



HAL
open science

Does the oath enhance truth-telling in eyewitness testimony? Experimental Evidence

Nicolas Jacquemet, Céline Launay, Stéphane Luchini, Danica Mijovic-Prelec, Drazen Prelec, Jacques Py, Julie Rosaz, Jason F. Shogren

► **To cite this version:**

Nicolas Jacquemet, Céline Launay, Stéphane Luchini, Danica Mijovic-Prelec, Drazen Prelec, et al.. Does the oath enhance truth-telling in eyewitness testimony? Experimental Evidence. 2024. halshs-04855141

HAL Id: halshs-04855141

<https://shs.hal.science/halshs-04855141v1>

Submitted on 24 Dec 2024

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 4.0 International License

CES

Centre d'Économie de la Sorbonne
UMR 8174

**Does the oath enhance truth-telling in eyewitness testimony?
Experimental Evidence**

Nicolas JACQUEMET, Céline LAUNAY, Stéphane LUCHINI,
Danica MIJOVIC-PRELEC, Drazen PRELEC, Jacques PY, Julie ROSAZ,
Jason F. SHOGREN

2024.13



Does the oath enhance truth-telling in eyewitness testimony? Experimental Evidence*

N. Jacquemet[†] C. Launay[‡] S. Luchini[§] D. Mijovic-Prelec[¶]
D. Prelec[¶] J. Py^{||} J. Rosaz^{**} J.F. Shogren^{††}

December 2024

Abstract

Eyewitness testimony is the most powerful form of evidence in a court of law. Eyewitnesses affect both the odds of conviction and the severity of sentences of the guilty. But eyewitnesses also lie, and false testimony is the primary cause of wrongful convictions. Most of the extant literature focuses on eyewitness reliability and credibility assessment, but very little is known about the efficiency of the main mechanism used in-field to foster eyewitness honesty: a solemn truth-telling oath—the most ancient and worldwide institution used in the solemn legal ceremony underpinning criminal cases. Herein we examine how the truth-telling oath actually affects the level of eyewitness deception. Using a controlled experimental test designed to address this question, we show that an eyewitness who is exogenously incentivized to lie and takes a solemn oath is significantly less likely to use deception. In contrast with the related literature focusing on the detection of lies, we show that an oath actually works to improve truth-telling. The oath is not just ceremonial, it plays a key role in improving efficiency within the court.

Keywords: Eye-witness testimony; Truth-telling oath; Controlled experiment.

*Authors are listed in alphabetical order. We gratefully acknowledge Rustam Romaniuc and Dimitri Dubois for their help in running the experiment. S.L would like to acknowledge the hospitality of the MIT Sloan Neuroeconomics Lab during his appointment at the Lab in 2011. In particular, he would like to thank Nouchine Hadjikhani, Alex Huang, Danica Mijovic-Prelec and Drazen Prelec for access to the experimental designs and behavioral and fMRI data (collected prior to 2011) on the impact of Oath on deception. We thank the financial support from the National Research Agency (program Investissements d’Avenir, ANR-10-LABX-93-0 and ANR-17-EURE-0001).

[†]Paris School of Economics and U. Paris 1 Panthéon-Sorbonne (CES), France. nicolas.jacquemet@univ-paris1.fr

[‡]CLLE, U. of Toulouse, CNRS, Toulouse, France. celine.launay@univ-tlse2.fr.

[§]Aix-Marseille U. (Aix-Marseille School of Economics), CNRS and EHESS, France. stephane.luchini@univ-amu.fr.

[¶]MIT Neuroeconomics laboratory, Sloan School of Management, Department of Economics, Department of Brain and Cognitive Sciences. Cambridge, US. mijovic@mit.edudprelec@mit.edu.

^{||}CLLE, U. of Toulouse, CNRS, Toulouse, France. Jacques.Py@univ-tlse2.fr

^{**}Burgundy School of Business, Lyon France. julie.rosaz@bsb-education.com.

^{††}Department of Economics, U. of Wyoming, Laramie, United States. JRamses@uwyo.edu.

Eyewitness testimony is the most powerful form of evidence in court, affecting both convictions and the severity of sentencing (Loftus, 2003). In fact, most wrongful convictions are not the result of unreliable witnesses (Wixted et al., 2018; Diamond et al., 2020) but rather deceptive ones, well ahead of false or misleading forensic evidence (Gross and Shaffer, 2012). While a significant literature has emerged since the 1920’s focusing on the ability to detect false testimony while maximizing the discriminative power of lie detection tools (Farah et al., 2014; Vrij and Fisher, 2016), no method has yet been empirically proven to highly discriminate truth from deception in the field (Farah et al., 2014).

This paucity of direct evidence on truth-telling efficiency holds even for the solemn oath — the most ancient procedure used in criminal systems to regulate and reduce deception (Silving, 1959; Spurr, 2001). Recent evidence in behavioral science however shows that compliance with a truth-telling oath is instrumental in fostering truth-telling (see, e.g., Jacquemet et al., 2019, 2021; Zickfeld et al., 2023, Zickfeld et al., 2024a for a meta-analysis and Zickfeld et al., 2024b for a large-scale mega-study and a theoretical discussion of the mechanisms leading to such behavioral effects).

Herein we provide an experimental test of the efficiency of the oath as a deception regulator of eyewitness testimony. Our research design relies on a new experimental procedure based on a multidisciplinary approach, using insights from both legal psychology, from which we borrow the design of the eye-testimony task, and experimental economics, using performance-based incentive to induce preferences towards either truth-telling or deception. In all conditions, participants viewed a one-minute video and then took a unique two-alternative forced-choice recognition test on 28 details from the video. The main outcome of interest is the level of accuracy in the recognition test.

We rely on performance-based monetary incentives to generate a benchmark of both the level of accuracy that can be achieved by participants given their memory limitations (*Truth*), and the extent of deception that occurs when participants are externally incentivized to lie (*Deception*). We also assess the ecological validity of this design by comparing these two incentivized treatment with a control condition in which incentives are removed (*Control*). Based on existing experimental evidence, we hypothesized that participants will respond to incentives by *(i)* providing as accurate answers as possible when they are incentivized to do so, *(ii)* lie to some extent when deception is incentivized, but *(iii)* do not lie maximally because of an intrinsic preference to tell the truth. All three hypotheses are confirmed in the data, which allows us to assess the effect of the oath on deceptive eye-witnesses.

The main manipulation of interest is introduced immediately after viewing the video, when participants are directed to a private desk in which they receive a form to sign. In the *Oath* conditions, participants receive a form entitled “solemn oath” that contains a unique sentence with a single prescription that reads “*I, ..., the undersigned do solemnly swear that during the whole experiment, I will tell the truth and always provide honest answers*”. This treatment is combined

with the manipulation of incentives according to a factorial design, resulting in a 3 (retrieval condition: *Control*, *Truth*, *Deception*) \times 2 (oath: *Oath*, *No oath*) between subjects design. While the level of accuracy is remarkably constant when participants who are exposed to an oath are at the same time incentivized to tell the truth, the oath manages to discipline a large part of the lies arising from the external incentives to do so. Based on a counter-factual analysis, we estimate that 35pp of potential deception is removed by the intrinsic preference for truth-telling. The oath further removes 35pp of the deception that would occur otherwise.

