

Loss aversion and lying behavior

Ellen Garbarino, Robert Slonim, Marie Claire Villeval

▶ **To cite this version:**

Ellen Garbarino, Robert Slonim, Marie Claire Villeval. Loss aversion and lying behavior. Journal of Economic Behavior and Organization, 2019, 258, pp. 379-393. 10.1016/j.jebo.2018.12.008 . halshs-01981542

HAL Id: halshs-01981542
https://shs.hal.science/halshs-01981542

Submitted on 22 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S016726811830338X
Manuscript_5c77f4ba5fd6f1e063bc91119d6b43a0

Loss Aversion and lying behavior

Ellen Garbarino[a], Robert Slonim[b], Marie Claire Villeval[c]

October 15, 2018

Abstract:
We theoretically show that loss-averse agents are more likely to lie to avoid receiving a low payoff after a random draw the lower the *ex-ante* probability of this bad outcome. The *ex-ante* expected payoff increases as the bad outcome becomes less likely, and hence the greater is the loss avoided by lying. We demonstrate robust support for this theory by reanalyzing the results from the extant literature and with a new experiment that varies the outcome probabilities and is run double-anonymous.

Keywords: Loss aversion, dishonesty, lying, experiments
JEL-codes: C91, C81, D03

[a] University of Sydney Business School, Department of Marketing, Abercrombie building, Sydney, NSW 2006, Australia. Email: ellen.garbarino@sydney.edu.au

[b] University of Sydney, Department of Economics, Merewether building, Sydney, NSW 2006, Australia; IZA, Bonn, Germany. Email: robert.slonim@sydney.edu.au

[c] Univ Lyon, CNRS, GATE (UMR5824), 93, Chemin des Mouilles F-69130, Ecully, France; IZA, Bonn, Germany. E-mail: villeval@gate.cnrs.fr

We thank the authors of the studies on lying who shared their data with us. We are grateful to C. Anthony, C. de Dreu, C. Gravert, D. Hugh-Jones, A. Kajackaite, M. Rigdon, I. Rodriguez-Lara, D. Rustagi, S. Shalvi, J. van de Ven, and participants at the European Workshop of Experimental and Behavioral Economics in Cologne, the ESA world meeting in Jerusalem, the SITE conference at Stanford University, the ESA European meeting in Bergen, the 10th M-BEES Conference in Maastricht, the Psychology and Economics workshop in Friedrichshafen, and at seminars at Aix-Marseille University, University of Amsterdam, University of Alaska in Anchorage, Central University of Finance and Economics in Beijing, University of Geneva, University of Innsbruck, Jiao-Tong University in Xi'an, Middlesex University, University of Nottingham, Peking University, Penn State University, Renmin University in Beijing, University of Trier, Tsinghua University, Turku University, and University of Vienna for thoughtful comments. Financial support from the University of Sydney and the FELIS program of the French National Agency for Research (ANR-14-CE28-0010-01) is gratefully acknowledged. This research was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR). It has been initiated while E. Garbarino and R. Slonim were visiting GATE at Ecole Normale Supérieure de Lyon.

© 2018 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

1. Introduction

Dishonesty is pervasive and deters economic growth and development.[1] Going beyond the initial economics-of-crime approach (Becker, 1968), economists have broadened their study of the behavioral determinants of dishonesty. This emerging literature shows that the extent to which individuals are or are not willing to follow the moral course of action can be affected by ethical values, social preferences, institutions, and norms (*e.g*., Gneezy 2005; Ariely 2012; Gächter and Schultz 2016; Mann *et al*. 2016).[2] Most of the literature on the effect of incentives have focused on the level of incentives and of their nature (*e.g*., individual, team-based or competitive),[3] while less is known about how the structure of the payoff matrix, such as the probability of each payoff, affects dishonesty. This is surprising, as economists have demonstrated that individuals actively try to avoid unlikely low payoffs (*e.g*., by investing in effort or by buying insurance). Hence, it is likely that some individuals may also choose to behave dishonestly when given the opportunity to avoid a lower payoff (*e.g*., cheating to get a bonus, exaggerating insurance claims after an accident, *etc*.). To fill this gap, our two contributions are to theoretically and empirically study the role of payoff probabilities on dishonesty accounting for the moral costs for lying.

Theoretically, we explore a determinant of dishonesty that is predicted by loss aversion and which operates through the *probability* of the privately observed outcomes that people can either report on honestly or dishonestly. We show that loss aversion predicts that the extent to which individuals behave dishonestly is sensitive to the probability of observed outcomes because this

[1] For example, the overall gross tax gap estimate by the IRS in the U.S represents about 16% of estimated actual tax liability. The global hidden financial assets in offshore tax havens represent up to $280 billion in lost income tax revenues. The World Bank Institute estimates that corruption is equal to about 2% of global GDP. Evaded taxes, corruption and embezzlement deter investment, and therefore hinder growth.

[2] For recent surveys on the experimental economic literature on dishonest behavior, see Rosenbaum *et al*. (2014), Irlenbusch and Villeval (2015), Abeler *et al*. (2018), Dhami (2017), and Jacobsen *et al*. (2018).

[3] Studies examining the level of incentives on dishonesty include, *e.g*., Fischbacher and Föllmi-Heusi (2013) and Kajackaite and Gneezy (2017). Studies examining the nature of incentives on dishonesty include, *e.g*., Jacob and Levitt (2003), Conrads *et al*. (2013), Danilov *et al*. (2013), Charness *et al*. (2014) and Faravelli *et al*. (2015).

probability affects the payoff that is expected to be observed; the expected payoff in turn affects the loss-averse utility of honest and dishonest reports. If individuals suffer more from losses than they enjoy equivalent gains, as loss aversion contends (Kahneman and Tversky 1979; 1992), then they gain more utility from being dishonest the further the realized outcome is below the payoff that they expected to observe. Thus, the less likely a bad outcome is expected to occur, the more this bad outcome will be below the expected outcome, and hence the greater the loss in utility that is avoided by lying.

Only a few studies have investigated the role of loss aversion in lying, theoretically and empirically. Most experimental studies of loss aversion and lying manipulate the endowment to vary the reference state and most find support for greater dishonesty in the loss than gain frame (Shalvi 2012; Cameron and Miller 2013; Schindler and Pfattheicher 2016; Balasubramanian *et al*. 2017; Grolleau *et al*. 2017; for an exception, see Charness *et al*. 2018).[4] Instead of manipulating the reference state, we vary the distribution of outcomes. Two recent studies on lying also manipulate the distribution of true states. In Abeler *et al.* (2018), subjects draw one chip out of 50 chips that pays either a low or a high payoff. They find that people cheat more when the number of high payoff chips is 30 rather than 5. In Gneezy *et al*. (2018), subjects take out one piece of paper from an envelope containing pieces of paper with different values. The manipulation of the prior distribution of states shows that people make more partial lies when the probability of the highest outcome decreases. While these findings are consistent with loss aversion, the authors discuss their results as evidence of the role of reputation in lying. The logic of the reputation model is that an agent's utility

[4] Using real-effort tasks, Cameron and Miller (2013) and Grolleau *et al*. (2017) find that individuals misreport their performance more in a loss than in an equivalent gain frame. Schindler and Pfattheicher (2016) also find, using both a die-rolling task and a coin-tossing task, that more people cheat in a loss than gain frame. In contrast, Charness *et al*. (2018) find that people cheat in a gain frame, but not in a loss frame. In a different approach, Shalvi (2012) examines lying to affect the odds of winning and shows that individuals dishonestly increase the likelihood of wining by turning negative gambles into positive gambles. Finally, Balasubramanian *et al*. (2017) consider instead a reference point related to expected daily income on MTurk and find reduced dishonesty at higher incentive levels driven by those subjects who have already exceed their stopping point.

depends on the probability that he is perceived as a liar when he reports a given outcome; when the likelihood of the high payoff outcome becomes small, the probability that the high report is honest goes down because a high report generates more suspicion from the audience, hence dishonest agents with image concerns may refrain from lying to gain reputation (on reputation and lying, see also Khameltski and Sliwka 2017; Dufwenberg and Dufwenberg 2018). While consistent with reputation, the evidence from these studies is also consistent with showing that more lying occurs when the probability of getting the bad outcome is lower.[5]

Our first contribution is to provide a theoretical model of loss aversion and lying in which we analyze the variation in the reference state resulting from differences in the probabilities of privately observed outcomes. In our model, a dishonest report generates an intrinsic moral cost as well as a gain in utility from the payoffs, where variations in the probability of the outcomes affect the relative size of the gains and losses, compared to the reference state, that occur between reporting honestly and dishonestly. Hence, our model provides a formal theoretical relationship between loss aversion and lying in the presence of moral costs of dishonesty. While this model is compatible with reputation concerns, we instead focus on the role of loss aversion. Our model differs from the loss aversion model of Abeler *et al*. (2018) in the way the reference point is defined; Abeler *et al.* use a notion of the reference point *à la* Köszegi and Rabin (2006) with gain-loss utility both in the monetary dimension and in the lying cost dimension, whereas in our model the reference point is defined as the expected outcome of the random device if the agent reports what he observes.

Our second contribution is empirical and experimental. We first re-examine the extensive experimental literature on dishonesty in which the outcome of a random draw is reported to the experimenter. In this literature, the probabilities of the outcomes and the rate of lying vary

[5] Subjects reported an outcome with 50% chance of being paid, with the amount of payment manipulated. The random draw was made only after subjects entered their report. The results reject that people lie more when the fixed payment is higher (Abeler *et al*. 2018).

dramatically across studies, allowing us to test the implications of our model. We start from the same studies as those surveyed by Abeler *et al*. (2018), but the objectives of the two reviews differ: while Abeler *et al*. (2018) try to identify the individual, economic and methodological determinants of misreporting, we focus on the impact of the distribution of probabilities of the low and high outcomes on the lying rate. We further test the theoretical implications of our model by conducting a new experiment that controls for factors other than the probabilities of the payoff outcomes that subjects can observe and that minimizes the individuals' concerns for reputation.

Empirically studying the likelihood that individuals cheat requires a valid statistical measure of dishonesty. Because statistics on fraud in natural settings focus on people who have been detected, which introduces a selection bias, researchers have increasingly turned to controlled experiments. The most common technique to detect cheating involves subjects generating an outcome from a random device (*e.g*., flipping a coin, rolling a die, drawing a ticket) where the outcome can only be observed by the subject (*e.g*., Fischbacher and Föllmi-Heusi 2013; Shalvi *et al*. 2011). If a subject privately observes an outcome with a payoff that is less than the maximum possible payoff, then he has a financial incentive to dishonestly report a higher outcome than observed.[6]

Virtually every paper using this technique compares the proportion of each reported outcome with the theoretical proportion if subjects reported truthfully. While this approach indicates whether the proportion of subjects reporting any particular outcome differs from the expected proportion if all subjects reported honestly, it does not describe the distribution of the proportion of subjects that behaved dishonestly. It is thus difficult to compare the extent of lying across studies beyond indicators for when dishonesty does and does not significantly occur. This limitation motivates our use of a technique that was developed in Garbarino *et al.* (2018) to estimate the full distribution of

[6] The artificiality of the laboratory may, however, affect behavior. Reassuringly, the literature shows that dishonesty in the lab correlates with dishonesty in the field (Barfort *et al*. 2015; Cohn *et al*. 2015; Potters and Stoop 2016; Hanna and Wang 2017; Dai *et al*., 2018; Cohn and Maréchal, Forthcoming).

the proportion of individuals who lie. This technique provides an estimate on the full distribution of lying, including the mean and confidence intervals, indicating lower and upper bounds on the proportion of people lying.

Using this new estimation technique, we first test our theory on the extant lying literature. We estimate the mean percent of individuals reporting dishonestly from all studies exploring lying using random devices we were able to obtain data from. Our review includes 81 studies representing 389 treatments and 36,668 individuals from 44 countries. We find a highly significant and robust negative correlation, ranging from -0.42 to -0.52 depending on the specifications, between the probability of observing a low payoff outcome and the mean lying propensity. The mean proportion of individuals that behave dishonestly to avoid a low payoff outcome is higher when this outcome is less likely.[7]

While this evidence is consistent with the loss aversion hypothesis, concerns about one's reputation might lead to similar results since they predict an equilibrium in which individuals are more likely to lie to get a higher payoff the less likely the claim for a higher payoff might be perceived as lying by observers. We address the potential confounding of reputation and loss aversion in two ways. First, we test with the data from the extent literature as to whether the impact of the probability of the bad outcome varies with the level of observability of the subject's identity by others when making reports. Unsurprisingly, we find support for a main effect of reputation: subjects are less likely to lie in treatments with moderate and high observability than in treatments with low observability. However, even after controlling for the level of observability, the relationship between the probability of the bad outcome and lying remains significant. Nor do we find an interaction between observability and the probability of the bad outcome, demonstrating that the

[7] For example, subjects lie more in a die task where only one value earns zero payoff (like in Fischbacher and Föllmi-Heusi (2013) in which the mean lying rate to avoid the zero payoff is 61%) than when five out of six sides earn zero (like in Kajackaite and Gneezy (2017) in which the mean lying rate is only 15%). Our results hold regardless of whether we define a bad payoff outcome as the lowest possible outcome or all outcomes except the highest one.

effect of the probability of the bad outcome on lying is the same regardless of the level of observability.

Second, to directly test the loss aversion lying hypothesis, to remove across experiment differences in the extant literature, and to minimize the potentially confounding influence of reputation, we designed a novel experiment using a "mind coin-tossing game" in which we vary only the likelihood of earning either $0 or $2. 978 subjects participated in this experiment that was run on-line on MTurk and was double-anonymous. Individuals have to predict the outcome to themselves before tossing a coin in private. This task is repeated three times and then subjects have to report the number of correct predictions, which determines their payoff. Across subjects, we vary the number of correct predictions (either 1, 2 or 3) required to earn the $2, which varies the probability of the undesired outcome (earning $0) from 12.5% to 50% to 87.5%.

