs; a Kolmogorov-Smirnov test confirms participants lie significantly in the non-observed treatments ($p < 0.001$).

[Insert Figure 7 here]

Whereas 14% report a nine in the Numbers observed treatment (which is not significantly higher than the actual fraction of 11% who observed a nine; $p = 0.106$, binomial test), 22% report a nine in the Basic non-observed treatment, which is significantly more than the theoretical prediction ($p < 0.001$). The difference between the Basic and Numbers treatments in reporting a nine is significant ($p = 0.033$, Fisher test). That is, in the non-observed treatment some of the participants who lie do not report the maximal payoff. This result is predicted by Proposition 2 and Hypothesis 4, assuming that some participants care about their reputation. In the non-observed treatment, to signal to the experimenter that she does not lie, a participant who lies may choose to claim high but not maximal numbers, such as eight or nine. Also supporting the reputational concern prediction, we find that the overall level of lying is higher in the Basic non-observed treatment than in the Numbers observed treatment (average reported number/payoff of 7.81 vs. 7.02, respectively, $p = 0.016$, MWU), indicating that some participants do not lie in the observed treatment because the experimenter may know that they lied.[13] Hence, we conclude:

[13] In the observed game, a lie must lead to $\rho = 0$. So there is less incentive to tell a partial lie in the observed game (in the non-observed game, someone might tell a partial lie in order to improve her reputation relative to a full lie). Even without variation in ρ, partial lies are possible in the observed game (due to variations in the cost function), but observability makes all lies less important. Reputation does play a role in the observed game because the reputation term adds α to honest reports (and 0 otherwise).

Result 7 *In line with reputational concerns, a larger fraction of participants lies partially in the non-observed treatment than in the observed treatment.*

Previous evidence showing partial lying in variations of the non-observed treatment were interpreted as a desire to maintain a positive self-image. Most notably, Mazar, On, and Ariely [15], conclude that "A little bit of dishonesty gives a taste of profit without spoiling a positive self-view." While replicating the partial lying finding of Mazar et al. and others, our results do not support their interpretation. Instead, our results suggest that partial lies are primarily due to reputational concerns, since the partial lying that might be caused by self-image concerns is low in the observed games and partial lying substantially increases in the non-observed game.

Hypothesis 6 predicts that a lower prior probability of the highest outcome will increase the number of values reported dishonestly. The Low Probability and 100–states non-observed treatments were designed to test this prediction. In the Low Probability treatment we reduced the probability of a "ten" to 1%. Figure 7b presents the results from the Low Probability treatment.

In line with the model's predictions, lowering the prior of the highest outcome increased the range of values reported dishonestly. While in the Basic treatment, only the fractions of reports of nine and ten are higher than the prior (22% and 37% compared to 10% prior; $p < 0.001$, binomial test), in the Low Probability treatment, participants overreport eight, nine and ten (marginally) statistically significantly and seven is overreported but not significantly so. Here, 16% report a seven (compared to 11% prior, $p = 0.120$), 17% an eight (compared to 11% prior, $p = 0.063$), 34% a nine (compared to 11% prior, $p < 0.001$) and 6% a ten (compared to 1% prior, $p < 0.001$). That is:

Result 8 *As predicted by Hypothesis 6, when a payoff of 10 has a 1% chance, a larger number of values is reported dishonestly.*

Hypothesis 7 predicts that a lower prior probability of the highest outcome will increase the probability of partial lies. In the Basic non-observed treatment, only nine and ten are overreported relative to the expected fraction. Nine was reported by 23 (22.33%) out of 103 participants; compared with the expected 10% who observe a nine. Thus we estimate the partial lying to be 12.33% in this treatment. In the Low Probability non-observed treatment, seven, eight, nine and ten are overreported. 11% should observe each seven, eight and nine, but 17 (15.89%), 18 (16.82%) and 36 (33.64%) out of 107 participants claim to do so. Therefore, in the Low-Probability treatment, the estimated partial lying is 33.36% $(71 - 35.31)/107$, with 71 being the sum of participants claiming 7–9, and 35.31 expected fraction of participants who observe 7–9.

Result 9 *As predicted by Hypothesis 7, when a payoff of ten has a 1% chance, the fraction of partial lies increases relative to when the payoff of 10 has 10% chance.*

The Low Probability treatment demonstrates that making the highest possible outcome relatively less likely than the other outcomes makes participants report a larger range of outcomes dishonestly. In the 100-States treatment we test whether lowering the absolute probability of the highest state has a similar effect. Figure 7c presents the results and Figure 7d presents the aggregate results.

To test Hypothesis 8 that splitting the states increases the number of values reported dishonestly, we compare the data from 100-states with the Basic treatment. As described above, in the Basic treatment only nine and ten are overreported (two out of ten possible outcomes). In the 100-state treatment, we find that 22 out of 100 outcomes are overreported relative to the expected 1%. The cutoff at which numbers are reported dishonestly is lower in the 100-state condition than in the 10-state condition. Significant overreporting starts at 62 (out of 100) in the 100-state condition the 100-state condition. It starts at 9 (out of 10) in the ten-state condition. We summarize this finding in Result 10.

Result 10 *As predicted by Hypothesis 8, the range of values reported dishonestly in the 100-state treatment is larger than in the ten-states treatment.*

Finally, to test Hypothesis 9 that splitting the states will increase the probability of partial lies, we compare estimated partial lying in the Basic treatment (12.33%) with the 100-states treatment. We estimate partial lying in 100-state treatment by analyzing pooled intervals and identifying which intervals are overreported. We split the data into 11 groups: 1–9, 10–19, . . . , 90–99, and 100. We find that 60–69, 80–89 and 90–99 are overreported with 60.58% reporting those outcomes. We conclude that partial lying amounts to 30.58% $(60.58 - 30)$ in the 100-state treatment.

Result 11 *As predicted by Hypothesis 9, splitting the states increases the fraction of partial lies.*

To obtain Result 11 we isolate the highest state and pooled lower intervals in groups of ten. Isolating the highest state is consistent with our objective of identifying whether reducing the probability of the highest state leads to a lower fraction of agents making the highest claim. If we pool together states 1–10, 11–20, ..., 91–100, then there is statistically significant overreporting only in the two highest pools, which is consistent Proposition 8 and, in particular, the ideas that splitting states does not qualitatively change lying behavior.