Method

General design

Once entering the lab, participants viewed a one-minute video depicting a typical non-violent robbery scene (see Appendix A, Section A.2, for more details on the video-stimulus, and Section A.1 for more details about the procedure of the experiment). Immediately after viewing the video, participants were led one by one to a private desk where a monitor instructed them to sign a form before entering the final phase of the experiment. In the *No-oath* conditions, participants were provided with a form labeled “Consent form”. In the *Oath* conditions, they were provided with a form labeled “Solemn oath” that contained the following sentence: “I swear to tell the truth and to always provide honest answers” (the form is presented in Appendix A, Section A.6). Signing the form was compulsory to participate in the experiment, as is the case with the implementation of the oath in courts.

Each participant in all conditions then entered a second experimental room where another monitor was waiting. Participants were instructed to perform a computerized counting task for 20mn, as an unrelated “distractor” task (Chan and LaPaglia, 2013), for which they earned a fixed monetary compensation of 12€. The distractor task served two purposes: (*i*) to clear participants’ short term memories of the robbery video they watched earlier, and (*ii*) to create the sense that each participant “earned” his or her compensation (earned money is preferred to windfall money to engage participants, e.g., Jacquemet et al., 2009; and to decrease the role of social preferences, Cherry et al., 2002). Participants then took a unique two-alternative forced-choice recognition test on 28 details from the video, for which they were instructed to be as accurate as possible. For example, they were asked “*the victim dropped an object in front of [a] the car’s front wheel [b] the car’s back wheel*” (see Appendix B, Section A.4 for the full questionnaire). The correct answer was randomly affected to response [b] in 75% of the questions to avoid that purely random behavior would be payoff maximizing. Participants were not given a time limit to answer any question. No corrective feedback was provided.

To provide extrinsic motivation to make deceptive responses without explicitly instructing participants to do so, which is the basis of most paradigms of lie detection (Farah et al., 2014; Vrij et al., 2019), we provided participants with monetary incentives to lie. In this *Deception* condition,

Table 1: Summary of the experimental design

<i>No oath</i>				<i>Oath</i>		
Condition	Incentives	<i>N</i>	Oath	Condition	<i>N</i>	Oath
<i>Control</i>	No	31	NO	<i>Control under oath</i>	40	YES
<i>Truth</i>	-30cts / incorrect answer	36	NO	<i>Truth under oath</i>	40	YES
<i>Deception</i>	-30cts / correct answer	38	NO	<i>Deception under oath</i>	37	YES

Note. Summary of the experimental design and the number of participants (*N*) of all 6 between-subjects conditions.

participants face an honesty cost: they were instructed that each correct answer will cause them to lose 30cts, subtracted from the compensation they earned in the distractor task.

We complement this evidence with two baseline conditions. First, the *Truth* condition implemented explicit monetary incentives to tell the truth: participants were told that each *incorrect* answer will cause them to lose 30cts. This condition thus provides a behavioral benchmark of retrieval performance when participants are (externally) motivated to testify truthfully. To ensure that monetary incentives to tell the truth did not alter memory performances, we added a *Control* condition in which participants were told their answers (either correct or incorrect) would not affect their earnings, as tested in classic legal psychology experiments. At the end of the session, participants answered a brief demographic questionnaire (see Appendix A, Section A.3-A.5 for the material used in the experiment).

Participants

The participants are volunteers recruited through an online platform to participate in a study in the Montpellier experimental laboratory (whose on-line registration system is based on ORSEE; Greiner, 2015). All participants signed an online consent form when they registered on the online platform. We did not apply any exclusion restriction in the invitation to participate, and confirm participation on a first-in-first served basis.

In total, 222 subjects participated in this experiment, with 15 to 20 participants in each experimental session. There is an almost perfect balance in the sex of our participants, with 52.3% of male participants. The age of participants is relatively low as compared to the general population, with a mean age equal to 27 years old and a standard deviation of 7.22. This is usual with experimental lab subjects pools in which students are typically over-represented (see, e.g. Jacquemet and L’Haridon, 2018). The diversity of religious feelings in our sample matches well the one observed in the general population in France, with 34.5% of the sample self-reporting no religious affiliation and Catholics (21.4%) and Muslims (30.9%) being the two most important groups (followed by Jews, 0.9%, Buddhist, 0.9% and Protestants 4.1%, the remaining 8.2% choosing the ‘Other’ answer). In total, 12 sessions were carried out, 2 for each condition. The sample

sizes for each experimental condition are presented in Table 1.

Results

Validation of experimental conditions

We begin by examining the baseline conditions: *Truth* without and with the oath. Figure 1.a shows the average accuracy by item in the *Truth* condition compared to the *Truth under oath* condition (in blue). In the *Truth* condition without the oath, memory performance ranges from 19.4% to 100.0% across items, with participants correctly recognizing 67.3% of items on average (see Appendix B, Section B.1 for detailed evidence on the proportion of correct responses by item and condition). This result aligns closely with what is typically observed in real-life forensic investigations (Kebbell et al., 1996).

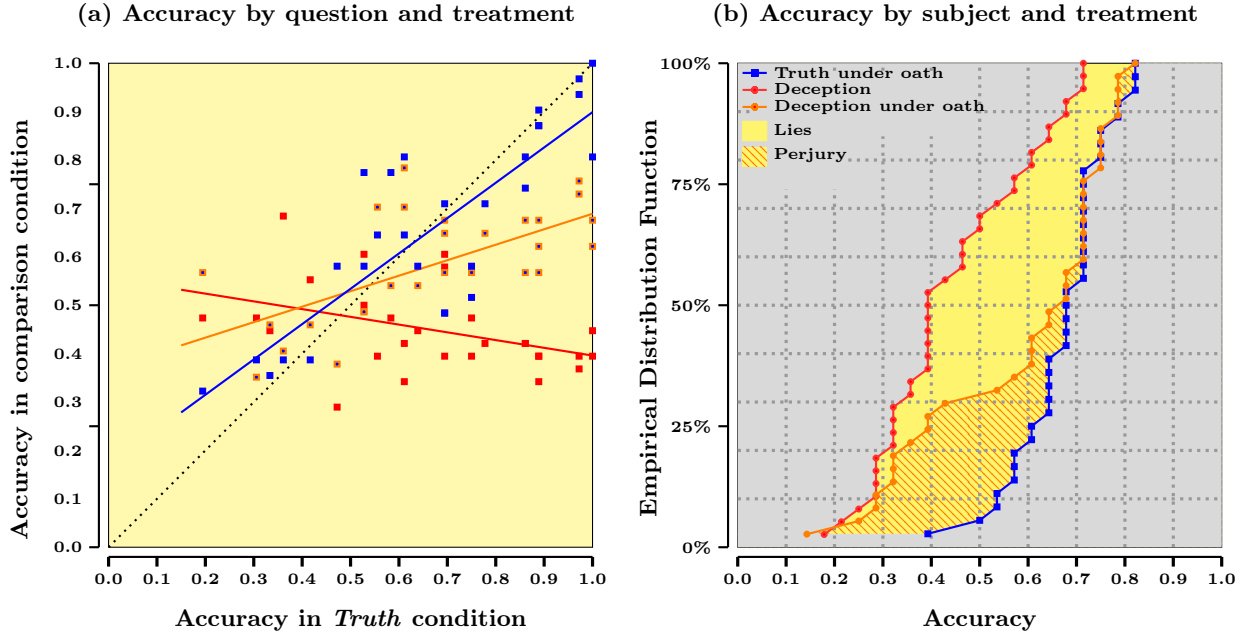
As anticipated, adding the oath to the *Truth* baseline condition does not significantly alter the average level of accuracy, which remains close at 67.8%. This is further illustrated by the blue regression line in Figure 1.a, showing comparable recognition rates with and without the oath. Similarly, this average level of accuracy remains consistent in the *Control* conditions, where no monetary incentives are present (see Appendix B, Section B.1 for detailed evidence on the proportion of correct responses by item and by condition). These findings confirm the ecological validity of the main experiment by ensuring that the addition of monetary incentives does not distort the natural behavior observed in baseline conditions.