The results of our experiment show that, consistent with our loss aversion prediction, people lie more often to avoid the undesired outcome when its probability is lower. On average 76% of the subjects lie to get a $2 rather than $0 payoff when the likelihood of getting the low payoff is 12.5%, whereas 66% lie when the likelihood of the low payoff is 50% and only 44% lie when the likelihood is 87.5%.

Our review of the extant literature and our experiment confirm that individuals are more likely to lie when the undesired outcome is more unlikely, consistently with our theoretical model of loss aversion. This suggests practical implications in terms of policy making; in particular, audits and controls should focus on settings with low probability of occurrence.

The remainder of this paper is organized as follows. Section 2 presents the theory. In Section 3 we estimate the distribution of the percentage of subjects who misreport in the existing literature. Section 4 presents our new experiment. Section 5 discusses our results and concludes.

2. A theory of Loss Aversion and the probability of outcomes on lying

This section shows theoretically that in a setting with two possible outcomes, agents with loss averse preferences and an intrinsic cost of lying who privately observe the low payoff outcome will be increasingly likely to lie the smaller the *ex-ante* probability of their experiencing this low payoff outcome. To show this, we examine the comparative static effect of changing the probability of observing the low payoff on the difference in the utility of the agent dishonestly reporting the not-observed high payoff *vs*. honestly reporting the observed low payoff.

Our model assumes homogeneous preferences over payments. Let x_1 and x_2 ($x_2 > x_1$) be an agent's monetary payoffs if he reports the low and high payoff outcomes, respectively, where the *ex-ante* probability of observing the outcomes that pay x_1 and x_2 is p and 1-p. In this context, the *ex-ante* expected observed payoff is $x_e = px_1 + (1\text{-}p)\, x_2$; *i.e.*, x_e is the expected observed payoff before observing how the random event is resolved. We show that if an agent has loss-averse preferences (Kahneman and Tversky, 1979; 1992) and uses the *ex-ante* expected observed payoff x_e as his reference point, then after observing the low payoff outcome he will be increasingly more likely to dishonestly report the high payoff outcome than to honestly report the low payoff outcome the less likely the *ex-ante* probability of observing the low payoff outcome (*i.e.*, the smaller p is). Intuitively, as the likelihood of observing the low payoff outcome decreases, the *ex-ante* expected observed payoff (the reference point) increases, so the sense of loss related to honestly reporting a bad outcome increases. As the probability of the low payoff outcome decreases, the person's expectation of receiving the higher payoff increases and so he increasingly suffers a greater loss if he reports the low payoff outcome, and hence is more inclined to dishonestly report the higher payoff outcome to avoid that loss. As the likelihood of observing the *ex-ante* low payoff outcome decreases, the gain in utility for a loss-averse agent who reports dishonestly to avoid the loss increases, increasing the benefits of a dishonest *vs*. honest report. We now formally show this.

We assume an agent's utility from reporting a certain outcome compared to the observed outcome is a function of the intrinsic moral cost of dishonesty $m()$ and the reference-dependent loss-averse utility $g()$ and $h()$ of the monetary payoff:

$$U(x_R \mid x_O) = g(x_2 - x_e)*\mathrm{I}(x_R \geq x_e) - h(x_e - x_1)*\mathrm{I}(x_R < x_e) - m(x_R - x_O)*\mathrm{I}(x_R \neq x_O) \quad (1)$$

where I() is an indicator function that equals 1 if true and 0 if false, x_O is the agent's payoff associated with his privately observed outcome from the random process, and x_R is the agent's payoff associated with his reported outcome.

The functions $g()$, $h()$ and $m()$ are continuous and twice differentiable; $g()$ and $h()$ are the utility of gains and losses, respectively, and $m()$ is the moral cost of reporting an outcome that was not observed by the agent (this is an intrinsic cost that does not depend on how others perceive the individual). We assume:

$$m'(x) \geq 0 \text{ and } m''(x) \geq 0 \quad (2)$$

Assumption (2) reflects the (weakly) increasing moral cost of lying as a function of the gain from misreporting. For instance, Mazar *et al*. (2008), Lundquist *et al*. (2009), and Fischbacher and Föllmi-Heusi (2013) find evidence consistent with $m'(x)>0$, while Kajackaite and Gneezy (2017) cannot reject that $m'(x)=0$. Note that using a fixed moral cost (like in Khameltski and Sliwka, 2017) would not change our conclusion on the effect of p on the decision to misreport since there are only two possible reports in the model.[8]

If the agent observes the low payoff outcome (then, $x_O = x_1$), he has a monetary incentive to dishonestly report observing the higher payoff outcome (thus, $x_R = x_2$). It follows from equation (1)

[8] The model could also be extended to capture the fact that reporting a low payoff outcome honestly when facing temptation may generate a positive feeling (*e.g*., pride, sense of honor) or reinforcement of positive moral identity. To represent the moral benefit of reporting the low payoff outcome honestly, one could add the following term to equation (1): $+n(x_2 - x_O)*\mathrm{I}(x_R = x_O)$. This term would remain only when the agent reports truthfully the low payoff outcome, and it would increase in the size of the temptation $(x_2 - x_O)$.

that when observing the low payoff outcome, the agent's utility if he reports the low or high payoff outcome, respectively, is:[9]

$$U(x_1 \mid x_1) = - h(x_e - x_1) \quad (3)$$

$$U(x_2 \mid x_1) = g(x_2 - x_e) - m(x_2 - x_O) \quad (4)$$

It follows that the difference in utility U(D) between dishonestly reporting observing the high payoff outcome (and earning x_2) and honestly reporting observing the low payoff outcome (and earning x_1) is:

$$\text{U(D)} = U(x_2 \mid x_1) - U(x_1 \mid x_1) = g(x_2\text{-}x_e) + h(x_e\text{-}x_1) - m(x_2\text{-}x_1) \quad (5)$$

Equation 5 highlights how the utility gained from the higher payoff by an agent with reference-dependent loss-averse preferences when reporting dishonestly *vs*. honestly can be decomposed into two components. First, he gains utility from the monetary gain above the reference point, $g(x_2\text{-}x_e)$, and second, he gains utility from avoiding the monetary loss below the reference point $h(x_e\text{-}x_1)$.

Substituting for $x_e = px_1 + (1\text{-}p)x_2$ and rearranging, the difference in utility between dishonestly reporting observing the high payoff outcome and honestly reporting observing the low payoff outcome is:

$$\text{U(D)} = g(p^*(x_2\text{-}x_1)) + h((1\text{-}p)^*(x_2\text{-}x_1)) - m(x_2\text{-}x_1) \quad (6)$$

Differentiating U(D) with respect to the probability of the low payoff outcome (p), we have:

$$\text{U'(D)} = (x_2\text{-}x_1)[g'(p(x_2\text{-}x_1)) - h'((1\text{-}p)(x_2\text{-}x_1))], \text{ and} \quad (7)$$

[9] We assume that if an agent observes the high outcome, $x_O = x_2$, he will always receive higher utility honestly reporting the high payoff outcome ($x_R = x_2$,) than dishonestly reporting the low payoff outcome ($x_R = x_1$) because underreporting would decrease his payoff and generate a lying cost. In all of the 389 treatments we have examined, in only two treatments do subjects report the low payoff outcome statistically more often than the expected frequency of the true distribution they should have observed (Utikal and Fischbacher 2013 in their experiment with a small (n=12) sample of nuns and the 1-Coin Telephone treatment of Abeler *et al.* (2014), where the winning outcome was reported less than 50% of the time). Of course, if there is on average upward lying in a treatment this does not preclude the presence of downward lying. However the evidence suggests downward lying is rare; for example in their Observed treatments, where subjects can be individually observed lying or not, Gneezy *et al.* (2018) identified only one subject out of 602 who underreports. See further evidence of the rarity of downward lying in footnote 13. Thus, even if we acknowledge that in principle downward lying by agents observing the high payoff outcome could occur because of reputation concerns, in our setting we do not theoretically examine the possibility of lying in the case where subjects observe the higher payoff outcome. Abeler *et al.* (2018) list the theoretical models able to rationalize downward lying, although their model of loss aversion similarly predicts no downward lying (see their Appendix B2).

$$U''(D) = (x_2-x_1)^2[g''(p(x_2-x_1)) + h''((1-p)(x_2-x_1))] \quad (8)$$

To examine the effect of a change in p on reporting, we isolate the distinct forces driving their overall comparative static effects. First, we isolate the effect of loss aversion (*i.e.*, the effect that for equivalently sized monetary losses and gains, losses have a larger effect on utility than gains) by assuming linear marginal utility over gains and losses. Second, we isolate the reference-dependent effect of diminishing marginal utility of losses and gains by assuming the utility of losses and gains are identical (*i.e.*, $g(x) = h(x)$). After examining these two cases, we then examine the overall impact, assuming both loss aversion and reference dependence.

Loss Aversion without reference-dependent diminishing marginal utility:

To isolate the loss aversion effect, we assume linear utility:

$$g(x) = kx \ (k>0)$$
$$h(x) = (1+a)g(x) \text{ and } a > 0 \quad (9)$$

$g(x)$ and $h(x)$ are linear utility functions (*i.e.*, no diminishing marginal utility) and with $a > 0$ $h(x)$ is steeper than $g(x)$, reflecting loss aversion. The marginal utility of reporting dishonestly is:

$$U'(D) = k(x_2-x_1) - (1+a)k(x_2-x_1) = -ak(x_2-x_1) < 0 \quad (7')$$

Intuitively, as the probability of observing the bad payoff outcome increases, the increase in utility in the gain domain, $k(x_2-x_1)$, is always less than the decrease in utility in the loss domain, $(1+a)k(x_2-x_1)$. This follows immediately from the steeper slope of losses than gains assumed with loss-averse preferences. Thus, as the probability of observing the bad payoff outcome increases, the net gain in utility from the monetary payment falls, due to loss aversion; if it falls below the moral cost of lying, then the agent will have overall greater utility from reporting honestly. This gives our first theoretical result concerning loss aversion on reporting dishonestly: as the probability of observing the bad payoff outcome decreases, increasing utility will be received from reporting dishonestly.

Reference-dependence and diminishing marginal utility without loss aversion

To isolate the reference-dependent diminishing marginal utility effect, we assume that (a) the utility of losses and gains are identical for the same absolute size of the gain and loss, and (b) the utility of gains and losses exhibit diminishing marginal utility the further payoffs are from the reference point:

$$g(x) = h(x),\ g'(x) > 0,\ g''(x) < 0 \tag{10}$$

In this case, we have:

$$U'(D) = (x_2\text{-}x_1)[g'(p(x_2\text{-}x_1)) - g'((1\text{-}p)(x_2\text{-}x_1))], \text{ and} \tag{7''}$$

$$U''(D) = (x_2\text{-}x_1)2[g''(p(x_2\text{-}x_1)) + g''((1\text{-}p)(x_2\text{-}x_1))] \tag{8'}$$

Given diminishing marginal utility, $g''(x) < 0$, it immediately follows that $U''(D) < 0$ for all p. Further, equation (7'') indicates that if $p = 0.5$ then $U'(D) = (x_2\text{-}x_1)[g'(.5(x_2\text{-}x_1)) - g'(.5(x_2\text{-}x_1))] = 0$. Thus, given $U''(D) < 0$ for all p and $U'(D) = 0$ when $p = 0.5$, it immediately follows that:

$$\begin{aligned} &\text{For } p < 0.5\text{: } U'(D) > 0;\ U''(D) < 0 \\ &\text{For } p = 0.5\text{: } U'(D) = 0;\ U''(D) < 0 \\ &\text{For } p > 0.5\text{: } U'(D) < 0;\ U''(D) < 0 \end{aligned} \tag{11}$$

Intuitively, when p is less than 0.5, the reference point x_e, $x_e = px_1 + (1\text{-}p)x_2$, is closer to the higher payoff x_2 than the lower payoff x_1, *i.e.*, $|x_2\text{-}x_e| = p(x_2\text{-}x_1) < |x_1\text{-}x_e| = (1\text{-}p)(x_2\text{-}x_1)$, and thus we are on the steeper part of the utility function in the domain of gains than in the domain of losses due to the assumed diminishing marginal utility; thus, as p increases, more utility is gained in the gain domain than is lost in the loss domain, and thus $U'(D) > 0$. In contrast, when p is greater than 0.5, the reference point x_e is now closer to the lower payoff x_1 than the higher payoff x_2, *i.e.*, $|x_1\text{-}x_e| = (1\text{-}p)(x_2\text{-}x_1) < |x_2\text{-}x_e| = p(x_2\text{-}x_1)$, and thus we are on the steeper part of the utility function in the domain of losses rather than in the domain of gains; as p increases, more utility is gained in the loss domain than is lost in the gain domain, and thus $U'(D) < 0$. This provides our second theoretical result concerning reference-dependence and diminishing marginal utility: as the probability p of observing

the bad payoff outcome decreases, increasingly more utility is received from reporting dishonestly when p is greater than 0.5, and less utility is received from reporting dishonestly when p is less than 0.5.

General case: Reference-dependence with Loss Aversion

Both reference dependence and loss aversion predict that as the probability p of observing the bad outcome decreases, increasingly more utility is received from reporting dishonestly when p is greater than 0.5. However, when p is less than 0.5, loss aversion continues to predict that as p decreases, utility increases when reporting dishonestly; in contrast, reference dependence predicts that utility is decreasing. Given that loss aversion predicts a linearly decreasing effect of p on the utility of misreporting (equation 7') whereas reference dependence predicts a diminishing positive effect (equation 7''), and given U'(D) = 0 at $p = 0.5$, continuity implies that there exists a p^* ($0 \leq p^* < 0.5$) such that for all $p > p^*$ the loss aversion effect is greater than the reference dependence effect.