7 Conclusion

Understanding when and why people (do not) cheat is important for many economic activities. Standard economic models under classic assumptions of selfishness and profit maximizing partially answer the question by appealing to the extrinsic cost of lying. But recent literature presents strong evidence that intrinsic costs of lying are also important. In this paper we formalize an important aspect of this intrinsic costs of lying – how does the size of the lie affect the decision. We discuss three possible kinds of lying cost: a cost related to the distance between the true outcome and what is reported; a cost related to the monetary gains generated by the lie; and a reputational cost associated with the probability that a statement is perceived to be dishonest. While the first two dimensions of the cost of lying were discussed informally in the literature on sizes of lies, the third one was not.

The model we construct allows us consider the influence of the size of the lie on outcomes and generate novel predictions that we test experimentally. In line with the properties of the equilibrium of our model, we find evidence for a cutoff value: If the payoff associated with the outcome is high enough people do not lie, and lies occur only when the payoff is below this cutoff. In equilibrium the reputation for being honest is decreasing in the claim. In the experiment, as predicted by the model, there are always dishonest claims of the maximal value. When reputational concerns are made less important (e.g., by making the outcome non-observed), there is more partial lying. Another support for the reputation argument is that when making the maximal outcome less likely ex ante, the frequency of partial lies increases.

We conclude that the size of the lie in terms of reputation seems to have an important impact on the decision whether to lie. Our findings indicate that the other two dimensions – the outcome and the monetary gain – have smaller effects on behavior. For example, as Result 1 states, the outcome dimension has no effect on the number of people who choose to lie and a small effect on partial lying.

Put together, our paper is offering a formal treatment and testing of an important aspect of lying behavior.

References

[1] Johannes Abeler, Anke Becker, and Armin Falk. Representative evidence on lying costs. *Journal of Public Economics*, 113:96–104, 2014.

[2] Johannes Abeler, Daniele Nosenzo, and Collin Raymond. Preferences for truth-telling. Technical report, Oxford University, July 19 2016.

[3] Alain Cohn, Ernst Fehr, and Michel André Maréchal. Business culture and dishonesty in the banking industry. *Nature*, 516(7529):86–89, 2014.

[4] Anna Dreber and Magnus Johannesson. Gender differences in deception. *Economics Letters*, 99:197–199, 2008.

[5] Martin Dufwenberg, Jr. and Martin Dufwenberg, Sr. Lies in disguise – a theoretical analysis of cheating. Technical report, University of Arizona, September 2016.

[6] Sanjiv Erat and Uri Gneezy. White lies. *Management Science*, 58(4):723–733, 2012.

[7] Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, 10(2):171–178, 2007.

[8] Urs Fischbacher and Franziska Föllmi-Heusi. Lies in disguise: An experimental study on cheating. *Journal of the European Economic Association*, 11:525–547, 2008.

[9] Uri Gneezy. Deception: The role of consequences. *American Economic Review*, 95(1):384–394, March 2005.

[10] Ben Greiner. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125, 2015.

[11] Agne Kajackaite and Uri Gneezy. Lying costs and incentives. Technical report, UCSD, 2016.

[12] Kiryl Khalmetski and Dirk Sliwka. Disguising lies – image concerns and partial lying in cheating games. Technical report, University of Cologne, April 2016.

[13] Markus Kröl and Devesh Rustagi. Got milk? motivation for honesty and cheating in informal markets: Evidence from india. Technical report, Research Center SAFE-Sustainable Architecture for Finance in Europe, Goethe University Frankfurt, 2016.

[14] Tobias Lundquist, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1):81–92, 2009.

[15] Nina Mazar, On Amir, and Dan Ariely. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research*, 45(6):633–644, 2008.

[16] Shaul Shalvi, Jason Dana, Michel JJ Handgraaf, and Carsten KW De Dreu. Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, 115(2):181–190, 2011.

[17] Matthias Sutter. Deception through telling the truth? Experimental evidence from individuals and teams. *The Economic Journal*, 119(534):47–60, 2009.

Table 1: The relation between number reported and payoff in the Numbers Mixed treatment

Number	7	3	1	8	4	9	5	10	6	2
Pay in €	1	2	3	4	5	6	7	8	9	10

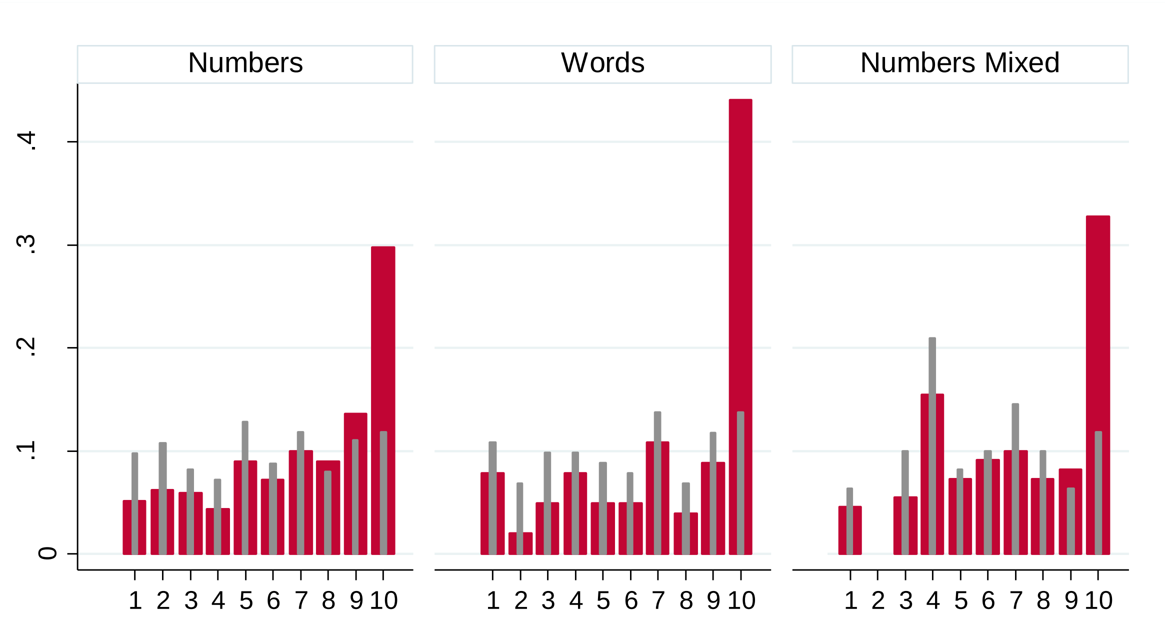
Table 2: The relation between word reported and payoff in the Words treatment

Word	vilkas	miskas	dangus	stirna	rojuje	siaure	kiskis	alyvos	obelis	pietus
Pay in €	1	2	3	4	5	6	7	8	9	10

Table 3: Summary of treatments and number of participants in each

Treatment		Number of participants
Observed	Numbers	390 (54.9% female)
	Numbers Mixed	110 (62.7% female)
	Words	102 (60.8% female)
Non-Observed	Basic	103 (52.4% female)
	Low Probability	107 (52.3% female)
	100-States	104 (54.8% female)

Figure 1: Distribution of reported payoffs in the observed treatments



Note: The thick dark red bars show the reported payoffs, whereas the thin light gray bars show the actual payoffs.