Manipulation check: assessing the effectiveness of deception induction

We now focus on the key condition of interest: *Deception*, both with and without the oath. Monetary losses tied to giving correct responses lead to a significant decline in retrieval performance. As shown in Figure 1.a (red panel), the average accuracy in the Deception condition drops to 44.8%, compared to 67.3% in the *Truth* condition. This result aligns with findings from behavioral economics research on deceptive behavior. For example, Gneezy (2005) observed that individuals are more likely to convey false information when incentivized to do so, and Fischbacher and Föllmi-Heusi (2013) found that participants over-report outcomes in incentivized die-roll experiments.

An important difference between the existing deceptive behavior literature and our experiment is that our participants do not always know what the truth is when they testify. In our experiment, participants are constrained by their memory abilities so that they can only lie as long as they know the correct answer. The regression line in red in Figure 1.a indicates that deceptive behavior mainly occurs for video items that are well recognized in the *Truth* condition (the accuracy gap between the two conditions is highest for items whose accuracy is high in the *Truth* condition). Figure 1.b presents the cumulative distribution of the average accuracy at the participant level (Empirical Distribution Functions, EDF). The EDF in the *Truth under oath* condition (which is identical to the EDF observed in the *Truth* condition, see Appendix B, Section B.3) first order

Figure 1: Accuracy by treatment at the aggregate and individual levels



Note. Panel (a) presents the distribution of memory performance according to whether monetary incentives for truth-telling are combined with the oath (blue) or deception with honesty costs either without the oath (red) or with the oath (orange) are implemented. Each square refers to an item in the questionnaire, for which the *x*-axis reports the average accuracy observed in *Truth* and the *y*-axis reports the average accuracy observed in the three corresponding conditions. An item for which the accuracy is identical in both conditions would appear on the 45 degree line. A regression line is plotted for each of the three conditions against *Truth*: the blue regression line compares memory performance in the *Oath* condition to the *Truth* baseline; the red regression line compares the *Deception* condition to the *Truth* baseline, and the orange regression line compares the *deception under oath* condition to the *Truth* baseline. A regression line on the 45 degree line indicates no difference in accuracy at the aggregate level between the two conditions being compared. Panel (b) presents the data at the individual level and shows the Empirical Distribution Functions (EDF, i.e., the empirical cumulative distribution) of the proportion of recognized items computed at the subject level. The gap observed between the *Deception* and the *Truth* conditions (without oath) measures the degree of deception. The gap between *deception under oath* and *Truth* measures the extent of perjuries — i.e., lying behavior under oath. Since we find that the EDF of the *truth under oath* condition is very similar to that of the *Truth* (without oath) condition, we omit it from the Figure to ease the reading (see Appendix B, Sections B.2 and B.3).

dominates the EDF in the *Deception* condition ($p < .001$; see Appendix B, Section B.4), meaning that subjects do respond to the monetary incentives to lie in the *Deception* condition whatever their intrinsic level of accuracy (as measured by their performance in the *Truth* condition).

Deceptive eye-witnesses under oath

The orange regression line in Figure 1.a shows the oath induces a sharp decrease in deceptive behavior in the *Deception under oath* condition (as compared to the red line) although items that are well-recognized in the *Truth* condition (in blue) are recognized with a lower level of accuracy in the *Deception under oath* condition. This is confirmed by the statistical comparison of the EDF displayed in Figure 1.b: the distribution of lying behavior in *Deception under oath* first order dominates the EDF in the *Deception* condition ($p = .024$), but is not significantly first order

dominated by the EDF in *Truth* ($p = .165$).

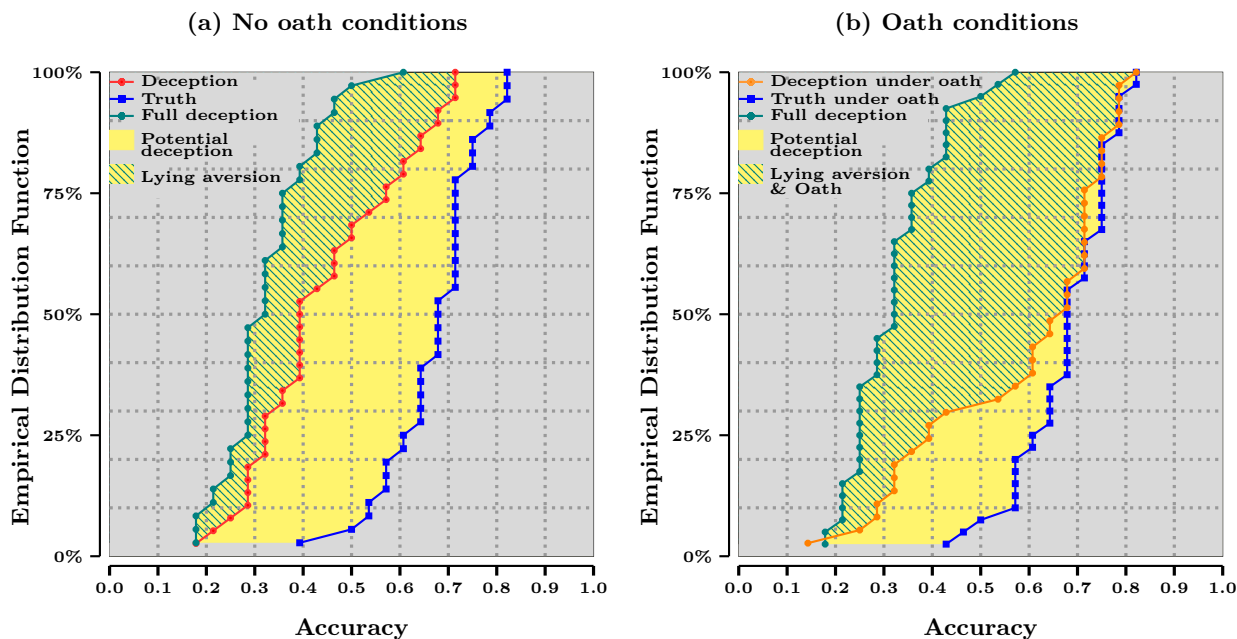
The results also reveal that lying behavior that persists under oath is not uniformly distributed across participants. Instead, we identify three distinct groups: at the lower end of the accuracy distribution, participants in the *Deception under oath* condition behave similarly to those in the *Deception* condition (without oath, in red). This group, representing 20% of the sample, chooses to lie under oath, effectively engaging in perjury. At the other end of the spectrum, 50% of participants exhibit retrieval performance at the upper end of the accuracy distribution, closely resembling that observed in the *Truth* baseline condition (without oath, in blue). For this group, the oath successfully shifts decisions away from incentive-driven deception. Finally, 30% of participants fall in between these two extremes, choosing a level of deception that is intermediate between the two benchmarks. These participants navigate the conflicting motivations of the oath and monetary incentives by selecting a moderate level of deception. A similar pattern has been observed in tax-evasion games (Jacquemet et al., 2020)