To explore where the loss aversion effect is greater than the reference-dependent diminishing marginal utility effect, p^*, we consider the functional form of utility where agents have CRRA preferences, *i.e.*, $g(x) = x(1\text{-}r)/(1\text{-}r)$ where r is the coefficient of relative risk aversion. To address loss aversion, we assume $h(x) = (1+\alpha)g(x)$, with $\alpha > 0$. With these preferences, it is easy to show that $p^* = 1 / [(e^{ln(\alpha)}/r)+1]$. This result indicates that as the degree of loss aversion α increases, *ceteris paribus* p^* decreases and the range in which the loss aversion effect is greater than the reference dependent effect increases. Moreover, as the relative degree of risk aversion r increases (*i.e.*, agents have greater diminishing marginal utility), *ceteris paribus* p^* increases and the range in which the reference dependent effect is greater than the loss aversion effect increases.

Figure 1 provides a graphical example of the difference in utility from reporting dishonestly *vs*. honestly, U(D), for an agent with CRRA utility. We assume that losses are twice as steep as gains (α=2) as is commonly estimated (*e.g.*, Kahneman and Tversky 1992; Abdellaoui *et al*. 2007). We

consider $r = 0.3$, $r = 0.5$ and $r = 0.7$, within the range commonly estimated in the literature (*e.g.*, Holt and Laury 2002; Eckel and Grossman 2008). Figure 1 shows that p^* is greater for agents with greater relative risk aversion; for example, $p^* = 0.09$, 0.20 and 0.27 for $r = 0.3$, 0.5 and 0.7 respectively. Consistent with Results 1 and 2 (and Result 3 below), Figure 1 also shows that as p decreases from 1 to p^* the agent receives increasingly more utility from reporting dishonestly than honestly, and that as p further decreases below p^* the relative benefits of reporting dishonestly start to decrease. Finally, if we assume the CRRA functional form of utility and standard values for relative risk aversion and loss aversion, then the loss aversion effect dominates the diminishing marginal utility effect for values of $p > p^*$ in the range of $0.09 \leq p^* \leq 0.25$.

The combined theoretical prediction for agents that have reference-dependent loss-averse preferences with diminishing marginal utility gives our last result concerning reference-dependence and loss-aversion with diminishing marginal utility: as the probability p of observing the low payoff outcome decreases, increasingly more utility is received from reporting dishonestly when p is greater than p^*. Moreover, p^* decreases the more loss averse and the less risk averse the agent is.

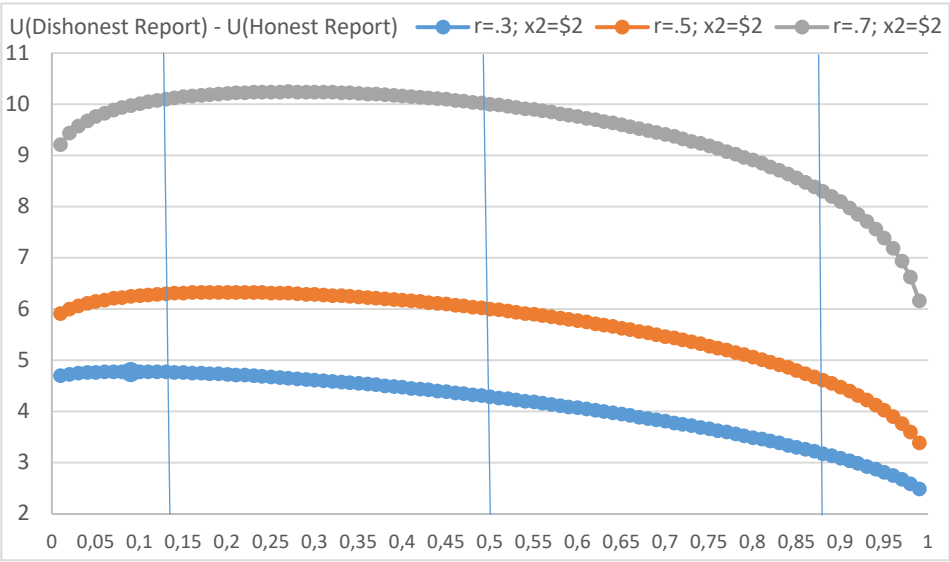


Figure 1: Utility of reporting dishonestly vs. honestly, U(D), as a function of the probability of the low payoff outcome

Notes: Utility functions of losses and gains assume CRRA utility with losses twice as steep as gains, assuming the low payoff is \$0 (x_1=\$0) and the high payoff is \$2 (x_2=\$2). Vertical lines represent the values of p (0.125, 0.5, 0.875) used in the experiment reported in section 5.

Note that while our theoretical model focuses on the importance of loss aversion on lying, it can be extended to account for reputation concerns (see a first attempt in this direction in Appendix 1).

3. Re-examination of the existing literature

To test the hypothesis of a negative correlation between the likelihood of a low payoff outcome and the percent of dishonest reports, we first apply the technique developed in Garbarino *et al*. (2018) to estimate the mean percent of individuals reporting dishonestly across previous studies. The idea is that the expected percent of subjects who lied is the sum (over all possible observations given a given number T of subjects reported the high outcome) of the percent of subjects who were dishonest weighted by the probability that T subjects observed the low outcome. Using the full distribution of possible outcomes, we are able to determine (a) the mean expected percent of subjects who were dishonest conditional on having received the low payoff outcome, (b) the minimum and maximum percent that were dishonest (the lower and upper bounds on the CI), and (c) using the CDFs, whether two (or more) treatments significantly differ for any statistic of interest over the full distribution, using non-parametric tests. To estimate mean lying rates we assume that any subject who reports the low outcome x_1 reports truthfully.[10]

To create our dataset, we started from the studies identified in Abeler *et al.*'s (2018) meta-analysis. These studies included those in which subjects have to report the private realization of a

[10] Rigdon and d'Esterre (2015) have unpublished data in which they were able to observe the true outcome subjects observed (unbeknownst to the subjects) along with the reported performance in a task. In personal correspondence (March 2017), the authors informed us that zero subjects lied to report the lower monetary outcome. Kröll and Rustagi (2016) asked 72 milkmen to roll a die 40 times using a Bluetooth enabled die that transmits the actual outcome of each roll to the smartphone of the experimenter, with subjects unaware of this. In personal communication (June 2017), they informed us that underreporting, saying you got the low outcome when you actually got the high outcome, occurred in only 9 cases out of 2880 (0.3% of all observations); and in 8 out of the 9 cases, under-reporting was by 1 point and in the ninth case it was by 2 points.

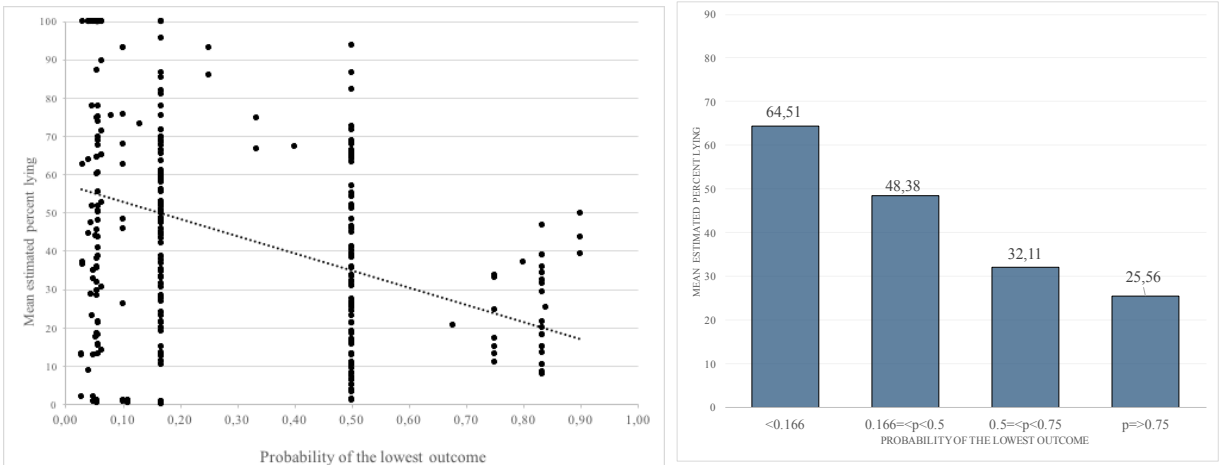
random device (a die, a coin, a ticket or a card) that is unobservable to the experimenter, and without any risk of detection during or after the experiment. This includes 'mind games' in which subjects report the number of times they were able to correctly predict (to themselves) the outcome of a random device. We exclude studies in which the subject's payoff depends on strategic interactions with other players (but we retain those in which the report affects other people), and treatments in which there is no variation in the monetary incentives associated with all possible realizations of the random device.

In total, our analysis includes 81 studies and 389 treatments, involving 36,668 individuals from 44 countries and representing 257,682 decisions (see Appendix Table A1). Seventy-one percent of these studies have been conducted by a team involving at least one economist; 53% of the treatments have been run in a standard laboratory setting with standard student-subjects pools; 13% are from artefactual field experiments conducted as in a lab setting but involving atypical subjects.

When there are only two possible outcomes (for example in a single coin tossing task or when there are only two payoffs for all possible die rolls) we calculate the percent of subjects who lie to get the high payoff when they observed a low outcome. This calculation requires only knowing the total number of subjects, the number of subjects reporting the high outcome, and the probability of getting the low outcome. When there are more than two possible outcomes, we estimate two statistics: first, the percent of people who lied to avoid the lowest payoff, and second, the percent who lied to get the highest payoff, *i.e.* framing all payoffs, except the highest one, as bad outcomes. In the case of repeated decisions (up to 20 coin tosses or card draws and up to 75 die rolls), we define the lowest and the highest outcomes as those associated with the bottom 5% and the top 5% outcomes of the theoretical distribution, respectively.[11]

[11] For example, when subjects have to predict the outcome of 20 coin tosses and earn money that monotonically increases with the number of correct predictions, the bad (good) outcome is defined as getting at most 6 (at least 14, resp.) correct guesses (with p=0.057). The 5% threshold is arbitrary, but it allows us to avoid studying extremely rare events such as 2^{-20} in the case of predicting all 20 coin tosses incorrectly or $1-2^{-20}$ in the case of predicting all 20 coin tosses correctly.

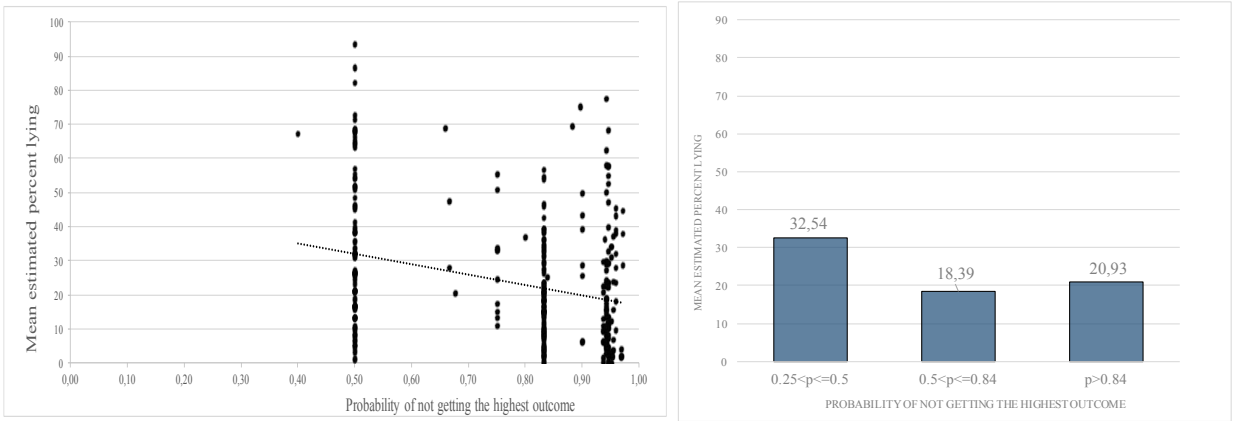
Figure 2 displays our estimated mean percent of subjects lying to avoid the lowest payoff in the reviewed studies as a function of the likelihood of the lowest outcome. In panel A, the *x*-axis represents the probability of the lowest outcome and the *y*-axis represents the mean percent lying. Each treatment in every study gives one observation. Panel B aggregates this information for various probability intervals. In Figure 3, we consider instead the relationship between the estimated mean percent of subjects lying to avoid the low payoff, when the latter is defined as any payoff except the highest one, and the likelihood of not getting the highest outcome.



A B

Figure 2: Mean estimated percent of subjects lying to avoid the lowest payoff by the likelihood of the lowest outcome, literature review

Note: In Panel B, the bar with $p < 0.166$ corresponds to 139 treatments and 11,270 subjects, the bar with $0.166 \leq p < 0.5$ corresponds to 102 treatments and 12,129 subjects, the bar with $0.5 \leq p < 0.75$ corresponds to 107 treatments and 7,399 subjects, the bar with $p \geq 0.75$ corresponds to 37 treatments and 4,502 subjects. In total, we have 385 treatments instead of 389 in this figure because in one study information was not available at the treatment level.



A B

Figure 3: Mean estimated percent of subjects lying to avoid a low payoff by the likelihood of not getting the highest outcome, literature review

Note: In Panel B, the bar with $0.25 < p \leq 0.5$ corresponds to 107 treatments and 7,395 subjects, the bar with $0.5 < p \leq 0.84$ to 141 treatments and 15,656 subjects, the bar with $p > 0.84$ to 141 treatments and 13,617 subjects.

These figures show a clear negative relationship between the estimated mean lying rate and the likelihood of the bad outcome, regardless of whether it is defined as the lowest outcome or any outcome except the highest.