Figure 2: Fraction of lying conditional on the actual payoff

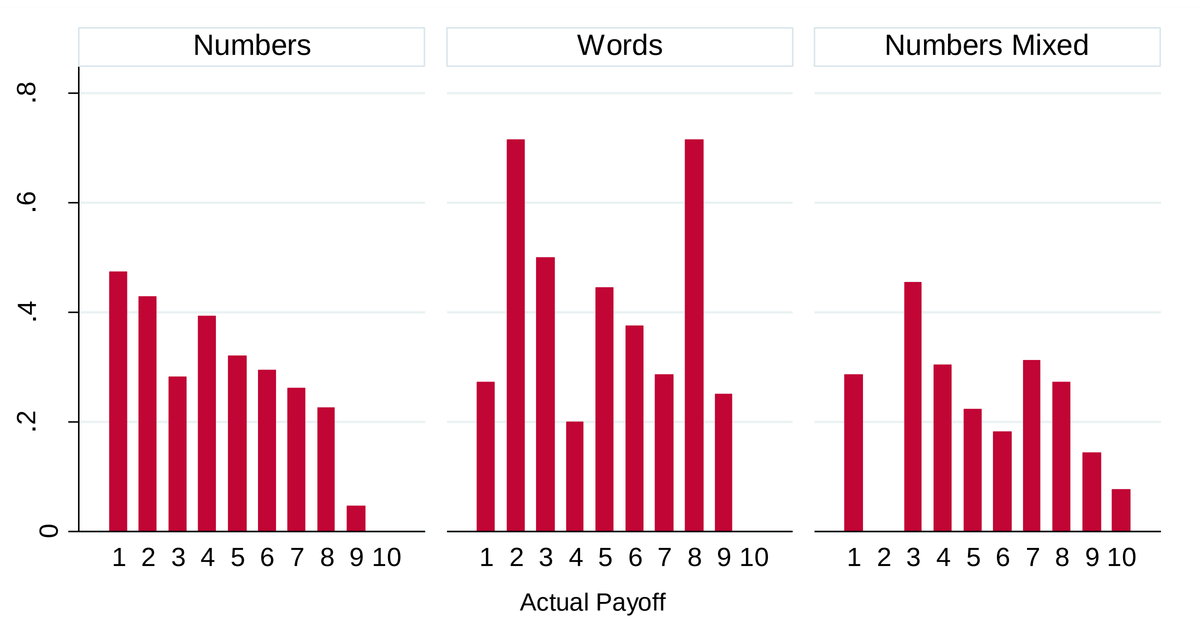


Figure 3: Average payoffs reported by participants who lie

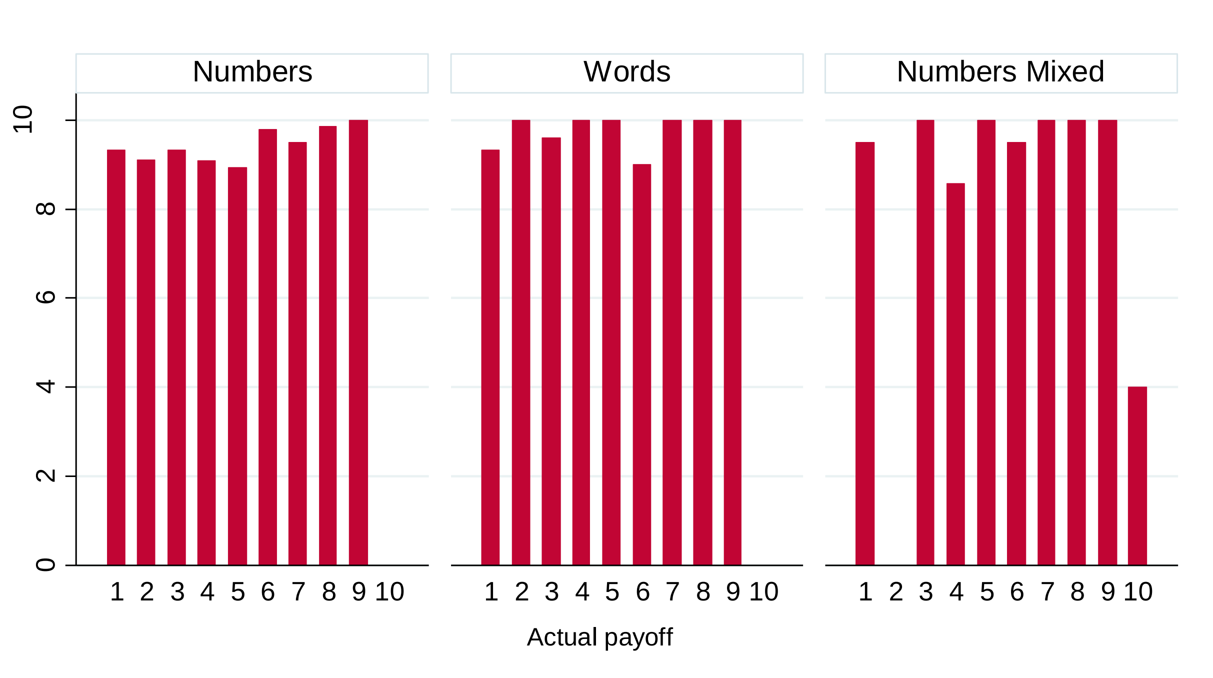


Figure 4: Distribution of payoffs reported by the participants who lie