We conducted additional analyses to evaluate the power and robustness of our behavioral findings. First, the results from unconditional statistical tests on the effects of honesty costs and the oath are confirmed using OLS and Poisson regression models, which control for a range of demographic and personal variables, including gender, age, religious background, prior participation in experiments, number of siblings, and birth order (see Appendix B, Sections B.5 and B.6). Second, we find that the introduction of monetary incentives for deception leads to a decrease in self-reported happiness, suggesting a tension between truth-telling and financial self-interest (see Extended Data B.7). Third, self-reported honesty closely mirrors the behavioral findings: it declines with monetary incentives for deception and increases when subjects are under oath, except for approximately one-third of participants (see Extended Data B.8). Finally, we validate the efficacy of the oath in promoting truth-telling using an alternative implementation, where participants had the option to voluntarily comply with the request to sign the oath (see Extended Data B.9).

Assessing the effectiveness of the oath

Accumulated evidence in the behavioral sciences shows that people do not lie maximally due to an intrinsic motivation to tell the truth (Prelec et al., 2015; Gneezy et al., 2018; Abeler et al., 2020; Köbis et al., 2019). We now compare this intrinsic motivation to the external motivation to tell the truth induced by the oath. To do so, we construct a benchmark of full deception in the experiment—the level of accuracy that would result if participants lied maximally within the constraints of their memory. Participants who are incentivized to be dishonest (through the induced honesty cost) and who fully respond to these extrinsic incentives should choose answers they believe to be incorrect. Accordingly, we construct a counterfactual distribution by assuming that participants in the *Truth* condition provide a baseline measure of true memory performance, which is then inverted to model full deception. Figure 2.a illustrates this counterfactual distribution, representing the theoretical

Figure 2: Accuracy by subject and treatment: Response to honesty costs by treatment



Note. The figures compare the observed EDF of accuracy to a theoretical construct (in green) reporting the EDF of full deception that would result from the observed accuracy in the *Truth* condition. This is built by reversing the observed answers of participants in the *Truth* condition, i.e., by operating a central symmetry of the EDF in the *Truth* condition (in blue) with regards to the point (0.5;0.5). The participant at the top of the EDF in the *Truth* condition becomes the participant at the bottom of the distribution of the full deception EDF. The area in yellow measures the extent of potential deceptive behavior would participants respond fully to honesty costs — i.e., 100% deception conditional on the retrieval performance observed in *Truth*. The area in shaded green measures how lying aversion affects deceptive behavior in **Panel (a)**, and the combined effect of lying aversion and the oath in **Panel (b)**.

empirical distribution function (EDF) of individual accuracy under maximal deception.

The comparison with the actual distribution in the *Deception* condition confirms an intrinsic propensity to tell the truth for most participants: the two distributions coincide only for a small subset of participants, who fully respond to the honesty costs. Under the assumption of full deception, the average level of accuracy would be 32.7%. By contrast, mean accuracy in the *Deception* condition is 44.8%, compared to 67.3% in the *Truth* condition. These results suggest that intrinsic motivation to tell the truth mitigates 35.0% of potential deception (see Appendix B, Section B.10 for computational details).

In Figure 2.b, we extend this analysis to the conditions with the oath, using the *Truth under oath* condition as the baseline for true memory performance. Under the assumption of full deception, the mean accuracy level would drop to 32.2%, compared to 67.8% in *Truth under oath* and 58.7% in *Deception under oath*. Combining the oath with intrinsic lying aversion eliminates 74.4% of potential deception. Of this, the oath itself accounts for a reduction of 39.4% (74.4% - 35.0%), leaving only 25.6% of perjury intact.

Discussion

In the legal psychology literature, deception has been considered as an estimator variable, beyond the control of the justice system (similar to lighting conditions or suspect ethnic group; in contrast to a system variable that is controlled by the legal system, such as police interviewing techniques, Wells, 1978; Wells and Olson, 2003). But with the solemn truth-telling oath, the justice system can control deception. While the oath was discussed early on in the history of research in legal psychology as an enhancer of truth-tellers' performances (e.g., Stern, 1902), most of the recent literature rather focuses on lie detection (e.g., Farah et al., 2014). The gap is surprising given the key role the oath plays to help solemnize court proceedings (Spurr, 2001). Our results show that a oath reduces perjury significantly, while maintaining a truth teller's performance. It provides strong support to the idea that the oath remains essential to the legal system: it provides a non-monetary commitment to tell the truth over and above intrinsic motivation. Given that the oath helps to create a collective morality (Prodi, 1992), this also gives empirical support to the recent and widespread calls to introduce formal oath taking procedures into business ethics and environmental protection (Cohn et al., 2014).

References

- Abeler, J., Nosenzo, D., Raymond, C., 2020. Preferences for truth-telling. *Econometrica* 87, 1115–1153.
- Chan, J.C.K., LaPaglia, J.A., 2013. Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciences* 110, 9309–9313. doi:10.1073/pnas.1218472110.
- Cherry, T., Frykblom, P., Shogren, J., 2002. Hardnose the Dictator. *American Economic Review* 92, 1218–1221.
- Cialdini, R.B., 2007. *Influence: The Psychology of Persuasion*. volume 55. Collins New York.
- Cohn, A., Fehr, E., Maréchal, M., 2014. Business culture and dishonesty in the banking industry. *Nature* 516, 86–89. doi:10.1038/nature13977.
- Diamond, N.B., Armson, M.J., Levine, B., 2020. The Truth Is Out There: Accuracy in Recall of Verifiable Real-World Events. *Psychological Science* 31, 1544–1556. doi:10.1177/0956797620954812.
- Farah, M.J., Hutchinson, J.B., Phelps, E.A., Wagner, A.D., 2014. Functional MRI-based lie detection: Scientific and societal challenges. *Nature Reviews Neuroscience* 15, 123–131.
- Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in Disguise. An experimental study on cheating. *Journal of the European Economic Association* 11, 525–547. doi:10.1111/jeea.12014.
- Gneezy, U., 2005. Deception: The Role of Consequences. *American Economic Review* 95, 384–394. arXiv:4132685.