Due to the large variance in lying across studies for a given probability, we have further analyzed the data using linear regressions in which the dependent variable is the mean estimated lying rate with each treatment serving as one observation. Robust standard errors are clustered at the study level and sample weighting allows us to control for the large diversity of sample sizes across studies.[12] In models (1) to (3) in Table 1, the bad outcome is defined as the lowest payoff outcome, whereas in models (4) to (6) it is defined as any outcome except the highest.

Our main independent variable of interest is the probability of the bad outcome that aims at capturing the effect of loss aversion. However, one potential additional explanation for the effect of the probability of the bad outcome on lying, as mentioned earlier, could be reputation concerns: subjects may feel more worried about appearing to have lied if they report the higher outcome when the probability of the bad outcome is higher. To test for reputation effects, we examine the interaction of p and the degree of observability of the outcome by others (Appendix 1 presents the extension of the theoretical model for various levels of observability). If reputation is driving the relationship between p and dishonest reports, then we would anticipate the effect of p varying with observability; seeing the smallest (or no) effect when there is no or low observability and the largest effect when observability is high. We define three categories of observability of subjects' behavior by others in the experiments: "low observability" (our reference category) characterizes treatments

[12] Note that we also estimated regressions with bootstrapping (allowing for up to 500 replications), instead of introducing weights. The results remain qualitatively similar.

in which there is no direct interaction between the subject and the experimenter (studies conducted online, or in a laboratory setting but with a double-blind procedure or using mind games without verbal reporting);[13] "high observability" characterizes treatments in which the subjects are observed by another subject when rolling the die or tossing the coin, or where they have to verbally report the outcome to an experimenter (either in the laboratory or in the field, face-to-face or on the phone); "moderate observability" characterizes the intermediate cases (studies conducted in the field or in the laboratory without double anonymity and without verbal reporting). 42% of the treatments (162) have low observability, 49% (193) have moderate observability and 9% (34) have high observability.

Models (2) and (5) augment models (1) and (4) by including the expected values of the lowest/bad outcome and of the good/highest outcome, converted from local currencies to US Dollars by 2015 Purchasing Power Parities.[14] They also include a control for conducting the experiment in a standard university lab setting and a dummy variable indicating whether the study has been conducted by a team involving economists, as opposed to studies conducted by other social scientists. This dummy variable assesses whether subjects may be more reluctant to lie in experiments conducted by economists because of the lack of deception in economic experiments; alternatively, they may be more likely to lie if they believe that economists expect payoff maximization behavior from the subjects. Finally, in models (3) and (6) we add country dummies.

Table 1. Likelihood of the bad outcome as a determinant of the mean lying rate

Dep. Variable:	Lying to avoid the lowest outcome			Lying to avoid a bad outcome		
Mean lying rate	(1)	(2)	(3)	(4)	(5)	(6)
p	-0.488***	-0.524***	-0.499***	-0.418**	-0.425**	-0.434***
	(0.092)	(0.089)	(0.075)	(0.166)	(0.189)	(0.152)
Moderate observability	-0.017	-0.136**	-0.177***	0.352	0.202	-0.115
	(0.073)	(0.067)	(0.060)	(0.246)	(0.226)	(0.160)

[13] We consider mind games as having "low observability" because it is clear to the subjects that a lie cannot be detected by the experimenter who cannot observe the actual guess, whereas with a die-rolling task some subjects may fear scrutiny (presence of hidden cameras, or use of special devices connected to the experimenter's cell phone).
[14] Suppose a die task in which each dot pays $1, except when the subject rolls a 6. Conditional on getting the lowest outcome, the expected value of this lowest outcome is $0, while the expected value of a good outcome is $3. Conditional on getting the highest outcome, the expected value of the highest outcome is $5, while the expected value of a bad outcome is $2.

High observability	-0.299**	-0.402***	-0.392***	-0.161*	-0.235**	-0.187**
	(0.125)	(0.112)	(0.112)	(0.092)	(0.090)	(0.077)
*p** Moderate observability	0.161	0.186	0.116	-0.420	-0.361	0.017
	(0.169)	(0.133)	(0.103)	(0.303)	(0.283)	(0.193)
*p** High observability	0.084	0.265	0.287	-0.136	0.002	-0.040
	(0.334)	(0.282)	(0.305)	(0.276)	(0.235)	(0.236)
EV of the lowest outcome	-	0.007	0.005	-	-	-
		(0.006)	(0.004)			
EV of a good outcome	-	-0.001	-0.001	-	-	-
		(0.001)	(0.001)			
EV of a bad outcome	-	-	-	-	0.010**	0.011**
					(0.004)	(0.005)
EV of the highest outcome	-	-	-	-	-0.003**	-0.003**
					(0.001)	(0.001)
Laboratory	-	0.136***	0.175***	-	0.126***	0.132***
		(0.50)	(0.52)		(0.036)	(0.033)
Economics affiliation	-	0.102**	0.033	-	0.070	0.027
		(0.041)	(0.041)		(0.050)	(0.029)
Country dummies	No	No	Yes	No	No	Yes
Constant	0.647***	0.577***	0.659***	0.593***	0.546***	0.565***
	(0.053)	(0.055)	(0.082)	(0.112)	(0.138)	(0.132)
Nb observations	385	385	385	389	389	389
Nb clusters	81	81	81	81	81	81
R^2	0.242	0.309	0.492	0.298	0.386	0.590

Note: Table 1 reports OLS estimates with robust standard errors (in parentheses) clustered at the study level and sample weighted. Each treatment gives one observation. 4 observations are missing in models (1) to (3) because information on lying to avoid the lowest outcome was not available at the treatment level in one study. *p* is the probability of the lowest outcome in models (1) to (3) and the probability of a bad outcome in models (4) to (6). ***, **, and * indicate significance at the 0.01, 0.05 and 0.10 level, respectively.

Consistent with Figures 2 and 3, all specifications reported in Table 1 show that the mean lying rate increases significantly (most of the time at the 1% level) when the likelihood of the bad outcome decreases. This effect is robust to either measure of lying (to avoid the lowest payoff or to get the highest payoff). The magnitude of the estimated effect is substantial; for instance, if we compare a die roll with a 5/6 chance to observe the bad outcome to a coin toss with a 50% chance to observe the bad outcome, *ceteris paribus*, the estimate with all the controls (Model 6) predicts 14.47 percentage points ([(5/6)-(1/2)]*0.434) more subjects will lie in the coin toss condition. We re-estimated the six models in Table 1 adding the square of the probability of the lowest (or bad) outcome to test for evidence of the reference dependence diminishing marginal utility implication that an increase in the probability could have a positive effect for low values of *p* (see Figure 1). However, the squared term is never significant. This suggests that either *p** is close to zero, we do not have enough observations

with low enough values of p and thus not enough power to detect this effect, or the loss aversion effect dominates the reference dependence effect for the lowest values of p we observe.

Table 1 also indicates a main effect of observability/reputation: moderate and high observability leads to less lying than low observability, consistent with the concern for appearing dishonest being greater when the feeling of scrutiny may be higher. However, Table 1 shows no significant interaction effects between observability and the probability of the lowest or bad outcome in any model. This suggests that reputation alone cannot explain the relationship between the probability of the bad outcome and lying. To investigate this question further, we have estimated the same six models for each category of observability separately. Table A2 in Appendix 1 reports the 18 estimated coefficients of the probability of the lowest outcome in models (1) to (3) and of the probability of the bad outcome in models (4) to (6). The probability is negative and significant at the 5% or 1% level in almost all regressions when observability is low or moderate. When observability is high the probability is no longer significant but the model lacks statistical power to detect any effect (there are only 34 observations and little variation in the probability since for 19 of these observations -56%- it is equal to 0.5). If reputation was the only effect driving the relationship between the probability and lying, we would expect to see the smallest coefficient of the probability when there is low observability and the largest coefficient when observability is high. However, the ordering of the coefficient in Table A2 points to the opposite direction, as it appears to decrease as observability increases. This shows again that reputation alone is unlikely to drive the effect of the probability of the low payoff outcome on lying, while the results are consistent with our model of loss aversion.

The regressions in Table 1 also show that the expected values have an impact consistent with expected utility maximizers with increasing marginal utility over money; in particular, as the expected value of the bad payoff increases (so the marginal utility of an extra dollar if lying is

smaller), fewer subjects lied; and as the expected payoff of the highest outcome increases (so there is a greater increase in payoffs to lie), more subjects lied. These results are directional in the analysis of lying to avoid the lowest payoff, and significant in the analysis of lying to get the highest payoff. However, when we replace the expected values with alternative definitions of the payoffs, we do not find significant effects, thus in general we find little evidence the size of the monetary stakes had a robust effect on lying, consistent with Abeler *et al*. (2018).[15]

Overall, the findings from our analysis of the literature are consistent with our theoretical result about the role of loss aversion in lying. Our second approach to isolate the effect of loss aversion is to run a new experiment where we minimize reputation effects and test the role of changing probabilities directly.

4. A within-study test of loss aversion and lying

In this section, we test whether the relationship between the probability of observing the bad outcome and the percent of subjects lying holds when we manipulate the probability of the outcomes, remove the variations in conditions that exist between studies in the literature (*e.g*., different recruitment, subjects, tasks and payoffs), and minimize potential reputation concerns. To accomplish this we use an online experiment that varies the probability distribution of outcomes, *ceteris paribus*, to earn a fixed payoff.

Experimental design and procedures[16]

[15] In alternative specifications (available upon request), we replaced the expected values by the maximum and the minimum payoffs. These variables were never significant, except the lowest payoff in model (6), with a positive sign. We also replaced the expected maximum payoff with the difference in the expected payoffs from reporting the maximum and minimum payoffs. Again, these variables were never significant. Since the R^2 is not higher in these alternative specifications, we omit reporting them.

[16] The full instructions for the experiment are available in the online Appendix. Web link accessible as of 11/08/2018 https://econusyd.az1.qualtrics.com/jfe/form/SV_bI9ZO3D0l3uo61v

This experiment is based on a mind coin tossing game.[17] Subjects have the opportunity to earn money for themselves depending on their ability to predict the outcome of three coin tosses. They are instructed to toss a coin three times, but before each toss they have to predict which side of the coin will be face-up. They have to report only the number of correct guesses. Thus, they can inflate their actual success by misreporting the number of correct guesses. The major advantage of the mind game process is that the subjects know that their outcome cannot be scrutinized: the experimenter cannot observe the prediction or the actual outcome of the coin toss. To minimize reputation concerns, the experiment was also run online and double-anonymous; it was common knowledge that the experimenter never knew the identity of the subjects. Therefore, individuals should have no reason to disguise a lie to maintain a positive image with other subjects or the experimenters (of course, intrinsic moral cost concerns may still exist).

The study involves three between-subject treatments that vary the probability of observing the lowest and the highest payoff. In the "0002" treatment, subjects earn \$0 if they report 0, 1, or 2 correct guesses and they earn \$2 if they report 3 correct guesses. The probability of observing the low payoff is 87.50% and the probability of observing the high payoff is 12.50%. In the "0022" treatment, subjects earn \$0 if they report 0 or 1 correct guess and they earn \$2 if they report 2 or 3 correct guesses. Here, the probability of observing the low payoff or the high payoff is 50%. Finally, in the "0222" treatment, subjects earn \$0 if they report 0 correct guesses and they earn \$2 if they report at least one correct guess. Here, the probability of observing the low payoff is 12.50% and the probability of observing the high payoff is 87.50%. Procedures and payoffs are common knowledge.

The experiment was run using Amazon MTurk with 978 U.S. subjects.[18] We varied the number of subjects across treatments based on a preliminary statistical power analysis, with the objective of

[17] Mind games have been used mainly with a die (Jiang 2013; Shalvi and de Dreu 2014; Barfort *et al*. 2015; Potters and Stoop 2016; Kajackaite and Gneezy 2017), rarely with coin tosses (Shalvi 2012).
[18] The use of MTurk ensures a double anonymous procedure reinforcing the lack of observation by others. It enables the

getting approximately the same number of likely opportunities to get the $0 payoff in each treatment, and hence the same number of respondents with a financial motive to lie. 80 persons participated in the 0002 treatment, 327 in the 0022 treatment and 571 in the 0222 treatment. After the task was explained (see the online instructions for full details), they tossed a coin and reported the number of correct predictions. We also collected a number of socio-demographic characteristics (including age, gender, highest educational attainment, category of household annual pretax income, mean weekly expenditures).[19] A summary of our subjects' characteristics by treatment and the effects of these characteristics on payoffs can be found in Appendix Tables A2 and A3.[20] On average, the experiment lasted 7 minutes 45 seconds (S.D.= 8 minutes and 24 seconds) and subjects earned $3.27 (S.D.= 0.64) which included a $1.50 participation fee.

Results

Our model predicts that subjects will be more likely to lie to avoid earning $0 in the 0222 treatment than in the 0022, and more likely to lie in the 0022 treatment than in the 0002 treatment. The data confirm these two predictions: consistent with the loss aversion model presented above, individuals lie more on average to increase their own payoff when observing the low payoff is more unlikely.

Table 2 displays summary statistics by treatment, including the distributions of reported outcomes, the *p*-values of χ^2 goodness-of-fit tests comparing the empirical distributions of reports

collection of a large number of observations from a broad demographic pool. Given the simplicity of the task, we feel confident that subjects understood the instructions. Moreover, studies have shown that the quality of data and behavior on MTurk do not differ from that of other sources (*e.g*., Paolacci *et al*. 2010; Horton *et al*. 2011; Suri and Watts 2011).

[19] We also collected beliefs about the social appropriateness of reporting different outcome combinations using a procedure inspired from Krupka and Weber (2013), and we elicited the person's belief in good luck using a twelve-item questionnaire from psychology (Darke and Freedman 1997). Since neither of these potential moderators had substantial or consistent effect, they have been removed from the remaining analyses.