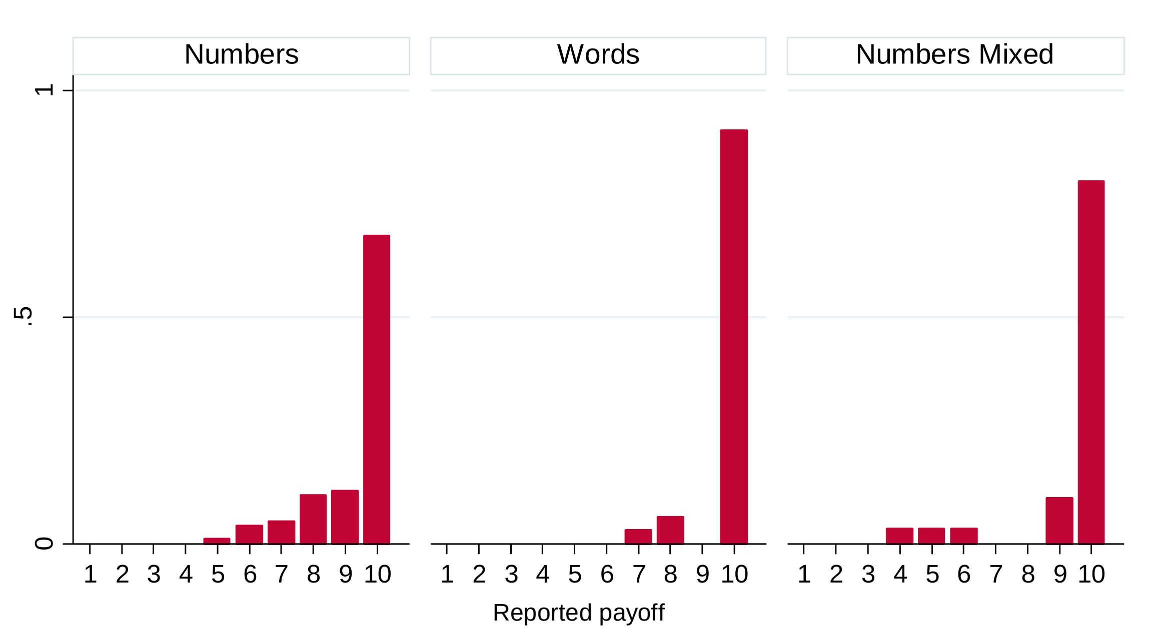
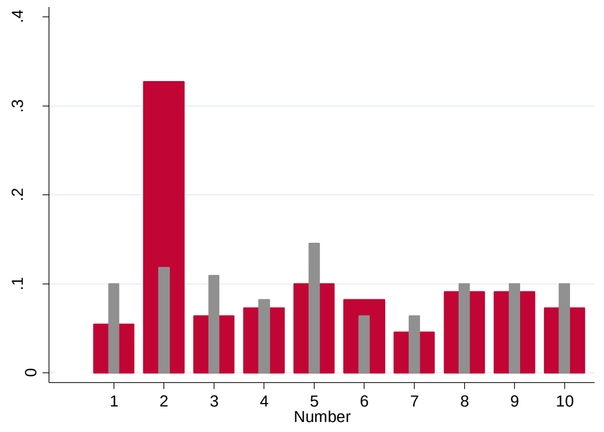
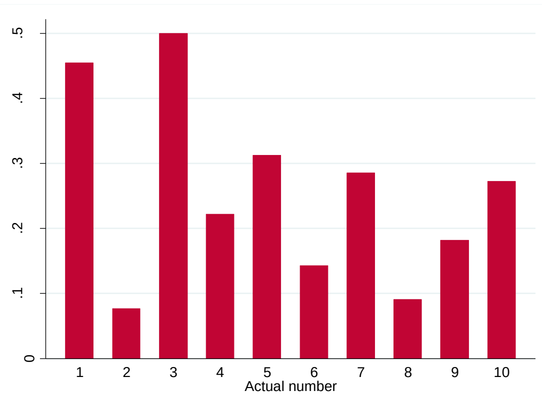


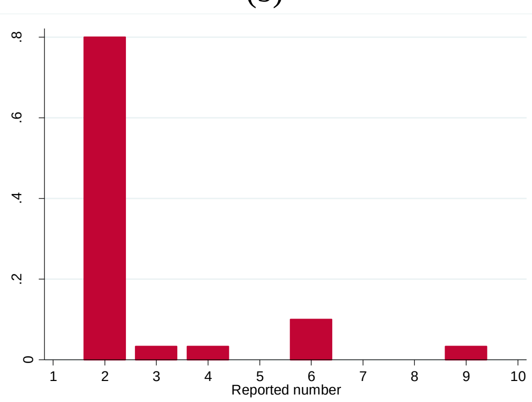
Figure 5: Lying in the Numbers Mixed treatment with respect to the number observed



(a)



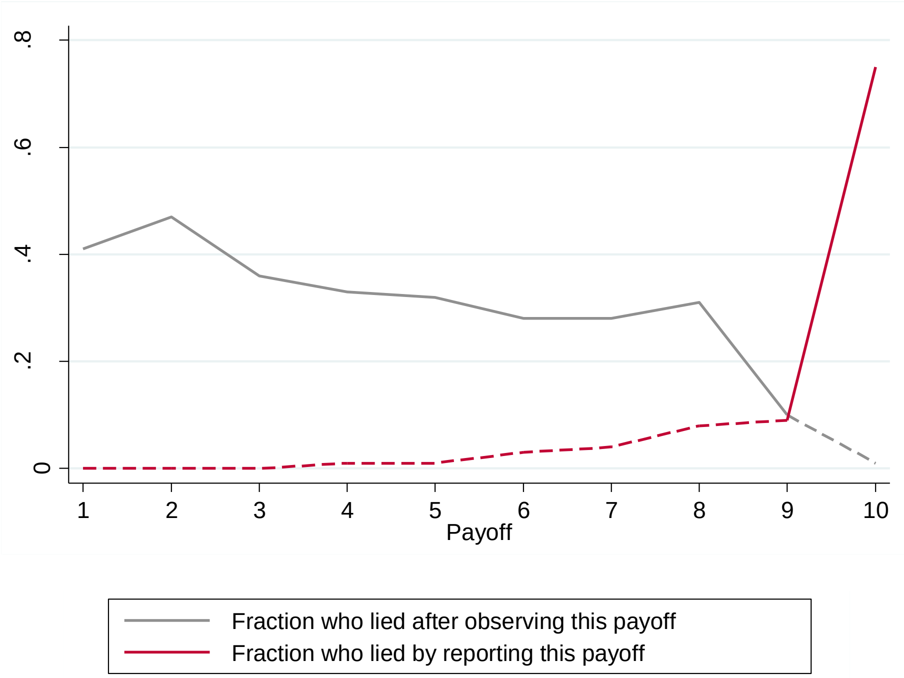
(b)



(c)