- Gneezy, U., Kajackaite, A., Sobel, J., 2018. Lying Aversion and the Size of the Lie. *American Economic Review* 108, 419–453.
- Greiner, B., 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1, 114–125.
- Gross, S.R., Shaffer, M., 2012. Exonerations in the United States, 1989-2012. Technical Report 1st. National Registry of Exonerations.
- Jacquemet, N., James, A.G., Luchini, S., Murphy, J.J., Shogren, J.F., 2021. Do truth-telling oaths improve honesty in crowd-working? *PLoS ONE* 16, e0244958. doi:10.1371/journal.pone.0244958.
- Jacquemet, N., Joule, R.V., Luchini, S., Shogren, J., 2009. Earned wealth, engaged bidders? Evidence from a second-price auction. *Economics Letters* 105, 36–38.
- Jacquemet, N., Joule, R.V., Luchini, S., Shogren, J.F., 2013. Preference Elicitation under Oath. *Journal of Environmental Economics and Management* 65, 110–132. doi:10.1016/j.jeem.2012.05.004.
- Jacquemet, N., L'Haridon, O., 2018. *Experimental Economics: Method and Applications*. Cambridge University Press.
- Jacquemet, N., Luchini, S., Malézieux, A., Shogren, J., 2020. Who'll stop lying under oath? Experimental evidence from tax evasion games. *European Economic Review* 20, 103369.
- Jacquemet, N., Luchini, S., Rosaz, J., Shogren, J.F., 2019. Truth-Telling under oath. *Management Science* 65, 426–438.
- Joule, R.V., Beauvois, J.L., 2010. *La soumission librement consentie. Comment amener les gens à faire librement ce qu'ils doivent faire?* Presses Universitaires de France, Paris.
- Kebbell, M.R., Wagstaff, G.F., Covey, J.A., 1996. The influence of item difficulty on the relationship between eyewitness confidence and accuracy. *British Journal of Psychology* 87, 653–662. doi:10.1111/j.2044-8295.1996.tb02614.x.
- Köbis, N.C., Verschuere, B., Bereby-Meyer, Y., Rand, D., Shalvi, S., 2019. Intuitive honesty versus dishonesty: Meta-analytic evidence. *Perspectives on Psychological Science* 14, 778–796.
- Launay, C., Py, J., 2017. Capturing the scene: Efficacy test of the re-enactment investigative instruction. *The Journal of Forensic Practice* 19, 174–189.
- Loftus, E., 2003. Our changeable memories: Legal and social implications. *Nature Reviews Neuroscience* 4, 422–422. doi:10.1038/nrn1132.
- MacKerron, G., 2012. Happiness economics from 35 000 feet. *Journal of Economics Surveys* 26, 705–735.
- Prelec, D., Luchini, S., Huang, A., Long, C., Hadjickani, N., Mijovic-Prelec, D., 2015. Brain Mechanism of self-signaling, under oath. Presentation at Workshop on "Self-deception, Self-signaling, and Self-Control," IAST, Toulouse, June 22-23, Toulouse. .

- Prodi, P., 1992. *Il Sacramento Del Potere: Il Giuramento Politico Nella Storia Costituzionale Dell'Occidente*. volume 15. il Mulino, Bologna, Italia.
- Sekhon, J., 2011. Multivariate and Propensity Score Matching Software with Automated Balance Optimization. *Journal of Statistical Software* 42, 1–52.
- Silving, H., 1959. The Oath: I. *Yale Law Journal* 68, 1329–1390.
- Spurr, J., 2001. A Profane History of Early Modern Oaths. *Transactions of the Royal Historical Society* 11, 37–63.
- Stern, L., 1902. Zur Psychologie der Aussage [On the psychology of testimony]. *Zeitschrift für die gesamte Strafrechtswissenschaft* 22, 315–370.
- Vrij, A., Fisher, R., 2016. Which lie detection tools are ready for use in the criminal justice system. *Journal of Applied Research in Memory and Cognition* 5, 302–307.
- Vrij, A., Hartwig, M., Granhag, P.A., 2019. Reading Lies: Nonverbal Communication and Deception. *Annual Review of Psychology* 70, 295–317. doi:10.1146/annurev-psych-010418-103135.
- Wells, G., 1978. Applied eyewitness-testimony research: System variables and estimator variables. *Journal of Personality and Social Psychology* 36, 1546–1557.
- Wells, G.L., Olson, E.A., 2003. Eyewitness Testimony. *Annual Review of Psychology* 54, 277–295. doi:10.1146/annurev.psych.54.101601.145028.
- Wilford, M.M., Chan, J.C.K., Tuhn, S.J., 2014. Retrieval enhances eyewitness suggestibility to misinformation in free and cued recall. *Journal of Experimental Psychology: Applied* 20, 81–93. doi:10.1037/xap0000001.
- Wixted, J.T., Mickes, L., Fisher, R.P., 2018. Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science* 13, 324–335.
- Zickfeld, J.H., Karg, S.T.S., Engen, S.S., Gonzalez, A.S.R., Michael, J., Mitkidis, P., 2024a. Committed (dis)honesty: A systematic meta-analytic review of the divergent effects of social commitment to individuals or honesty oaths on dishonest behavior. *Psychological Bulletin* 150, 586–620. doi:10.1037/bu10000429.
- Zickfeld, J.H., Ścigała, K.A., Elbæk, C.T., Michael, J., Tønnesen, M.H., Levy, G., Ayal, S., Thielmann, I., Nockur, L., Peer, E., Capraro, V., Barkan, R., Bø, S., Bahník, Š., Nosenzo, D., Hertwig, R., Mazar, N., Weiss, A., Koessler, A.K., Montal-Rosenberg, R., Hafenbrädl, S., Nielsen, Y.A., Kanngiesser, P., Schindler, S., Gerlach, P., Köbis, N., Jacquemet, N., Vranka, M., Ariely, D., Martuza, J.B., Feldman, Y., Białek, M., Woike, J.K., Rahwan, Z., Seidl, A., Chou, E., Kajackaite, A., Schudy, S., Glogowsky, U., Czarna, A.Z., Pfattheicher, S., Mitkidis, P., 2024b. Effectiveness of ex ante honesty oaths in reducing dishonesty depends on content. *Nature Human Behaviour* doi:10.1038/s41562-024-02009-0.
- Zickfeld, J.H., Ścigała, K.A., Weiss, A., Michael, J., Mitkidis, P., 2023. Commitment to honesty oaths decreases dishonesty, but commitment to another individual does not affect dishonesty. *Communications Psychology* 1, 1–18. doi:10.1038/s44271-023-00028-7.

A Extended Methods

A.1 Procedure

Each participant was randomly assigned to one of the six conditions. Each condition followed a similar six-step procedure. The experimenters remain the same through all experimental sessions.

Step 1: Once entering the lab, participants were greeted by a female monitor and viewed the film-stimulus in groups of up to 20 people (see Section A.2). Participants were informed a memory test will follow. The female monitor stayed in the room the entire time, so that participants cannot discuss the film-stimulus.

Step 2: Immediately after viewing the video, each participant was led to a private desk with a male monitor who instructed him or her to sign a form before entering the next step of the experiment. In the no-oath conditions, the monitor presented a participant with a form labeled “Consent form”. In the oath conditions, they were provided with a form labeled “solemn oath” (see Section A.6). Signing the forms was compulsory to pursue the participation in the experiment.