[20] The various demographics are randomly dispersed across conditions. Pairwise tests show no significant differences between conditions on gender ($p>0.14$), age ($p>0.14$), education ($p>0.90$), income ($p>0.07$), or weekly spending ($p>0.80$). We cannot estimate lying at the individual level, so to test the effects of covariates we used the participant's payout as a proxy for lying, since those who lied receive the larger payout. The demographic and attitudinal variables have little effect on the participant's payout. Only age consistently reaches the 0.05 level of significance, such that younger people get higher payouts. Additional analyses find no significant interactions of the covariates with the conditions (all *p*-values > 0.05), demonstrating that none of the covariates moderate the effects of treatment on payout.

with the theoretical distribution, the *p*-values of binomial tests comparing each category of reported outcome with its theoretical relative frequency, the mean estimated lying rates using our technique, and the associated 95% confidence intervals.

Table 2 shows that the percentage of reports paying \$0 is 2.98% in the 0222 treatment, 16.82% in the 0022 treatment and 48.75% in the 0002 treatment, while the theoretical percentages should be 12.50%, 50% and 87.50%, respectively. Binomial tests and χ^2 goodness-of-fit tests show that the reports deviate significantly from the theoretical distribution. Our estimation based on the full distribution of reports (assuming that nobody lies downwards) shows that on average 44.18% of subjects who had a financial motive to lie lied in the 0002 treatment, 66.26% in the 0022 treatment, and 75.88% in the 0222 treatment. Pairwise comparisons indicate that the differences between the three treatments are significant at the $p < 0.01$ level, as there is no overlap between the 95% confidence intervals. Figure 4 illustrates these findings.

Table 2. Summary statistics

Treatments	0222	0022	0002
Number of reported correct guesses			
0	17 (2.98%)	13 (3.98%)	3 (3.75%)
1	165 (28.90%)	42 (12.84%)	11 (13.75%)
2	308 (53.94%)	221 (67.58%)	25 (31.25%)
3	81 (14.19%)	51 (15.60%)	41 (51.25%)
N	571 (100%)	327 (100%)	80 (100%)
p-values from χ^2 goodness-of-fit tests	<0.001	<0.001	<0.001
p-values from two-sided Binomial tests			
0 correct guess	<0.001	<0.001	0.016
1 correct guess	<0.001	<0.001	<0.001
2 correct guesses	<0.001	<0.001	0.299
3 correct guesses	0.229	0.094	<0.001
Percent who earned \$0	2.98%	16.82%	48.75%
Expected percent if no lying	12.50%	50.00%	87.50%
Mean percent of subjects lying	75.88	66.26	44.18
Confidence Interval, 95%	70.03-80.62	62.40-69.73	39.39-48.63

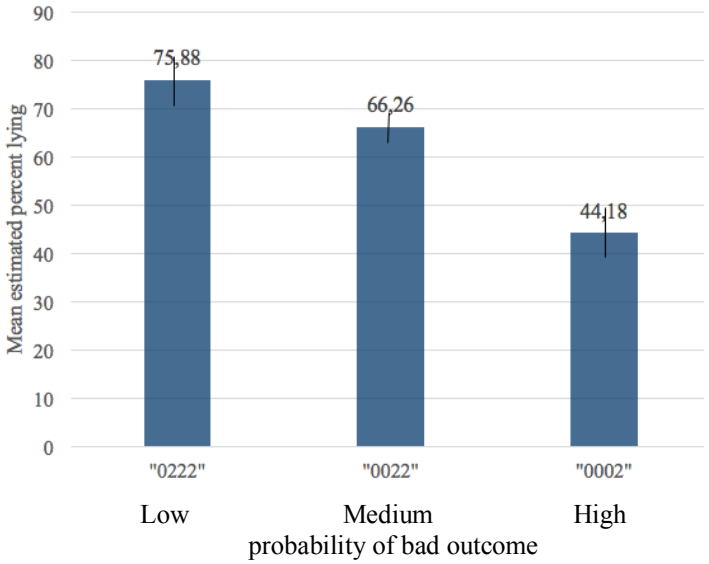


Figure 4. Mean estimated percent of subjects lying, by treatment

Discussion and Conclusion

Loss aversion predicts that the probability of observing each payoff outcome affects the level of dishonesty. We demonstrate theoretically that as the likelihood of observing an *ex-ante* low payoff outcome decreases, the gain in utility for a loss-averse agent who reports dishonestly to avoid the loss increases. Using a simple econometric technique to estimate the full distribution of dishonesty, we analyze data from the extant literature and a new experiment. We confirm that individuals lie more to avoid the undesired outcome when this undesired outcome is more unlikely to occur.

Several extensions of our study might further extend our understanding. To further test the theoretical model, new experiments could measure each subject's degree of loss and risk aversion to more directly investigate the relationship between loss aversion and lying.[21] Moreover, the role of loss aversion does not exclude the influence of other factors that may motivate people to lie more

[21] We did not measure risk and loss aversion because, first, the length of our study was already approaching the average time for running MTurk studies, and to appropriately estimate loss aversion would go beyond the upper limits of an MTurk study. Second, we estimate lying at the treatment level, thus we would need many more observations per subject to estimate lying at the individual level. Third, we were concerned that measuring loss or risk aversion before our lying task might affect behavior and that performing the lying task first might affect the measure of risk and loss aversion.

when a bad payoff outcome is more unlikely. Recent models consider how decision makers derive disutility based on the degree to which they are perceived to be cheaters by others and in their own eyes. Some participants might have a distaste for being perceived as acting suspicious. However, when examining the extant literature, we found no interaction effect between the degree of observability of subjects and the impact of the probability of the bad outcome. Even in studies where observability is very limited or nonexistent, lying is less likely when the bad outcome is more likely. Being conducted online and double-anonymous, our own experiment excludes any direct contact with, and any possibility of being observed by, the experimenter to minimize the salience that could possibly generate suspicion or risk of negative reputation if lying. Moreover, using a mind game guarantees that a lie cannot be detected at the individual level. We still find a strong relationship between the probability of the bad outcome and the mean lying rate. This finding supports our model but we acknowledge that it does not preclude that people who are more concerned about their moral identity may be more reluctant to lie when the probability of the bad outcome is higher, even in the absence of external observers.

Moreover, as developed in Abeler *et al*. (2018), other approaches can also contribute to explain this lying behavior, for example descriptive norms and conformity in lying costs or in reports, or distributional preferences (although in the experiments considered here there is no active counterpart). Indeed, when the probability of the high payoff outcome is higher, a subject that gets the low payoff outcome may be more willing to lie than when this probability is lower despite his lying cost because he can persuade himself that everyone in the same situation would do the same or because he is suffering from disadvantageous inequality relative to a higher number of individuals. Our model does not include beliefs about what the norm in the community is and it assumes homogeneous preferences of the agents over payments; this could be relaxed in further extensions.

Our study has a number of implications, both for academics and for policy-makers. Regarding academic implications, our analysis reveals the importance of controlling for the probability of low payoff outcomes when studying lying behavior. It also suggests that if experimentalists want to study situations where dishonesty might not otherwise be too high, they should use a large sample and set a low probability for the bad outcome. Regarding policy implications, our results suggest that to limit lying, choices should be framed to increase the perceived likelihood of the undesired outcome. They also suggest that costly efforts of detection should be concentrated on situations in which the likelihood of a bad outcome is relatively low.

References

Abdellaoui, Mohammed, Han Bleichrodt, and Corina Paraschiv (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science*, 53: 1659-1674.

Abeler, Johannes, Anke Becker, and Armin Falk (2014). Representative evidence on lying costs. *Journal of Public Economics*, 113: 96-104.

Abeler, Johannes (2015). A reporting experiment with Chinese professionals. Mimeo.

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (Forthcoming). Preferences for truth-telling. *Econometrica*.

Amir, Amitai, Tehila Kogut, and Yoella Bereby-Meyer (2016). Careful Cheating: People Cheat Groups Rather than Individuals. *Frontiers in Psychology*, 7, article 371.

Antony, Markus, Armin Falk, and Holger Gerhardt (2016). The Impact of Food and Water Deprivation on Economic Decision-Making. Poster, Society for Neuroeconomics 13th Annual Meeting, September, Miami.

Arbel, Yuval, Ronen Bar-El, Erez Siniver, and Yossef Tobol (2014). Roll a die and tell a lie–what affects honesty? *Journal of Economic Behavior & Organization*, 107: 153–172.

Ariely, Dan, Ximena Garcia-Rada, Lars Hornuf, and Heather Mann (2014). The (true) legacy of two really existing economic systems. University of Munich Discussion Paper, 2014-26.

Aydogan, Gökhan, Andrea Jobst, Fabian Loy, Norbert Müller, and Martin Kocher (2015). Oxytocin promotes dishonesty under competition. University of Munich.

Balasubramanian, Parasuram, Victor M. Bennett, and Lamar Pierce (2017). The wages of dishonesty: The supply of cheating under high-powered incentives. *Journal of Economic Behavior & Organization,* 137: 428-444.

Banerjee, Ritwick, Nabanita Datta Gupta, and Marie Claire Villeval (2017). The spillover effects of Affirmative Action on confidence and unethical behavior. *European Economic Review* 101: 567-604.

Bar-El, Ronen, and Yossef Tobol (2017). Honesty toward the Holy Day. *Journal of Behavioral and Experimental Economics* 68.

Barfort, Sebastian, Nikolai Harmon, Frederik Hjorth, and Asmus L. Olsen (2015). Dishonesty and Selection into Public Service in Denmark: Who Runs the World's Least Corrupt Public Sector? University of Copenhagen Discussion Paper.

Beck, Thorsten, Christopher Bühren, B. Frank, and Elina Khachatryan (2016). Lying in the face of monitoring, reciprocity, and commitment. Mimeo.

Becker, Gary S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy,* 76: 169–217.

Blanco, Mariana, and Juan Camilo Cárdenas (2015). Honesty after a labor relationship. Universidad del Rosario Discussion Paper, 2015-37.

Braun, Susanne, and Lars Hornuf (2015). Leadership and persistency in spontaneous dishonesty. IAAEU Discussion Paper, 10/2015.

Bryan, Christopher J., Gabrielle S. Adams, and Benoît Monin (2013). When Cheating Would Make You a Cheater: Implicating the Self Prevents Unethical Behavior. *Journal of Experimental Psychology*, 142 (4): 1001-1005.

Bucciol, Alessandro, and Marco Piovesan (2011). Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology*, 32 (1): 73–78.

Cadsby, C. Bram, Ninghua Du, and Fei Song (2016). In-group favoritism and moral decision-making. *Journal of Economic Behavior & Organization*, 128: 59-71.

Cameron, Jessica S., and Dale T. Miller (2009). Ethical Standards in Gain versus Loss Frames. In David de Cremer (Ed.), *Psychological Perspectives on Ethical Behavior and Decision Making*. Information Age Publishing, Charlotte, NC: 91-106.

Charness, Gary, David Masclet, and Marie Claire Villeval (2014). The dark side of competition for status. *Management Science,* 60 (1): 38-55.

Charness, Gary, Celia Blanco-Jimenez, Lara Ezquerra, and Ismael Rodriguez-Lara (Forthcoming). Cheating, incentives, and money manipulation. *Experimental Economics*.

Chytilova, Julie, and Vaclav Korbel (2014). Individual and group cheating behavior: a field experiment with adolescents. IES Working Paper, 06/2014.

Clot, Sophie, Gilles Grolleau, and Lisette Ibanez (2014). Smug Alert! Exploring self-licensing behavior in a cheating game. *Economics Letters*, 123 (2): 191–194.

Cohen, Taya R., Scott T. Wolf, Abigail T. Panter, and Chester A. Insko (2011). Introducing the GASP scale: a new measure of guilt and shame proneness. *Journal of Personality and Social Psychology*, 100 (5): 947.

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

Cohn, Alain, Michel André Maréchal, and Thomas Noll (2015). Bad Boys: How Criminal Identity Salience Affects Rule Violation. *Review of Economic Studies*, 82 (4): 1289–1308.

Cohn, Alain, and Michel André Maréchal (Forthcoming). Laboratory Measure of Cheating Predicts School Misconduct. *The Economic Journal*.

Conrads, Julian, and Sebastian Lotz (2015). The effect of communication channels on dishonest behavior. *Journal of Behavioral and Experimental Economics*, 58: 88–93.

Conrads, Julian (2014). The effect of communication channels on lying. Cologne Graduate School Working Paper, 5(6).

Conrads, Julian, Bernd Irlenbusch, Rainer Michael Rilke, Gari Walkowitz (2013). Lying and team incentives. *Journal of Economic Psychology*, 34: 1-7.

Dai, Zhixin, Fabio Galeotti, and Marie Claire Villeval (2018). Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportations. *Management Science,* 64 (3): 1081-1100.

Danilov, Anastasia, Torsten Biemann, Thom Kring, and Dirk Sliwka (2013). The dark side of team incentives: Experimental evidence on advice quality from financial service professionals. *Journal of Economic Behavior & Organization,* 93: 266-272.

Darke, Peter R., and Jonathan L. Freedman (1997). The belief in good luck scale. *Journal of Research in Personality*, 31 (4): 486-511.

Dato, Simon, and Petra Nieken (2015). Compensation and Honesty: Gender Differences in Lying. Discussion Paper Series in Economics and Management, 15-20.

Dhami, Sanjit (2017). Human ethics and virtues: Rethinking the homo-oeconomicus model. CESifo Working Paper 6836.

Dieckmann, Anja, Urs Fischbacher, Veronika Grimm, Matthias Unfried, Verena Utikal, and Lorenzo Valmasoni (2015). Trust and Beliefs Among Europeans: Cross-country Evidence on Perceptions and Behavior. IWQW Discussion Paper, 04/2015.

Diekmann, Andreas, Wojtek Przepiorka, and Heiko Rauhut (2015). Lifting the veil of ignorance: An experiment on the contagiousness of norm violations. *Rationality and Society*, 27 (3): 309–333.

di Falco, Salvatore, Brice Magdalou, David Masclet, Marie Claire Villeval, and Marc Willinger (2016). Can transparency of information reduce embezzlement? An experiment in Tanzania. IZA Discussion Paper, 9925, Bonn.