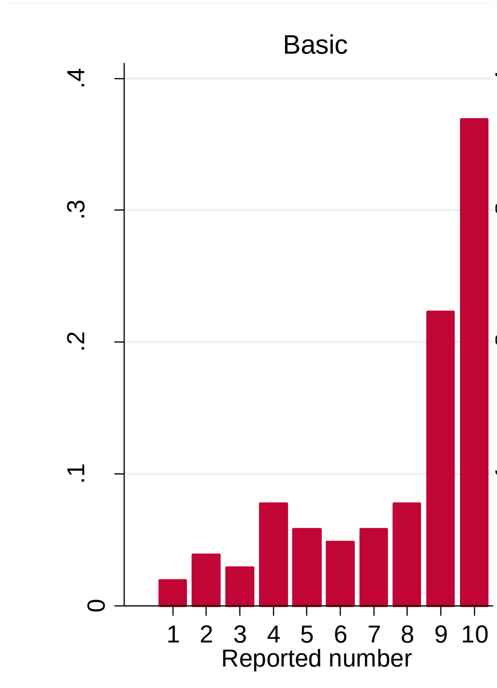
Note: Figure (a) presents the distribution of reported and actual numbers in the Numbers Mixed treatment. The thick dark red bars show the reported numbers, whereas the thin light gray bars show the actual numbers. Figure (b) presents the fraction of lying conditional on the actual number. Figure (c) presents the distribution of numbers reported by the participants who lie.

Figure 6: Fraction of participants who lie after observing a payoff versus fraction of participants who lie by reporting the payoff

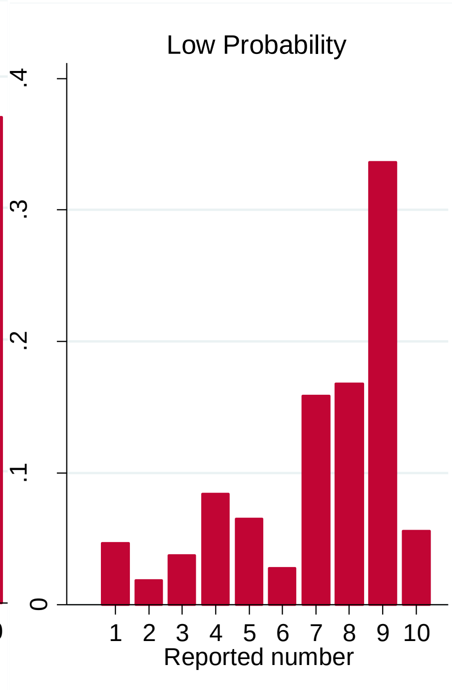


Note: Solid lines denote the behavior which is in line with the theory. Dashed lines denote “mistakes”.

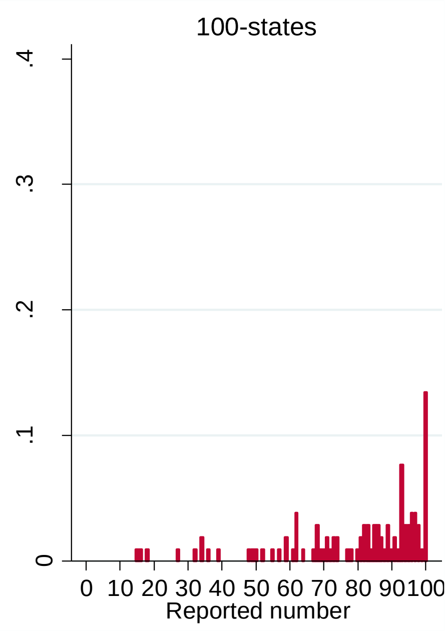
Figure 7: Distribution of reported payoffs in the non-observed treatments



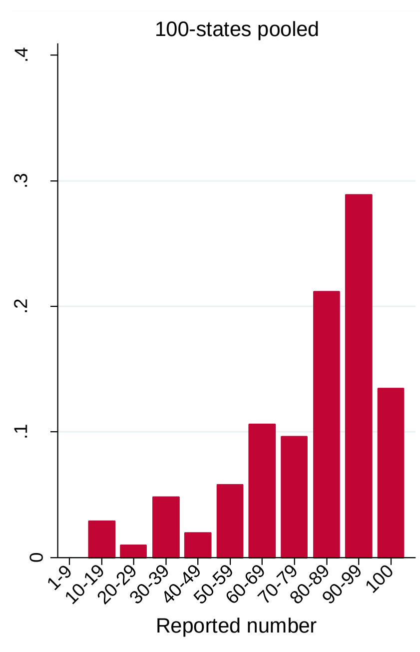
(a)



(b)



(c)



(d)

Note: Figure 7 presents distributions of reported payoffs/numbers in the Basic, Low Probability and 100-states non-observed treatments.

APPENDICES A and B online only

Appendix A: Additional data and analyses

Table A1: Raw data

Treatment	Observed payoff	Reported payoff	Observed payoff	Reported payoff	Observed payoff	Reported payoff
Numbers	1	1	1	10	2	10
	1	1	1	10	2	10
	1	1	2	2	2	10
	1	1	2	2	2	10
	1	1	2	2	2	10
	1	1	2	2	2	10
	1	1	2	2	2	10
	1	1	2	2	2	10
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	1	2	2	3	3
	1	5	2	2	3	3
	1	7	2	2	3	3
	1	7	2	2	3	3
	1	9	2	2	3	3
	1	10	2	2	3	3
	1	10	2	2	3	3
	1	10	2	6	3	3
	1	10	2	6	3	3
	1	10	2	7	3	3
	1	10	2	7	3	3
	1	10	2	8	3	3
	1	10	2	10	3	8
	1	10	2	10	3	8
	1	10	2	10	3	8
	1	10	2	10	3	10
	1	10	2	10	3	10

Treatment	Observed payoff	Reported payoff	Observed payoff	Reported payoff	Observed payoff	Reported payoff
Numbers	3	10	5	5	5	9
	3	10	5	5	5	9
	3	10	5	5	5	9
	3	10	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	5	10
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	4	5	5	6	6
	4	6	5	5	6	6
	4	8	5	5	6	6
	4	9	5	5	6	6
	4	9	5	5	6	6
	4	9	5	5	6	6
	4	9	5	5	6	6
	4	10	5	5	6	6
	4	10	5	5	6	6
	4	10	5	5	6	6
	4	10	5	6	6	6
	4	10	5	7	6	6
	5	5	5	8	6	6
	5	5	5	8	6	6
	5	5	5	8	6	9
	5	5	5	9	6	9