Step 3: Each participant then entered a second experimental room where the first female monitor was waiting; he or she were assigned to a group of up to 20 participants. They were seated in front of individual computers with wall partitions separating the computer terminals and instructed to read the instructions. Then, the monitor read out-loud all the experimental instructions from the relevant condition.

Step 4: Each participant then first performed a computerized counting task for approximately 20mn, as an unrelated “distractor” task, for which they earned a fixed monetary compensation of approximately 12 euros. This task was designed so that each participant could check and correct their answer until it was correct. The distraction task served two purposes: *(i)* to clear participants’ short-term memories of the film-stimulus, and *(ii)* to create the sense that each participant earned his or her compensation.

Step 5: Participants then took a unique two-alternative forced-choice recognition test on 28 items from the video. We conducted a pilot study (Launay and Py, 2017) to construct the recognition test, in which participants ($N = 40$) viewed the video and were then instructed to provide as much details as possible about the video. Items that were not recalled spontaneously too frequently (by less than 25% of the participants) were used to construct the recognition test. Those items dealt with different video details (see Wilford et al., 2014, for a similar procedure). Participants in all conditions took the recognition test for which they were instructed to be as accurate as possible (Chan and LaPaglia, 2013; Wilford et al., 2014, see Section A.4 for the full questionnaire). The correct answer was randomly affected to response (b) in 75% of the questions to avoid that purely random behavior would be payoff maximizing. Participants were presented with the 28 questions one at a time and were instructed the questions will be presented in a chronological order according to the film-stimulus. Participants were not given a time limit to answer any question and were not provided with any corrective feedback. Here the manipulation of the retrieval condition took place: depending on the condition, participants are further instructed that they will lose 0.30 euros for each correct (*Deception* conditions) or incorrect (*Truth* conditions) answer.

Step 6: Participants then answered a brief demographic questionnaire (see Section A.5).

A.2 Movie

The one-minute duration film-stimulus was developed for experimental purposes. The film was shot from a single angle and depicted a robbery involving two male assailants and one female victim. The film began with two men in a parking lot talking about a money-related problem, after which they disappeared off-screen. A young lady walks up to her car, opens the door and retrieves some items. One of the two men depicted earlier attacks her, pushes her into her car and steals her briefcase. Both men flee the parking lot by their own car. The original movie used in the experiment can be seen here: https://drive.google.com/open?id=1n8QB_Jm0FhJdYmsE4uFUL4LoeATMrniP.

A.3 Instructions

[The text below is distributed to participants before the start of the experiment, and read aloud. Variations according to the treatment are indicated in brackets.]

- Step 1. Thank you for participating in this study. You are about to participate in a study on eyewitness testimony. You will watch a short video, once. You have to watch it as if you were involved in the scene, a bit set back. Then, you will have to respond to questions about the scene. Please, pay attention, I am about to start the video.
- Step 2 [*No oath* conditions]. In that kind of experiment, it is very important that participants know the privacy policy. I am therefore asking you to sign this form. Signing this form is necessary to pursue the experiment.
- Step 2 [*Oath* conditions]. In that kind of experiment, it is very important that participants perform the task as honestly as possible. I am therefore asking you to commit in writing by signing this form. Signing this form is necessary to pursue the experiment.
- Step 3. Please read the instructions carefully, they will enable you to understand properly the experiment. When all the participants have finished reading the instructions, one experimenter will read-aloud the instructions. All the responses you will make will be analyzed anonymously. You will indicate your responses on the computer you are facing. The experiment is divided into two parts: “Part 1” and “Part 2”. The following instructions deal with the Part 1 of the experiment. You will be later informed of specific instructions for the Part 2 of the experiment. At the end of the experiment, you will receive the earning you collected in cash. From now on, please don’t speak out loud. If you have a question, raise your hand, an experimenter will respond to you privately.
- Step 4. During this part, you will be able to accumulate euros by undertaking a series of task. Each of this task will consist of counting how many “1” are present in a grid that contains a sequence of “0”and “1”. Here is an example (that contains 23 “1”):

```
010011011
011110110
011010001
110000101
000110110
```

You are about to be provided with a paper file that contains 12 such grids. Each of these grids is a task to be performed. On your computer screen will appear an input window that will remain active for 20 minutes.

The figure below shows an example of an input window. For each grid, please provide the number of “1” you counted. You may check if your response is correct by clicking on the icon “tester”.

Once you have successfully completed a task, that is, when your answer is correct, it will be automatically validated by the computer (both the icon and the input field will deactivate). You can then focus on the remaining tasks. It is not necessary that you follow the order of the numbered tasks.

IMPORTANT

- If you manage to complete all the tasks in the allocated time (20 minutes), you will earn 12 euros and will be able to pursue the experiment.
 - If you fail to complete all the tasks in the allocated time, you will not be able to pursue the experiment and you will only earn the travel flat rate.
- Step 5. Before entering the lab, you watched a video. You will be asked to testify on that video and therefore to remember best the scene. We will guide your recall by asking you specific questions in a chronological order according to the video. For each question, please choose the correct answer between two alternatives. Your objective is to be as accurate as possible.

[*Control* condition: Your answers will not affect your earnings]. [*Truth* condition: For each incorrect answer, you will lose 0,30 euros]. [*Deception* condition: For each correct answer, you will lose 0,30 euros]. Once you finished the task, you will be provided with the number of correct answers and the amount of your earnings.

A.4 Questionnaire

[Both the correct and the incorrect answers are provided in brackets]

1. When both men (the assailant-to-be and the accomplice) arrive on the parking lot, they arrive: [Correct answer: from the left of the screen; False answer: from the right of the screen]
2. When the assailant-to-be arrives in the field of vision, he: [Correct answer: is already on the phone; False answer: is about to be on the phone]
3. When the assailant-to-be is on the phone, he repeats: [Correct answer: “yeah, yeah”, False answer: “no, no”]
4. The assailant-to-be tells his accomplice: [Correct answer: “we have to move it”, False answer: “we have to get cracking”]
5. The assailant-to-be tells his accomplice: [Correct answer: “we need dough”, False answer: “we need cash”]
6. When both men finish their talk, the accomplice leaves: [Correct answer: toward the right side of the screen, False answer: toward the left side of the screen]