Djawadi, Behnud Mir, and René Fahr (2015). ". . . and they are really lying": Clean evidence on the pervasiveness of cheating in professional contexts from a field experiment. *Journal of Economic Psychology*, 48: 48–59.

Drupp, Morotz A., Menusch Khadjavi, and Martin F. Quaas (2016). Truth-Telling and the Regulator. Evidence from a Field Experiment with Commercial Fishermen. Mimeo.

Dufwenberg, Martin, and Martin A. Dufwenberg (2018). Lies in Disguise – A theoretical Analysis of Cheating. *Journal of Economic Theory,* 175: 248-264.

Duncan, Denvil, and Danyang Li (2018). Liar Liar: Experimental evidence of the effect of confirmation-reports on dishonesty. *Southern Economic Journal*, 84(3): 742-770.

Eckel, Catherine C., and Philip J. Grossman (2008). Men, women and risk aversion: Experimental evidence. in Charles Plott and Vernon Smith (Eds.), *Handbook of Experimental Economics Results*, New York, Elsevier, 1, chapter 113: 1061-1073.

Effron, Daniel A., Christopher J. Bryan, and J. Keith Murnighan (2015). Cheating at the End to Avoid Regret. *Journal of Personality and Social Psychology*, 109 (3): 395-414.

Falk, Armin, and James Heckman (2009). Lab Experiments are a Major Source of Knowledge in the Social Sciences. *Science*, 326 (5952): 535-538.

Faravelli, Marco, Lana Friesen, and Lata Gangadharan (2015). Selection, tournaments, and dishonesty. *Journal of Economic Behavior & Organization*, 110:160-175.

Fischbacher, Urs, and Franziska Föllmi-Heusi (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association*, 11 (3): 525–547.

Foerster, Anna, Roland Pfister, Constantin Schmidts, David Dignath, and Wilfried Kunde (2013). Honesty saves time (and justifications). *Frontiers in Psychology*, 4 (473): 1-2.

Fosgaard, Toke Reinholt, Lars Gaarn Hansen, and Marco Piovesan (2013). Separating Will from Grace: An experiment on conformity and awareness in cheating. *Journal of Economic Behavior & Organization*, 93: 279–284.

Gächter, Simon, and Jonathan F. Schulz (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531, 24 March: 496-499.

Garbarino, Ellen, Robert Slonim, and Marie Claire Villeval (Forthcoming). A Method to Estimate Mean Lying Rates and Their Full Distribution. *Journal of the Economic Science Association*.

Gneezy, Uri, Alexandre Imas, and Kristóf Madarász (2014). Conscience Accounting: Emotion Dynamics and Social Behavior. *Management Science,* 60 (11): 2645-2658.

Gneezy, Uri, Agne Kajackaite and Joel Sobel (2018). Lying Aversion and the Size of the Lie *American Economic Review*, 108 (2): 419-453.

Grolleau, Gilles, Martin G. Kocher, and Angela Sutan (2017). Cheating and loss aversion: do people lie more to avoid a loss? *Management Science,* 62 (12): 3428-3438.

Halevy, Rony, Shaul Shalvi, and Bruno Verschuere (2013). Being Honest About Dishonesty: Correlating Self-Reports and Actual Lying. *Human Communication Research*, 40 (1): 54–72.

Hanna, Rema, and Shing-Yi Wang (2017). Dishonesty and Selection into Public Service: Evidence from India. *American Economic Journal: Economic Policy*, 9 (3): 262-290.

Hilbig, Benjamin E., and Corinna M. Hessler (2013). What *lie*s beneath: How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology*, 49 (2): 263-266.

Hilbig, Benjamin E., and Ingo Zettler (2015). When the cat's away, some mice will play: A basic trait account of dishonest behavior. *Journal of Research in Personality*, 57: 72-88.

Holt, Charles A., and Susan K. Laury (2002). Risk aversion and incentive effects. *American Economic Review,* 92 (5): 1644-1655.

Horton, John J., David G. Rand, and Richard J. Zeckhauser (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics*, 14 (3): 399-425.

Houser, Daniel, John A. List, Marco Piovesan, Anya Samek, and Joachim Winter (2016). Dishonesty: From parents to children. *European Economic Review*, 82: 242–254.

Hruschka, Daniel, Charles Efferson, Ting Jiang, Ashlan Falletta-Cowden, Sveinn Sigurdsson, Rita McNamara, Madeline Sands, Shirajum Munira, Edward Slingerland, and Joseph Henrich (2014). Impartial Institutions, Pathogen Stress and the Expanding Social Network. *Human Nature*, 25 (4): 567-579.

Huber, Stefan, and Odilo W. Huber (2012). The centrality of religiosity scale (CRS). *Religions*, 3.3: 710-724.

Hugh-Jones, David (2016). Honesty, beliefs about honesty, and economic growth in 15 countries. *Journal of Economic Behavior & Organization*, 127: 99-114.

Irlenbusch, Bernd, and Marie Claire Villeval (2015). Behavioral ethics: how psychology influenced economics and how economics might inform psychology? *Current Opinions in Psychology,* 6: 87–92.

Jacob, Brian A., and Steven D. Levitt (2003). Rotten apples: an investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118: 843-877.

Jacobsen, Catrine, and Marco Piovesan (2016). Tax me if you can: An artefactual field experiment on dishonesty. *Journal of Economic Behavior & Organization*, 124: 7-14.

Jacobsen, Catrine, Toke Reinholt Fosgaard, and David Pascual-Ezama (2018). Why do we lie? A practical guide to the dishonesty literature. *Journal of Economic Surveys,* 32 (2): 357-387.

Jiang, Ting (2013). Cheating in mind games: The subtlety of rules matters. *Journal of Economic Behavior & Organization*, 93: 328–336.

Jiang, Ting (2015). Other-regarding Preferences and Other-regarding Cheating – Experimental Evidence from China, Italy, Japan and the Netherlands. Mimeo.

Kajackaite, Agne (2018). Lying about Luck versus Lying about Performance. *Journal of Economic Behavior & Organization*, 153: 194-199.

Kajackaite, Agne, and Uri Gneezy (2017). Incentives and cheating. *Games and Economic Behavior*, 102: 433-444.

Kahneman, Daniel, and Amos Tversky (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47: 263-291.

Kahneman, Daniel, and Amos Tversky (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5: 297-323.

Khalmetski, Kiryl, and Dirk Sliwka (2017). Disguising lies – Image concerns and partial lying in cheating games. CESIfo Working Paper 6347.

Kroher, Martina, and Tobias Wolbring (2015). Social control, social learning, and cheating: Evidence from lab and online experiments on dishonesty. *Social Science Research*, 53: 311–324.

Kröll, Markus, and Devesh Rustagi (2016). Shades of Dishonesty and Cheating in Informal Milk Markets in India. SAFE Working Paper, 134, Goethe University, Frankfurt.

Koszegi, Botond, and Matthew Rabin (2006). A Model of Reference Dependent Preferences. *The Quarterly Journal of Economics*, 121 (4): 1133-1165.

Krupka, Erin L., and Roberto A. Weber (2013). Identifying social norms using coordination games: why does dictator game sharing vary? *Journal of the European Economic Association*, 11 (3): 495-524.

Lewis, Alan, Alexander Bardis, Chloe Flint, Claire Mason, Natalya Smith, Charlotte Tickle, and Jennifer Zinser (2012). Drawing the line somewhere: An experimental study of moral compromise. *Journal of Economic Psychology,* 33 (4): 718-725.

Lundquist, Tobias, Tor Ellingsen, Erik Gribbe, and Magnus Johannesson (2009). The aversion to lying. *Journal of Economic Behavior &* Organization, 70: 81-92.

Mann, Heather, Ximena Garcia-Rada, Lars Hornuf, Juan Tafurt, and Dan Ariely (2016). Cut From the Same Cloth: Similarly Dishonest Individuals Across cultures. *Journal of Cross-Cultural Psychology*, 47 (6): 858-874.

Maréchal, Michel André, Alain Cohn, Giuseppe Ugazio, and Christian C. Ruff (2017). Increasing honesty in humans with noninvasive brain stimulation. *Proceedings of the National Academy of Science USA*, 114(17): 4360–4364.

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

Meub, Lukas, Till Proeger, Tim Schneider, and Kilian Bizer (2016). The victim matters – Experimental evidence on lying, moral costs and moral cleansing. *Applied Economics Letters*, 23 (16): 1162-1167.

Muehlheusser, Gerd, Andreas Roider, and Niklas Wallmeier (2015). Gender differences in honesty: Groups versus individuals. *Economics Letters*, 128: 25–29.

Muñoz-Izquierdo, Nora, Beatriz Gil-Gómez de Liaño, Francisco Daniel Rin-Sánchez, and David Pascual-Ezama (2014). Economists: cheaters with altruistic instincts. MPRA Discussion Paper, Jena.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis G. Ipeirotis (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5 (5): 411–419.

Pascual-Ezama, David, Fosgaard, Toke R., Juan Camilo Cardenas, Praveen Kujal, Robert Veszteg, Beatriz Gil-Gómez de Liaño, Brian Gunia, Doris Weichselbaumer, Katharina Hilken, Armenak Antinyan, Joyce Delnoij, Antonios Proestakis, Michael D. Tira, Yulius Pratomo, Tarek Jaber-Lopez, and Pablo Branas-Garza (2015). Context-dependent cheating: Experimental evidence from 16 countries. *Journal of Economic Behavior & Organization*, 116: 379–386.

Ploner, Matteo, and Tobias Regner (2013). Self-image and moral balancing: An experimental analysis. *Journal of Economic Behavior & Organization*, 93: 374–383.

Potters, Jan, and Jan Stoop (2016). Do cheaters in the lab also cheat in the field? *European Economic Review*, 87: 26-33.

Rauhut, Heiko (2013). Beliefs about Lying and Spreading of Dishonesty: Undetected Lies and Their Constructive and Destructive Social Dynamics in Dice Experiments. *PLoS One*, 8 (11).

Rigdon, Mary L., and Alexander D'Esterre (2015). The effect of competition on the nature of cheating behavior, *Southern Economic Journal* 81 (4): 1012-1024.

Rosenbaum, Stephen Mark, Stephan Billinger, and Nils Stieglitz (2014). Let's be honest: A review of experimental evidence of honesty and truth-telling. *Journal of Economic Psychology,* 45, 181–196.

Ruffle, Bradley J., and Yossef Tobol (2014). Honest on Mondays: Honesty and the temporal separation between decisions and payoffs. *European Economic Review*, 65: 126–135.

Ruffle, Bradley J., and Yossef Tobol (2017). Clever enough to tell the truth. *Experimental Economics*, 20(1): 130-155.

Schindler, Simon, and Stefan Pfattheicher (2017). The frame of the game: Loss-framing increases dishonest behavior. *Journal of Experimental Social Psychology,* 69: 172-177.

Shalvi, Shaul (2012). Dishonestly increasing the likelihood of winning. *Judgment and Decision Making*, 7 (3): 292-303.

Shalvi, Shaul, Jason Dana, Michel J.J. Handgraaf, and Carsten K.W. de Dreu (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, 115 (2): 181-190.

Shalvi, Shaul, Ori Eldar, and Yoella Bereby-Meyer (2012). Honesty Requires Time (and Lack of Justifications). *Psychological Science*, 23 (10): 1264-1270.

Shalvi, Shaul, and David Leiser (2013). Moral firmness. *Journal of Economic Behavior & Organization*, 93: 400-407.

Shalvi, Shaul, and Carsten K.W. de Dreu (2014). Oxytocin promotes group-serving dishonesty. *Proceedings of the National Academy of Sciences*, 111 (15): 5503-5507.

Shen, Qiang, Meijun Teo, Eyal Winter, Einav Hart, Soo H. Chew, and Richard P. Ebstein (2016). To Cheat or Not To Cheat: Tryptophan Hydroxylase 2 SNP Variants Contribute to Dishonest Behavior. *Frontiers in Behavioral Neuroscience*, 10: 82.

Škoda, Samuel (2013). *Effort and cheating behavior: An experiment*. Bachelor Thesis. Charles University in Prague.

Suri, Siddharth, Daniel G. Goldstein, and Winter A. Mason (2011). Honesty in an Online Labor Market. *Human Computation*: Papers from the 2011 AAAI Workshop (WS-11-11): 61-66.

Suri, Siddharth, and Duncan J. Watts (2011). Cooperation and Contagion in Web-Based, Networked Public Goods Experiments. *PLoS One*, 6 (3).

Thielmann, Isabel, Benjamin E. Hilbig, Ingo Zettler, and Morten Moshagen (2016). On Measuring the Sixth Basic Personality Dimension: A Comparison Between HEXACO Honesty-Humility and Big Six Honesty-Propriety. *Assessment*: 1-13.

Utikal, Verena, and Urs Fischbacher (2013). Disadvantageous lies in individual decisions. *Journal of Economic Behavior & Organization*, 85: 108-111.

van 't Veer, Anna E., Mariëlle Stel, and Ilja van Beest (2014). Limited capacity to lie: Cognitive load interferes with being dishonest. *Judgment and Decision Making* 9 (3): 199–206.

Waubert De Puiseau, Berenike, and Andreas Glöckner (2012). Investigating Cheating Behavior in Students Compared to the General Public. Poster.

Weisel, Ori, and Shaul Shalvi (2015). The collaborative roots of corruption. *Proceedings of the National Academy of Sciences*, 112 (34): 10651-10656.

Wibral, Matthias, Thomas Dohmen, Dietrich Klingmüller, Bernd Weber, and Armin Falk (2012). Testosterone Administration Reduces Lying in Men. *PLoS One*, 7 (10).

Zettler, Ingo, Benjamin E. Hilbig, Morten Moshagen, and Reinout E. de Vries (2015). Dishonest responding or true virtue? A behavioral test of impression management. *Personality and Individual Differences*, 81: 107-111.