Treatment	Observed payoff	Reported payoff	Observed payoff	Reported payoff	Treatment	Observed payoff	Reported payoff
Numbers	9	9	10	10	Words	1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	1
	9	9	10	10		1	8
	9	9	10	10		1	10
	9	9	10	10		1	10
	9	9	10	10		2	2
	9	9	10	10		2	2
	9	9	10	10		2	10
	9	9	10	10		2	10
	9	9	10	10		2	10
	9	9	10	10		2	10
	9	9	10	10		2	10
	9	10	10	10		3	3
	9	10	10	10		3	3
	10	10	10	10		3	3
	10	10	10	10		3	3
	10	10	10	10		3	3
	10	10	10	10		3	8
	10	10	10	10		3	10
	10	10	10	10		3	10
	10	10	10	10		3	10
	10	10	10	10		3	10
	10	10	10	10		4	4
	10	10	10	10		4	4
	10	10				4	4
	10	10				4	4
	10	10				4	4
	10	10				4	4
	10	10				4	4
	10	10				4	4

Treatment	Observed payoff	Reported payoff	Observed payoff	Reported payoff	Treatment	Observed payoff	Reported payoff
Words	4	10	8	10	Numbers Mixed	0	2
	4	10	8	10		0	2
	5	5	8	10		1	2
	5	5	8	10		1	6
	5	5	9	9		1	7
	5	5	9	9		1	7
	5	5	9	9		1	7
	5	10	9	9		1	7
	5	10	9	9		1	7
	5	10	9	9		3	1
	5	10	9	9		3	1
	6	6	9	9		3	1
	6	6	9	9		3	1
	6	6	9	10		3	1
	6	6	9	10		3	1
	6	6	9	10		3	2
	6	7	10	10		3	2
	6	10	10	10		3	2
	6	10	10	10		3	2
	7	7	10	10		3	2
	7	7	10	10		4	2
	7	7	10	10		4	2
	7	7	10	10		4	2
	7	7	10	10		4	2
	7	7	10	10		4	3
	7	7	10	10		4	3
	7	7	10	10		4	3
	7	7	10	10		4	3
	7	7	10	10		4	3
	7	10	10	10		4	3
	7	10				4	4
	7	10				4	6
	7	10				4	8
	8	8				4	8
	8	8				4	8
	8	10				4	8

Treatment	Observed payoff	Reported payoff	Observed payoff	Reported payoff	Observed payoff	Reported payoff
Numbers Mixed	4	8	7	5	10	2
	4	8	7	5	10	3
	4	8	7	5		
	4	8	7	5		
	4	8	7	5		
	4	8	7	5		
	4	9	7	5		
	5	2	8	2		
	5	2	8	2		
	5	4	8	2		
	5	4	8	10		
	5	4	8	10		
	5	4	8	10		
	5	4	8	10		
	5	4	8	10		
	5	4	8	10		
	6	2	8	10		
	6	6	8	10		
	6	9	9	2		
	6	9	9	6		
	6	9	9	6		
	6	9	9	6		
	6	9	9	6		
	6	9	9	6		
	6	9	9	6		
	6	9	10	2		
	6	9	10	2		
	7	2	10	2		
	7	2	10	2		
	7	2	10	2		
	7	2	10	2		
	7	2	10	2		
	7	5	10	2		
	7	5	10	2		
	7	5	10	2		
	7	5	10	2		

Treatment	Reported payoff	Reported payoff	Reported payoff	Treatment	Reported payoff	Reported payoff
Basic	1	8	10	Low Probability	1	7
	1	8	10		1	7
	2	8	10		1	7
	2	8	10		1	7
	2	8	10		1	7
	2	8	10		2	7
	3	9	10		2	7
	3	9	10		3	7
	3	9	10		3	7
	4	9	10		3	7
	4	9	10		3	7
	4	9	10		4	8
	4	9	10		4	8
	4	9	10		4	8
	4	9	10		4	8
	4	9	10		4	8
	4	9	10		4	8
	5	9	10		4	8
	5	9	10		4	8
	5	9	10		4	8
	5	9	10		5	8
	5	9	10		5	8
	5	9	10		5	8
	6	9	10		5	8
	6	9	10		5	8
	6	9	10		5	8
	6	9	10		5	8
	6	9	10		6	8
	7	9	10		6	8
	7	10	10		6	9
	7	10	10		7	9
	7	10			7	9
	7	10			7	9
	7	10			7	9
	8	10			7	9
	8	10			7	9

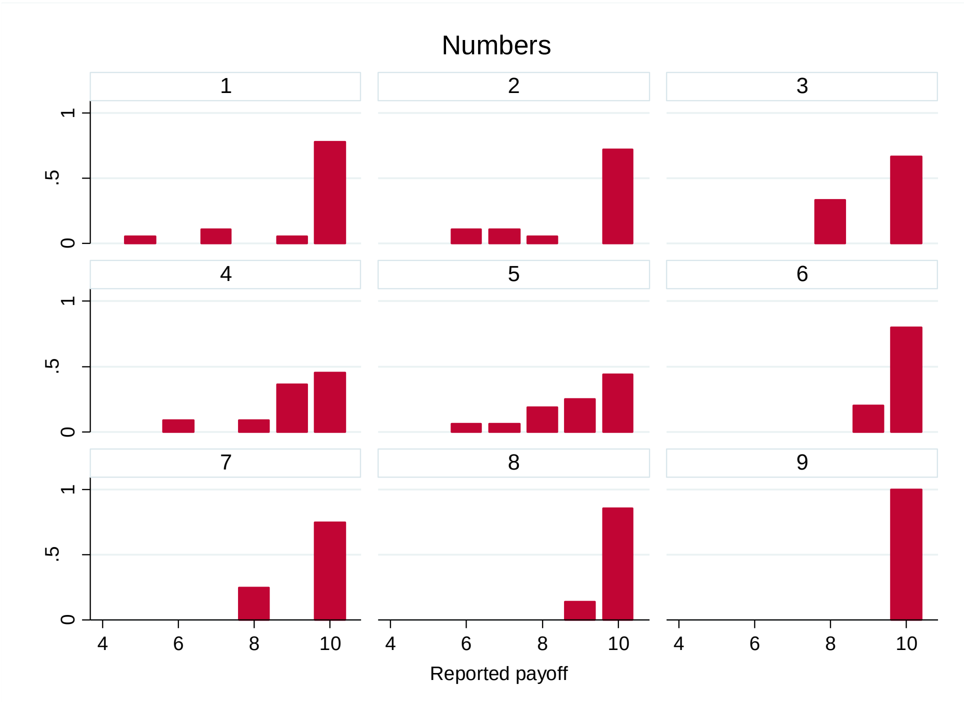
Treatment	Reported payoff	Treatment	Reported payoff	Reported payoff	Reported payoff
Low Probability	9	100-states	15	77	94
	9		16	78	94
	9		18	80	94
	9		27	81	95
	9		32	81	95
	9		34	82	95
	9		34	82	96
	9		36	82	96
	9		39	83	96
	9		48	83	96
	9		49	83	97
	9		50	84	97
	9		52	85	97
	9		55	85	97
	9		57	85	98
	9		59	86	98
	9		59	86	98
	9		61	86	99
	9		62	87	100
	9		62	87	100
	9		62	88	100
	9		62	89	100
	9		64	89	100
	9		67	89	100
	9		68	90	100
	9		68	91	100
	9		68	91	100
	9		69	92	100
	9		70	93	100
	10		71	93	100
	10		71	93	100
	10		72	93	100
	10		73	93	
	10		73	93	
	10		74	93	
			74	93	