7. The victim arrives on the parking lot from: [Correct answer: the right side of the screen, False answer: the left side of the screen]
8. When the victim arrives, she holds a soon-to-be stolen bag: [Correct answer: in her hand, False answer: on her shoulder]
9. When the victim walks towards her car, one can hear the sound of: [Correct answer: her heels; False answer: her phone]
10. The victim drops a small item on the floor: [Correct answer: before she opened her car's door; False answer: after she opened her car's door"]
11. The victim drops an object in front of: [Correct answer: the car's front wheel; False answer: the car's back wheel"]
12. The victim picks up the item she dropped: [Correct answer: without putting it in her bag; False answer: and put it in her bag"]
13. The victim's car is a: [Correct answer: three-doors car, False answer: five-doors car"]
14. The victim starts to put her stuff [Correct answer: in the front of the car; False answer: in the back of the car]
15. The men arrives to attack the woman: [Correct answer: from the left of the screen; False answer: from the right of the screen]
16. At the time of the attack, relative to the camera, the woman: [Correct answer: has her back turned to the camera; False answer: is from side-on relative to the camera]
17. The assailant pushes the victim: [Correct answer: in the front of the car, False answer: in the back of the car]
18. The assailant steals a bag: [Correct answer: black-colored, False answer: grey-colored]
19. Before getting away from the victim's car, the assailant touches the car's: [Correct answer: door, False answer: roof]
20. Once the assailant away from the victim's car, he waves to his accomplice: [Correct answer: with his right hand; False answer: with his left hand]
21. In order to flee, the assailant get in a: [Correct answer: Peugeot car; False answer: Renault car]
22. The car that picks up the assailant has: [Correct answer: three brake lights; False answer: two brake lights]
23. The assailants' car goes to the end of the car park and: [Correct answer: stops; False answer: turns]
24. After the attack, the victim picks up an item: [Correct answer: on the floor; False answer: on the car's package shelf]
25. After the attack, the victim touches: [Correct answer: the top of her head; False answer: her forehead]
26. After the attack, on the parking lot: [Correct answer: the victim is alone; False answer: somebody arrives at the bottom]

27. At the bottom of the parking lot, one can see: [Correct answer: street lights; False answer: city garbage cans]
28. The parking spaces are demarcated by: [Correct answer: three-digit numbers; False answer: two-digit numbers]



A.5 Socio-demographic questionnaire

Please fill up the questionnaire below. Your responses will remain anonymous, please respond honestly [for closed-ended questions, the possible answers appear in brackets].

1. Year of birth
2. Gender [man/woman]
3. Nationality
4. Do you live as a couple? [yes/no]
5. Are you a student?
6. Study major
7. Year of study
8. Have you ever participated in an experiment in economy? [yes/no]
9. How many siblings to you have (half-brother/sister included)?
10. Your rank among your siblings
11. Do you practice any sport? [yes/no]
 - if yes Sport type: team sport/individual sport
 - if yes Do you participate in tournaments?
 - if yes To which level?
12. Does religion constitute an important part of your life? [very little/little/somewhat/very]
13. In general, would you say you are: [very little/little/somewhat/very] believer
14. What is your religion?
15. Please refer to the following scale and indicate to what extent you are confident with your responses relative to the video [Likert scale from 1 to 7]
16. Please refer to the following scale and indicate to what extent you are currently happy [Likert scale from 1 to 7]
17. Please refer to the following scale and indicate to what extent you have been honest during this experiment [Likert scale from 1 to 7]
18. Please refer to the following scale and indicate to what extent other participants have been honest during this experiment [Likert scale from 1 to 7]
19. Please indicate the city in which you currently live

- 20. What language do you speak with your family?
- 21. What is your native language?
- 22. Had you ever seen this video? [yes/no]

A.6 Oath form

 <p>FACULTÉ D'ÉCONOMIE université montpellier 1</p>	 <p>UM1 Université Montpellier 1</p>
<div style="border: 1px solid black; padding: 5px; display: inline-block;">SOLEMN OATH</div>	
<p>I undersigned swear upon my honor that, during the whole experiment, I will:</p>	
<p>Tell the truth and always provide honest answers.</p>	
Montpellier,	Signature.....
<hr/> <p>Avenue Raymond DUGRAND - Espace Richter—C.S. 79606—34960 —Montpellier cedex 2—Tel. 04.67.15.84.50—Fax. 04.67.15.84.86</p>	

B Extended Data

B.1 Detailed data on accuracy by item and by condition

Figure 3 provides detailed evidence of the average accuracy of participants in the control condition across items. Table 2 provides descriptive statistics on observed accuracy in each condition.

Figure 3: Proportion of correct responses by item in the *Truth* condition

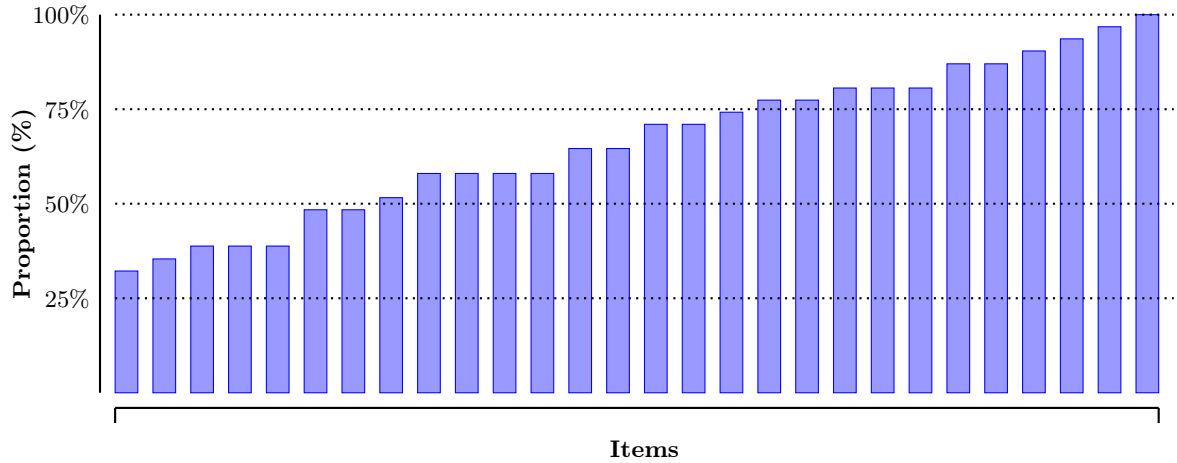
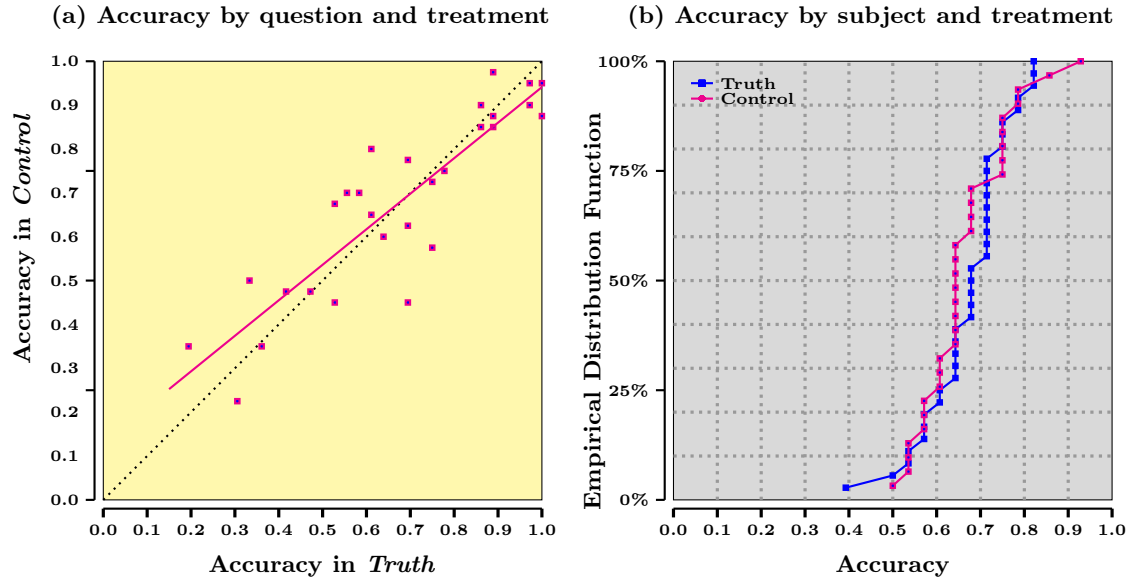