Zimerman, Liora, Shaul Shalvi, and Yoella Bereby-Meyer (2014). Self-reported ethical risk taking tendencies predict actual dishonesty. *Judgment and Decision Making*, 9 (1): 58–64.

Supplementary Materials for:

Loss aversion and lying behavior

Contents

Appendix 1. An extension of the model with reputation concerns

Appendix 2. Instructions of the experiment

(web link accessible as of 11/08/2018 https://az1.qualtrics.com/ControlPanel/?ClientAction=EditSurvey&Section=SV_bI9ZO3D0l3uo61v&SubSection=&SubSubSection=&PageActionOptions=&TransactionID=1&Repeatable=0)

Appendix 3. Tables:

Table A1: List of papers included in the review

Table A2: Estimates of the impact of the probability of the worst outcome and of the probability of the bad outcome on the mean lying rate, by category of observability in the extent literature

Table A3. Exogenous characteristics of the subjects in the experiment

Table A4: Influence of possible covariates on payout in the experiment

Appendix 1. An extension of the model with reputation concerns

We can extend the previous agent's utility function exposed in equation (1) with a reputational disutility for making a report to earn the high payoff (as theorized in recent research notably by Khameltsky and Sliwka, (2017), Dufwenberg and Dufwenberg (2018) and Gneezy *et al*. (2018)), $r(p,b)$:

$$U(x_R \mid x_O) = g(x_2 - x_e)*I(x_R \geq x_e) - h(x_e - x_1)*I(x_R < x_e) - m(x_R - x_O)*I(x_R \neq x_O) - r(p,b)*I(x_R = x_2) \quad (1')$$

$r(p,b)$ represents the potential loss in reputation from reporting the high payoff outcome which may seem suspicious of lying to other subjects or the experimenter who observe the report. b captures the observability of the report by others. We assume that $r()$ is increasing in p to capture that the more likely the bad outcome should have occurred, the more suspicious the report of a higher payoff will appear dishonest, and hence the greater the loss in reputation. We assume that $r(0,b) = 0$ for all b; if there is no chance to observe a low payoff, then there is no suspicion that the agent was behaving dishonestly. We further assume that $r(p,b)$ is an increasing function of the degree of observability b; if no one can observe the report (or identify the person making the report), then there is no loss of reputation (*i.e*., $r(p,0) = 0$), and as degree of observability increases (such as when the report has to be made face-to-face *vs*. on the phone, *vs*. on a computer) when the agent makes his report, the greater the possible suspicion (Individuals may feel more scrutinized when they have to report face-to-face to the experimenter than when they report on the phone or when the report is made on the computer without direct human interaction).

When observing the low payoff outcome, the agent's utility if he reports the high payoff outcome becomes:

$$U(x_2 \mid x_1) = g(x_2 - x_e) - m(x_2 - x_O) - r(p,b) \quad (4')$$

Thus, the difference in utility U(D) between dishonestly reporting observing the high payoff outcome and honestly reporting observing the low payoff outcome is:

$$U(D) = U(x_2 \mid x_1) - U(x_1 \mid x_1) = g(x_2 - x_e) + h(x_e - x_1) - m(x_2 - x_1) - r(p,b) \quad (5')$$

Differentiating U(D) with respect to the probability of the low outcome (p) yields:

$$U'(D) = (x_2 - x_1)[g'(p(x_2 - x_1)) - h'((1-p)(x_2 - x_1))] - r'(p,b) \text{ and} \quad (7''')$$

$$U''(D) = (x_2 - x_1)^2[g''(p(x_2 - x_1)) + h''((1-p)(x_2 - x_1))] - r''(p,b) \quad (8''')$$

It follows from (7''') that as the likelihood of the bad outcome p increases, agents lose increasingly more utility through reputation effects. We isolate the effect of such reputation in case of a change in p, assuming no reference dependence ($g(x) = kx$, $k>0$) and no loss aversion ($h(x) = g(x)$). In this case, the marginal utility of reporting dishonestly *vs*. honestly is:

$$U'(D) = k(x_2 - x_1) - k(x_2 - x_1) - r'(p,b) = -r'(p,b) \leq 0 \quad (7'''')$$

It thus follows that the agent will (weakly) be more likely to dishonestly report the higher payoff outcome the less likely the low payoff outcome may occur. However, note that if no one observes the report (*i.e*., $b=0$), then there is no effect on reputation and $U'(D) = -r(p,0) = 0$. Since loss aversion and reputation make similar predictions about the impact of the probability of the low payoff outcome, we employ two strategies in the empirical analysis. First, in our analysis of the extent literature we explore the interactive effects of a change in p overall and differing degrees of observability of the subjects (various b). Second, we present evidence of a new experiment using a mind game where subjects know they cannot be observed.

Appendix 2. Instructions for the experiment

Overview:

In addition to your normal pay for completing the task, you can get additional money for yourself. At the completion of this experiment, we will bonus you for any additional money you make. While the base pay for completing the task is modest, you might potentially get significant extra money.

How you will generate extra money:

The task involves predicting coin flips. The results of this coin flip will determine your extra pay.

Coin toss task:

You will be asked to toss a coin three times. If you do not have a coin at this time, please get a coin before proceeding to the next page.

Tossing the coin:

When you toss the coin, let it land on a flat surface so that one side of the coin is clearly facing-up and the other side is clearly facing-down.

Predicting the outcome of each coin toss:

Before each toss, predict which side of the coin will be face-up. To make your prediction, we would like you to make your prediction out loud before each coin toss. Thus, you can make just one of two possible predictions:

> "I predict Heads will be face-up" or
> "I predict Tails will be face-up."

Recording the outcome of your prediction:
In this task you will toss a coin a total of three times and report how many times you correctly predicted the coin toss. If you correctly predict the coin toss all three times, then report 3. If you correctly predict the coin toss twice, then report 2. If you correctly predict the coin toss just once, then report 1. And if you do not correctly predict any of the three coin tosses, then report 0.

Note that it does not matter in which order you make correct or incorrect predictions. For instance, if you correctly predicted the first toss only, or the second toss only, or the third toss only, then in all these cases you would report 1.

Valuing each correct prediction:

Depending on the outcome of your coin tosses you can get a $2 bonus payment.

The page will indicate how much you can get for each correct prediction. There will be an amount listed for getting 0 correct predictions, 1 correct prediction, 2 correct predictions and 3 correct

predictions. You will indicate your number of correct predictions by clicking on the appropriate option.

On this page we show you an example of how you will record the number of correct predictions. The payout schedule is shown below. This is only an example, the actual payout schedule may be different.

Example:

Click on the number of correct coin toss predictions

Payout Schedule:

0 correct prediction - Gain $0 additional reward
1 correct prediction - Gain $0 additional reward
2 correct predictions - Gain $2 additional reward
3 correct predictions - Gain $2 additional reward

You will thus only click on the number of correct predictions, and then click next to move to the next page.

Getting started:
If you have not done so yet, please get a coin. We would like you to not take any breaks while completing the rest of the task, so when you are ready, please continue to the prediction page and additional survey questions.

Prediction page differed between subject conditions
Prediction Page (0002 condition):
Please complete your three coin tosses now and click on the number of correct coin toss predictions. Depending on your outcome, you will receive payment of the amount listed.

- ❍ 0 correct prediction - Gain $0 additional reward (1)
- ❍ 1 correct prediction - Gain $0 additional reward (2)
- ❍ 2 correct predictions - Gain $0 additional reward (3)
- ❍ 3 correct predictions - Gain $2 additional reward (4)

Prediction Page (0022 condition):
Please complete your three coin tosses now and click on the number of correct coin toss predictions. Depending on your outcome, you will receive payment of the amount listed.

- ❍ 0 correct prediction - Gain $0 additional reward (1)
- ❍ 1 correct prediction - Gain $0 additional reward (2)
- ❍ 2 correct predictions - Gain $2 additional reward (3)
- ❍ 3 correct predictions - Gain $2 additional reward (4)

Prediction Page (0222 condition):
Please complete your three coin tosses now and click on the number of correct coin toss predictions. Depending on your outcome, you will receive payment of the amount listed.

- ❍ 0 correct prediction - Gain $0 additional reward (1)
- ❍ 1 correct prediction - Gain $2 additional reward (2)
- ❍ 2 correct predictions - Gain $2 additional reward (3)
- ❍ 3 correct predictions - Gain $2 additional reward (4)

Before we move on to the attitude and demographic questions, we would like to give you one more chance to earn more money for yourself. Like you, other participants in this study are reporting the outcomes of coin

tosses to earn money. We would like you to evaluate whether reporting more, less or the actual amount of money earned is "socially appropriate" and "consistent with moral or proper social behavior," or are "socially inappropriate" and "inconsistent with moral or proper social behavior." By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do. Another way to think about what we mean is that if the participant was to select a socially inappropriate choice, then someone else might be judgmental of this person for doing so. Please answer each question as accurately as possible, based on your opinion of what constitutes socially appropriate or socially inappropriate behavior. At the end of the experiment, we will randomly select one participant and one of the four questions below. If you are chosen and gave the same response to the question that is most frequently given by all other participants, then you will receive an additional $100 as a bonus to your MTurk account.

How socially appropriate or inappropriate is it to report...

	Very Socially Inappropriate (1)	Somewhat Socially Inappropriate (2)	Somewhat Socially Appropriate (3)	Very Socially Appropriate (4)
an outcome that earns $0 when the prediction would earn $0 (4)	❍	❍	❍	❍
an outcome that earns $2 when the prediction would earn $0 (1)	❍	❍	❍	❍
an outcome that earns $2 when the prediction would earn $2 (2)	❍	❍	❍	❍
an outcome that earns $0 when the prediction would earn $2 (7)	❍	❍	❍	❍

Before we finish, please complete the following set of questions about your experience, your views and yourself.

What type of coin did you use for the coin toss?

- ❍ Quarter (1)
- ❍ Dime (2)
- ❍ Nickel (3)
- ❍ Penny (4)
- ❍ Other (5)

If you predicted three coins tosses correctly, how much extra money would you have earned?

- ❍ $1 (1)
- ❍ $2 (2)
- ❍ $4 (3)
- ❍ $8 (4)
- ❍ $10 (5)

If we asked 1000 people to toss a coin three times, how many of them do you think would toss three heads?
_____ Please move the slide to your response (2)

How would you classify your political affiliation?

- ❍ Republican (1)
- ❍ Democratic (2)
- ❍ Independent (3)
- ❍ Other (4)

How much do you agree with the goals of the Republican National Party?
_____ Please move the slide to your response (1)

How much do you agree with the goals of the Democratic National Party?
_____ Please move the slide to your response (1)

How often do you vote in national elections (including Senate, House of Representative, and Presidential elections)?

- ❍ I never vote in national elections (1)
- ❍ I occasionally vote in national elections (2)
- ❍ I sometimes vote in national elections (3)
- ❍ I frequently vote in national elections (6)
- ❍ I usually vote in national elections (4)
- ❍ I always vote in national elections (5)

Please indicate your level of agreement with the following statements (1=Strongly disagree, 7=Strongly agree).

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
Luck plays an important part in everyone's life. (1)	❍	❍	❍	❍	❍	❍	❍
Some people are consistently lucky, and others are unlucky. (2)	❍	❍	❍	❍	❍	❍	❍
I consider myself to be a lucky person. (3)	❍	❍	❍	❍	❍	❍	❍
I believe in luck. (4)	❍	❍	❍	❍	❍	❍	❍
I often feel like it's my lucky day. (5)	❍	❍	❍	❍	❍	❍	❍
I consistently have good luck. (6)	❍	❍	❍	❍	❍	❍	❍
It's a mistake to base any decisions on how lucky you feel. (7)	❍	❍	❍	❍	❍	❍	❍
Luck works in my favor. (8)	❍	❍	❍	❍	❍	❍	❍
I don't mind leaving things to chance because I'm a lucky person. (9)	❍	❍	❍	❍	❍	❍	❍
Even the things in life I can't control tend to go my way because I'm lucky. (10)	❍	❍	❍	❍	❍	❍	❍
There is such a thing as luck that favors some people, but not others. (11)	❍	❍	❍	❍	❍	❍	❍
Luck is nothing more than random chance. (12)	❍	❍	❍	❍	❍	❍	❍

And finally a few questions about yourself... What is your gender?

- ❍ Male (1)
- ❍ Female (2)

What is your highest educational degree obtained?

- ❍ Primary school (1)
- ❍ Less than high school (2)
- ❍ High school diploma or equivalent (3)
- ❍ Undergraduate degree (4)
- ❍ Post-graduate degree (5)

What year were you born (e.g., 1980) (Dropdown box)

What is your approximate household annual pretax income?

- ❍ Less than $10,000 (1)
- ❍ Between $10,000 and $20,000 (2)
- ❍ Between $20,000 and $30,000 (3)
- ❍ Between $30,000 and $50,000 (4)
- ❍ Between $50,000 and $70,000 (5)
- ❍ Between $70,000 and $90,000 (6)
- ❍ Between $90,000 and $110,000 (7)
- ❍ Between $110,000 and $130,000 (8)
- ❍ Between $130,000 and $150,000 (9)
- ❍ More than $150,000 (10)

How much money do you spend in a typical week (this should be your daily expenses e.g., food, travel, mobile charges, purchases; but excluding rent, mortgage, educational fees, work expenses)?