Table A2: The fraction of people who lie after observing a payoff and the fraction who lie by reporting that payoff, for each treatment

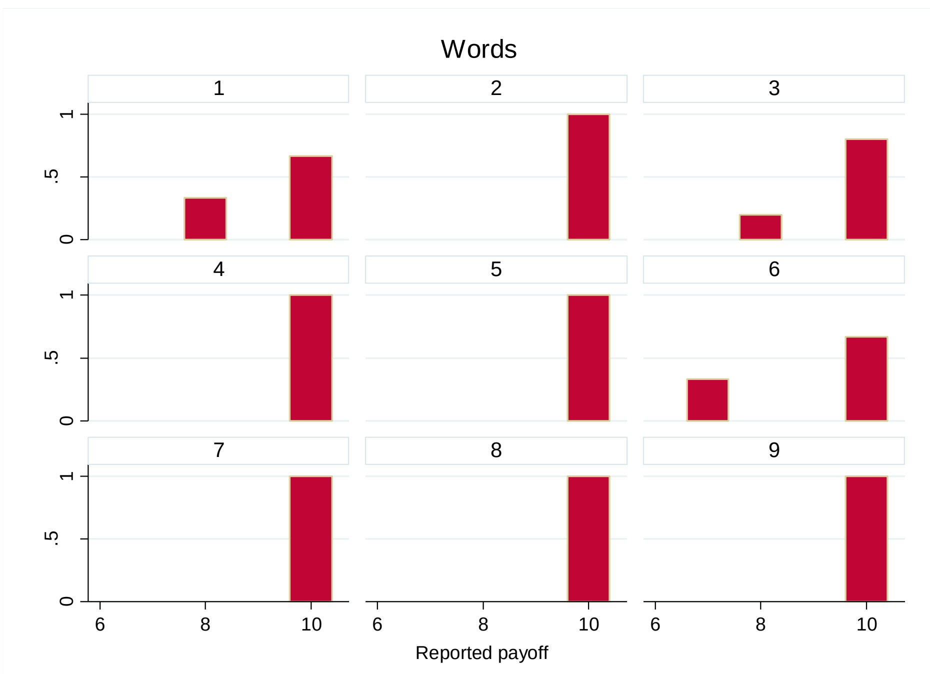
Payoff	Fraction who lied after observing this payoff				Fraction who lied by reporting this payoff			
	Numbers	Words	Numbers Mixed	Pooled	Numbers	Words	Numbers Mixed	Pooled
1	47.37%	27.27%	28.57%	41.07%	0%	0%	0%	0%
2	42.86%	71.43%	-	46.94%	0%	0%	0%	0%
3	28.13%	50%	45.45%	35.85%	0%	0%	0%	0%
4	39.29%	20%	30.43%	32.79%	0%	0%	3.33%	0.60%
5	32%	44.44%	22.22%	32.35%	0.97%	0%	3.33%	1.20%
6	29.41%	37.5%	18.18%	28.30%	3.88%	0%	3.33%	2.99%
7	26.01%	28.57%	31.25%	27.63%	4.85%	2.94%	0%	3.59%
8	22.58%	71.43%	27.27%	30.61%	10.68%	5.88%	0%	7.78%
9	4.65%	25%	14.29%	9.68%	11.65%	0%	10%	8.98%
10	0%	0%	7.69%	1.37%	67.96%	91.18%	80%	74.85%

Note: The results presented in the table are measured in payoffs. On the left side of the table, we report what fraction of participants lie after observing a particular payoff (from observing a payoff of 1 to 10 down the table) and on the right side of the table, we report the fraction of participants who lie by reporting a particular payoff (from reporting a payoff of 1 to 10 down the table).

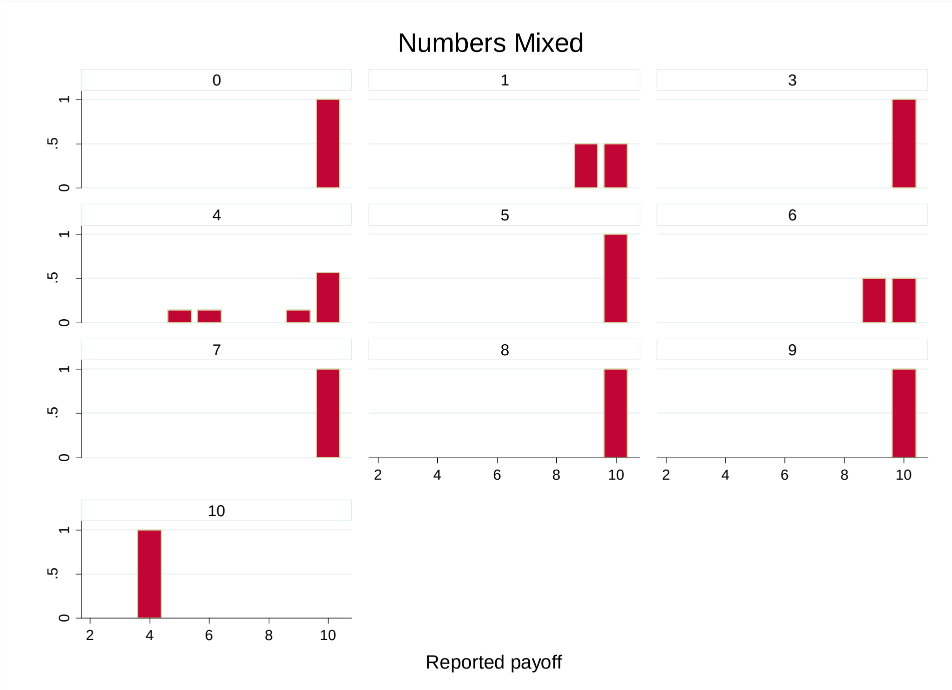
Figure A1: Distribution of payoffs reported by participants who lie, for each actual payoff



(a)



(b)



(c)

Note: Figures 5 a-c represent the distribution of reported payoff conditional on lying for each actual payoff separately. The number on the top of the cell stands for the actual payoff, and the numbers on the x-axis, for the reported payoffs. The number of observations amounts to 103, 34 and 30 in (a), (b) and (c), respectively.

Appendix B: Instructions

Translated from German

Treatment: Numbers

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 2.50 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 2.50 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

On the next screen you will see 10 boxes with numbers hidden behind them. The numbers in the boxes are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 and they are placed in a random order. We will ask you to click on one box.

Once you click on the box, you will see a number that we ask you to remember and later report to us.

The number you report determines how much money you will be paid. You will be paid the equivalent in Euros to the number you report. In other words, if you write “1”, you receive 1€, If you write “2”, you receive 2€, if you write “3”, you receive 3€ and so on.

If you have any questions please raise your hand and we will come to you!

/Screen 3/

Please click on one of the boxes: _____

Please remember the number you saw.

/Screen 4/

Please report the number that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the number that you saw: _____

Your cubicle number: ______

/Post-questionnaire/

Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____

Treatment: Words

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 2.50 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 2.50 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

On the next screen you will see 10 boxes with Lithuanian words hidden behind them. The words in the boxes are „vilkas“, „miskas“, „dangus“, „stirna“, „rojuje“, „siaure“, „kiskis“, „alyvos“, „obelis“, „pietus“and they are placed in a random order. We will ask you to click on one box. Once you click on the box, you will see a Lithuanian word that we ask you to remember and later report to us.

The word you report determines how much money you will be paid. You will be paid according to the word you report. If you write “vilkas”, you receive 1€, if you write “miskas”, you receive 2€; if you write “dangus”, if you write, you receive 3€; if you write “stirna”, you receive 4€; if you write “rojuje”, you receive 5€; if you write “siaure”, you receive 6€; if you write “kiskis”, you receive 7€; if you write “alyvos”, you receive 8€; if you write “obelis”, you receive 9€; if you write “pietus”, you receive 10€.

If you have any questions please raise your hand and we will come to you!

/Screen 3/

Please click on one of the boxes: _____

Please remember the word you saw.

/Screen 4/

Please report the word that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the word that you saw: _____

Your cubicle number: ______

/Post-questionnaire/

Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____

Treatment: Numbers Mixed

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 4 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 4 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

On the next screen you will see 10 boxes with numbers hidden behind them. The numbers in the boxes are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 and they are placed in a random order. We will ask you to click on one box.

Once you click on the box, you will see a number that we ask you to remember and later report to us.

The number you report determines how much money you will be paid. You will be paid according to the number you report. If you write “7”, you receive 1€; if you write “3”, you receive 2€; if you write “1”, you receive 3€; if you write “8”, you receive 4€; if you write “4”, you receive 5€; if you write “9”, you receive 6€; if you write “5”, you receive 7€; if you write “10”, you receive 8€; if you write “6”, you receive 9€; if you write “2”, you receive 10€.

If you have any questions please raise your hand and we will come to you!

/Screen 3/

Please click on one of the boxes: _____

Please remember the number you saw.

/Screen 4/

Please report the number that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the number that you saw: _____

Your cubicle number: ______

/Post-questionnaire/

Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____

Treatment: Basic

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 2.50 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 2.50 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

We will give you an envelope with 10 pieces of paper in it. The numbers on the paper pieces are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 and they are placed in a random order. We will ask you to take out one piece of paper and later to put it back in the envelope.

Once you take a piece of paper out of the envelope, you will see a number that we ask you to remember and later report to us.

The number you report determines how much money you will be paid. You will be paid the equivalent in Euros to the number you report. In other words, if you write “1”, you receive 1€. If you write “2”, you receive 2€, if you write “3”, you receive 3€ and so on.

If you have any questions please raise your hand and we will come to you!

/Screen 3/

Please take one number out of envelope.

Please remember the number you saw.

Please put the piece of paper back in the envelope.

/Screen 4/

Please report the number that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the number that you saw: _____

Your cubicle number: ______

/Post-questionnaire/

Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____
____
____

Treatment: Low probability

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 2.50 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 2.50 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

We will give you an envelope with 100 pieces of paper in it. The numbers on the paper pieces are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 and they are placed in a random order.

There are 11 pieces of paper with the number "1" on it, 11 pieces with the number "2", and so on until "9". There is only 1 piece of paper with the number "10" on it. We will ask you to take out one piece of paper and later put it back in the envelope.

Once you take a piece of paper out of the envelope, you will see a number that we ask you to remember and later report to us.

The number you report determines how much money you will be paid. You will be paid the equivalent in Euros to the number you report. In other words, if you write “1”, you receive 1€, If you write “2”, you receive 2€, if you write “3”, you receive 3€ and so on.
If you have any questions please raise your hand and we will come to you!
/Screen 3/

Please take one number out of envelope.

Please remember the number you saw.

Please put the piece of paper back in the envelope.

/Screen 4/

Please report the number that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the number that you saw: _____

Your cubicle number: ______

/Post-questionnaire/
Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____
____
____

Treatment: 100-states

/Screen 1/

Welcome to our experiment. Please read the instruction carefully. If you have a question, please raise your hand and we will come over to you. Please do not communicate with other participants during the experiment.

Every participant will receive 2.50 Euros for attending, which will be paid out independently of the decisions made in the experiment.

Furthermore, you will be able to earn additional money. At the end of the experiment, you will receive the income which you earned over the course of the experiment plus the 2.50 Euros for attending in cash.

Your decisions are private and no other participant will know about them.

/Screen 2/

We will give you an envelope with 100 pieces of paper in it. The numbers on the paper pieces are all the natural numbers between 1 and 100 (i.e. 1, 2, 3,…., 98, 99 and 100) and they are placed in a random order. We will ask you to take out one piece of paper and later put it back in the envelope. Once you take a piece of paper out of the envelope, you will see a number that we ask you to remember and later report to us.

The number you report determines how much money you will be paid. You will be paid the equivalent in Euros to the number you report divided by 10. In other words, if you write "1", you receive 0.1€. If you write "2", you receive 0.2€, if you write "3", you receive 0.3€ and so on.
If you have any questions please raise your hand and we will come to you!
/Screen 3/

Please take one number out of envelope.

Please remember the number you saw.

Please put the piece of paper back in the envelope.

/Screen 4/

Please report the number that you saw on the sheet of paper that we will give to you soon. Please also write down your cubicle number on that sheet.

We will collect the sheets from every participant, as soon as everybody is done and will prepare the payoffs.

/On paper/

Please report the number that you saw: _____

Your cubicle number: ______

/Post-questionnaire/

Please answer following questions.

Age: _____

Gender: _____

Major: _____

Semester: _____

Please describe briefly how you made your decisions in this experiment: ____