- ❍ $0-$29 (1)
- ❍ $30-$49 (2)
- ❍ $50-$79 (3)
- ❍ $80-$119 (4)
- ❍ $120-$174 (5)
- ❍ $175-$249 (6)
- ❍ $250-$350 (7)
- ❍ More than $350 (8)

What do you think we are trying to assess with this study? (open-ended)

Appendix 3. Tables

Table A1. List of studies included in the review

Authors (Year)	Nb treatments	Nb subjects	Country	Lab	Students	Nb coin tosses	Nb die rolls	Nb other devices	Observability
1. Abeler *et al*. (2014)	4	1102	Germany	Lab+telephone	Both	1 and 4	-	-	1/2
2. Abeler (2015)	1	60	China	Lab	Yes	-	-	1 ticket/10	1
3. Abeler *et al*. (2016)	8	1098	UK	Lab	Yes	-	-	1 chip/50	1/2
4. Amir *et al*. (2016)	11	403	Israel	Online	Yes	20	-	-	0
5. Anthony *et al*. (2016)	2	200	Germany	Lab	Yes	-	1	-	1
6. Arbel *et al*. (2014)	2	399	Israel	Lab	Yes	-	1	-	1
7. Ariely *et al*. (2014)	2	188	Germany	Art. field	No	-	40	-	0
8. Aydogan *et al*. (2015)	2	120	Germany	Lab	Yes	2	-	-	1
9. Balasubramanian *et al*. (2017)	10	320	India	Online	No	6	-	-	0
10. Banerjee *et al*. (2017)	8	1344	India	Art. Field	No	-	1	-	1
11. Bar-El and Tobol (2017)	4	358	Israel	Lab	Yes	-	1	-	1
12. Barfort *et al*. (2015)	1	862	Denmark	Online	Yes	-	4*10	-	0
13. Beck *et al*. (2016)	6	191	Germany	Lab	Yes	-	1	-	1
14. Blanco and Cardenas (2015)	2	103	Colombia	Lab	Yes	-	1	-	1
15. Braun and Hornuf (2015)	7	342	Germany	Lab	Yes	-	40	-	0
16. Bryan *et al*. (2013)	7	238	USA	Lab + online	Both	10	-	1 nb/10	0/2
17. Bucciol and Piovesan (2011)	2	160	Italy	Art. field	No	1	-	-	1
18. Cadsby *et al*. (2016)	1	90	China	Lab	Yes	-	1	-	1
19. Charness *et al*. (Forthcoming)	2	173	Spain	Lab	Yes	-	1	-	1
20. Chylitova and Korbel (2014)	3	444	Czech Rep.	Art. field	Yes	-	1	-	2
21. Clot *et al*. (2014)	2	98	Madagascar	Lab	Yes	-	1	-	1
22. Cohn *et al*. (2014)	8	435	Switzerland	Art.field+ online	Both	10	-	-	0
23. Cohn and Marechal (Forthcoming)	1	162	Switzerland	Art. field	No	10	-	-	1
24. Cohn *et al.* (2015)	4	285	Switzerland	Art. field	No	10	-	-	1
25. Conrads *et al*. (2013)	4	554	Germany	Lab	Yes	-	1	-	1/2
26. Conrads and Lotz (2015)	4	246	Germany	Lab + tel+ online	Yes	4	-	-	0/1/2
27. Dai *et al*. (2018)	2	471	France	Art. field+lab	Both	-	1; 3-sided	-	1
28. Dato and Nieken (2015)	1	288	Germany	Lab	Yes	-	1	-	1
29. Dieckmann *et al*. (2015)	5	1015	Multiple	Online	No	1	-	-	0
30. Diekmann *et al*. (2015)	6	466	Switzerland	Lab	Yes	-	1	-	1
31. Di Falco *et al.* (2016)	1	1080	Tanzania	Lab	Yes	1	-	-	1
32. Djawadi and Fahr (2015)	1	252	Germany	Art. Field	Both	-	-	1 chip	1
33. Drupp *et al*. (2016)	4	164	Germany	Mail	Both	4	-	-	0
34. Duncan and Li (2016)	1	203	USA	Online	No	-	1	-	0

35. Effron *et al*. (2015)	6	1704	USA	Online	No	13 and 20	-	-	0
36. Fischbacher and Föllmi-Heusi (2013)	7	1347	Switzerland	Lab	Yes	-	1	-	0/1
37. Foerster *et al*. (2013)	2	28	Germany	Lab	Yes	-	6; 8-sided	-	1
38. Fosgaard *et al*. (2013)	4	209	Denmark	Lab	Yes	1	-	-	1
39. Gächter and Schultz (2016)	23	2568	Multiple	Lab	Yes	-	1	-	1
40. Gneezy *et al*. (Forthcoming)	3	314	Germany	Lab	Yes	-	-	1	1
41. Halevy *et al.* (2013)	1	51	Netherlands	Lab	Yes	-	60	-	1
42. Hanna and Wang (2017)	2	826	India	Art. field+lab	Both	-	42	-	1
43. Hilbig and Hessler (2013)	6	765	Germany	Lab	Yes	-	1	-	1
44. Hilbig and Zettler (2015)	6	549	Germany	Lab+Online	Both	2 and 10	1	-	0/1/2
45. Houser *et al.* (2016)	4	249	USA	Art. field	No	2	-	-	1/2
46. Hruschka *et al*. (2014)	8	223	Multiple	Art. field	No	30	-	-	1
47. Hugh-Jones (2016)	30	1475	Multiple	On line	No	1	-	-	0
48. Jacobsen and Piovesan (2016)	3	149	Denmark	Art. field	No	-	1	-	1
49. Jiang (2013)	2	39	Netherlands	Lab	Yes	-	20	-	0
50. Jiang (2015)	4	216	Multiple	Lab	Yes	-	20	-	0
51. Kajackaite (2018)	1	91	Germany	Lab	Yes	-	-	1 ticket/100	1
52. Kajackaite and Gneezy (2017)	17	1303	USA, Germ	Lab	Yes	-	1	1 ticket/10	0/1
53. Kroher and Wolbring (2015)	6	382	Germany	Lab + online	Yes	-	1	-	0/1/2
54. Mann *et al*. (2016)	10	2179	Multiple	Art. field+lab	Both	-	20	-	1
55. Maréchal *et al*. (2017)	5	301	Switzerland	Lab	Yes	-	10	-	1
56. Meub *et al*. (2016)	2	94	Germany	Lab	Yes	-	1	-	1
57. Muehlheusser *et al*. (2015)	1	108	Germany	Lab	Yes	-	1	-	1
58. Munoz-Izquierdo *et al*. (2014)	3	270	Spain	Art. field	Yes	1	-	-	1
59. Pascual-Ezama *et al.* (2015)	48	1440	Multiple	Lab+verbal+no contact	Yes	1	-	-	0/1/2
60. Ploner and Regner (2013)	6	316	Germany	Lab	Yes	-	1	-	1
61. Potters and Stoop (2016)	2	102	Netherlands	Lab	Yes	-	-	20 cards	0
62. Rauhut (2013)	3	240	Switzerland	Lab	Yes	-	4*12	-	1
63. Ruffle and Tobol (2014)	1	427	Israel	Art. field	No	-	1	-	2
64. Ruffle and Tobol (2017)	1	156	Israel	Art. field	No	-	1	-	2
65. Schindler and Pfattheicher (2017)	4	386	Multiple	Lab+on-line	Both	1	75	-	1
66. Shalvi (2012)	2	178	Netherlands	Lab	Yes	20	-	-	0
67. Shalvi *et al*. (2011)	2	129	USA	Lab	Yes	-	1	-	1
68. Shalvi *et al*. (2012)	4	144	Israel	Lab	Yes	-	1	-	1
69. Shalvi and Leiser (2013)	2	126	Israel	Lab	Yes	-	1	-	1
70. Shalvi and de Dreu (2014)	8	120	Netherlands	Lab	Yes	10	-	-	0
71. Shen *et al*. (2016)	1	205	Singapore	Lab	Yes	-	1	-	1
72. Skoda (2013)	3	90	Czech Rep.	Lab	Yes	-	1	-	1
73. Suri *et al*. (2011)	3	674	India, USA	Online	No	-	1	-	1

74. Thielmann *et al.* (2016)	1	152	Germany	Online	No	2	-	-	0
75. Utikal and Fischbacher (2013)	2	31	Germany	Art. field+lab	Both	-	1	-	0
76. van 't Veer *et al*. (2014)	2	173	Netherlands	Lab	Yes	-	1	-	1
77. Waubert de Puiseau, Glöckner (2012)	4	407	Germany	On-line	Both	5	-	-	0
78. Weisel and Shalvi (2015)	1	36	Germany	Lab	Yes	-	20	-	1
79. Wibral *et al*. (2012)	2	91	Germany	Lab	Yes	-	1	-	1
80. Zetler *et al*. (2015)	1	134	Germany	Online	No	2	-	-	0
81. Zimerman *et* al. (2014)	1	189	Israel	Online	Yes	20	-	-	0
Total	389	36668	44	53% of lab treatments	68.5% of treatments	48% of treatments	47% of treatments	5% of treatments	Mean: 0.67

Appendix Table A2. Estimates of the impact of the probability of the worst outcome and of the probability of the bad outcome on the mean lying rate, by category of observability in the extent literature

This Table reports the estimated coefficients (and standard errors in parentheses) of the **probability of the lowest outcome** in models (1) to (3) and of the **probability of the bad outcome** in models (4) to (6) on the mean probability to lie to avoid the lowest outcome (models (1) to (3)) or on the mean estimated probability to lie to avoid the bad outcome (models (4) to (6)). They come from 18 regressions based on the same models as those reported in Table 1 but estimated for each category of observability separately.

In these linear regressions, the dependent variable is the mean estimated lying rate with each treatment serving as one observation. Robust standard errors are clustered at the study level and sample weighting allows us to control for the large diversity of sample sizes across studies. In models (1) to (3), the bad outcome is defined as the lowest payoff outcome, whereas in models (4) to (6) it is defined as any outcome except the highest.

In model (1) the only independent variable is the probability of the lowest outcome and in model (4) the probability of the bad outcome. In models (2) and (5), we add the expected value of the lowest/bad outcome and of the good/highest outcome, in 2015 Purchasing Power Parities. They include a control for conducting the experiment in a standard university lab setting and a dummy variable indicating whether the study has been conducted by a team involving economists. Finally, in models (3) and (6) we add country dummies.

Category of observability	Lying to avoid the lowest outcome			N	Lying to avoid a bad outcome			N
	(1)	(2)	(3)		(4)	(5)	(6)	
Low observability	-0.488***	-0.540***	-0.467***	162	-0.517***	-0.288	-0.559**	162
	(0.092)	(0.087)	(0.076)		(0.167)	(0.327)	(0.218)	
Moderate observability	-0.327**	-0.273**	-0.317**	189	-0.838***	-0.882***	-0.613**	193
	(0.138)	(0.120)	(0.123)		(0.273)	(0.267)	(0.173)	
High observability	-0.404	0.080	0.248	34	0.120	-0.079	0.070	34
	(0.334)	(0.224)	(0.482)		(0.174)	(0.163)	(0.379)	

Note: *** and ** indicate significance at the 0.01 and 0.05 level, respectively.

Appendix Table A3. Exogenous characteristics of subjects in the experiment

This Table reports the exogenous characteristics of the subjects to our experiment. "Income" is percentage of people in income categories. "Spending" is percentage of people in weekly spending categories. The various demographics are randomly dispersed across the conditions. Pairwise testing shows no significant differences between conditions on gender (ps>0.14), age (ps > 0.14), education (ps > 0.90), income (p >0.07), or weekly spending (ps > 0.80).

Treatments	0002	0022	0222
N	80	327	571
% Females	57.5	48.3	52.7
Mean (SD) age	36 (11.7)	38 (12.0)	38 (12.1)
Education (in %)			
- Less high school	0	.6	.5
- High school	28.7	29.2	31.0
- Undergraduate	58.8	57.1	49.5
- Post-graduate	12.5	13.2	19.0
Income (in %)			
< $30k	35.0	24.6	25.9
$30k - $50k	16.3	26.5	24.3
$50k - $70k	12.5	22.1	21.0
$70k - $90k	15.0	10.4	10.7
> $90k	21.2	16.4	18.1
Spending (in %)			
< $50	15.0	15.7	16.1
$50 - $79	15.0	13.8	11.5
$80 - $119	23.8	21.3	23.3
$120 - $174	15.0	20.7	19.3
$175 - $249	15.0	13.5	14.0
> $250	16.3	15.1	15.9

Appendix Table A4: Influence of possible covariates on payout in the experiment

Our current technique does not allow for regression-based analysis, so to test the effects of covariates we have used the participant's payout level as a proxy for lying, since those who lied will have gotten the larger pays. Regressions are Ordered Logit models. "Income" is the percentage of people in income categories in real life. "Spending" is the percentage of people in weekly spending categories. Factor analysis of the belief in luck scale (Darke and Freedman 1997) forms two factors representing a belief that there is such a thing as 'luck' and a belief that the respondent is lucky.

DV: Participant Payout Beta (SE)	Overall	0002	0022	0222
N	927	78	304	545
Females	0.020(0.21)	0.387(0.51)	-0.120(0.33)	-0.304(0.52)
Year born	0.027(0.01)***	0.048 (0.02)**	0.037(0.01)***	0.037(0.02)**
Education	-0.242(0.15)	-0.601(0.43)	0.640(0.26)**	0.389(0.38)
Income	0.094(0.06)	0.060(0.14)	0.229(0.10)**	0.029(0.14)
Spending	0.000(0.06)	-0.159(0.18)	0.023(0.10)	0.092(0.15)
Luck Beliefs				
- Luck exists	0.021(0.09)	0.079(0.25)	-0.019(0.13)	0.149(0.20)
- I am lucky	0.002(0.10)	0.290(0.25)	0.143(0.14)	0.216(0.27)

Notes: *** and ** indicate significance at the 0.01 and 0.05 level, respectively. The number of observations in this Table is slightly lower than the total number of subjects since some subjects did not complete the demographic questionnaire.

As can be seen in this Table, the demographic and attitudinal variables have little effect on the participant's payout. Only age consistently reaches the traditional significance cutoff, such that younger people get higher payouts. Additional analyses (not shown) find no significant interactions of the covariates with the conditions (all p-values > 0.05), demonstrating that none of the covariates moderate the effects of treatment on payout.