Table 2: Descriptive statistics on observed accuracy (in%) in each condition

<i>No oath</i>					<i>Oath</i>				
	Min	Max	Mean	sd.		Min	Max	Mean	sd.
<i>Control</i>	50.00	92.86	66.47	9.83	<i>Control under oath</i>	46.43	89.29	66.96	10.06
<i>Truth</i>	39.29	82.14	67.26	9.49	<i>Truth under oath</i>	42.86	82.14	67.77	9.41
<i>Deception</i>	17.86	71.43	44.83	15.28	<i>Deception under oath</i>	14.29	82.14	58.69	19.19

B.2 Comparison between the *Control* and the *Truth* conditions without oath

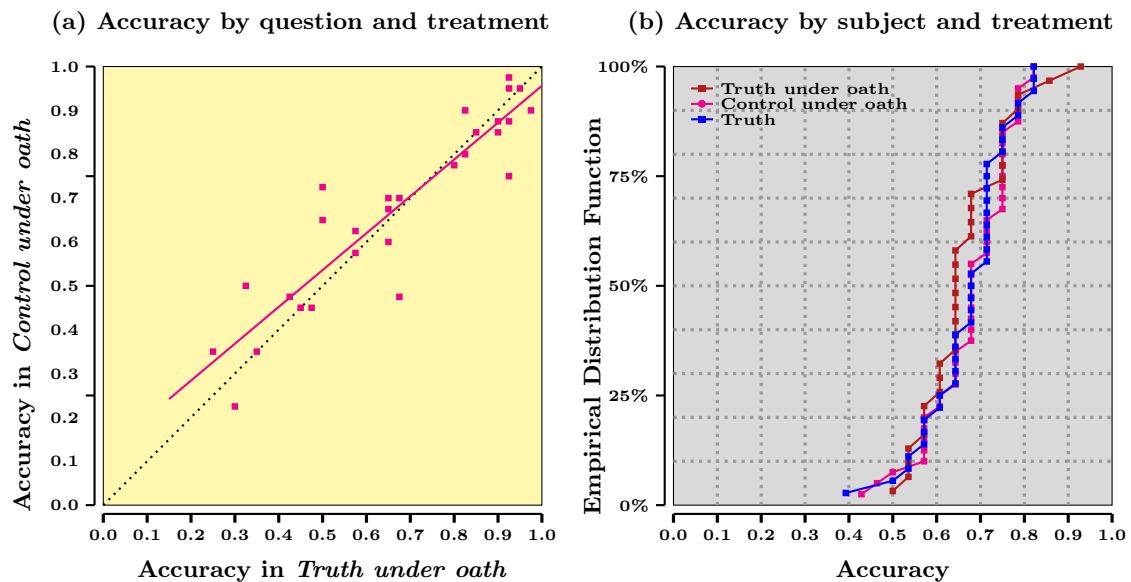
The figure below shows that the EDF of the *Truth* condition closely overlaps that of the *Control* condition. A bootstrap Kolmogorov-Smirnov test that provides correct coverage even when the distributions being compared are not entirely continuous (Sekhon, 2011) indicates that the two EDF are not significantly different ($p = .278$).



Note. Distribution of memory performance with and without monetary incentives for truth-telling when participants are not under oath. **Panel (a)** plots the average accuracy observed on each item in *truth* (without oath, on the *x*-axis) against the average accuracy observed in the *Control* condition (without oath), with no monetary incentives (on the *y*-axis). The regression line is on the 45 degree line and indicates no difference at the aggregate level. **Panel (b)** presents the EDF of individual accuracy for the same two treatments.

B.3 Comparison between the *Control* and the *Truth* conditions under oath

The figure below shows that the EDF of the *Truth under oath* condition closely overlaps that of the *Control under oath*. A bootstrap Kolmogorov-Smirnov test that provides correct coverage even when the distributions being compared are not entirely continuous (Sekhon, 2011) indicates that the two EDF are not significantly different ($p = .871$). Truth with and without oath closely overlap ($p = .943$).



Note. Distribution of memory performance with and without monetary incentives for truth-telling when participants are under oath. **Panel (a)** plots the average accuracy observed on each item in *truth under oath* (on the *x*-axis) against the average accuracy observed in the *control under oath* condition, without monetary incentives (on the *y*-axis). The regression line is on the 45 degree line and indicates no difference at the aggregate level. **Panel (b)** presents the EDF of individual accuracy for the same two treatments.

B.4 First order stochastic dominance tests

First order dominance tests provided in the text are carried out using a one-tailed bootstrap version of the Kolmogorov-Smirnov test on the EDF of individual accuracy. The test allows for ties, which provides correct coverage even when the distributions being compared are not entirely continuous (Sekhon, 2011). FOD in the Table below stands for “First Order Dominates”.

	Comparison		<i>p</i> -value
<i>Truth</i> (without oath)	FOD	<i>Deception</i> (without oath)	.001
<i>Deception under oath</i>	FOD	<i>Deception</i> (without oath)	.024
<i>Truth</i> (without oath)	FOD	<i>Deception under oath</i>	.165
<i>Truth under oath</i>	FOD	<i>Truth</i> (without oath)	.943
<i>Control under oath</i>	FOD	<i>Control</i> (without oath)	.936

B.5 OLS Regression Analysis

Dependent variable: Number of correct responses		
	Coefficient	<i>p</i> -value
Intercept	18.800	<.001
Male	-0.021	.663
Participated to previous experiments	0.886	.141
Age	-0.045	.265
Atheist (Ref. Christian and others)	0.427	.487
Muslim (Ref. Christian and others)	-0.017	.978
Strong believer (Religion)	-0.245	.724
Number of siblings	0.127	.273
Live in couple	0.674	.978
Condition		
<i>Truth</i>	0.127	.833
<i>Deception</i>	-6.270	<.001
<i>Truth under oath</i>	0.026	.965
<i>Deception under oath</i>	3.917	<.001

Note. OLS regression Estimates ($N = 222$). The dependent variable (number of correct responses per respondent) is regressed on treatment dummies and individual characteristics. The referent individual is female, Christian, single and did not participate to previous experiments. Treatment dummies are defined so as to mimic the results presented in the paper — using the *Control* condition with no monetary incentive as a reference. The coefficient on *Truth* shows that monetary incentives for performance have no effect on accuracy ($p = .833$). The coefficient on *Deception* shows that honesty costs significantly decrease accuracy ($p < .001$). The coefficient on *Deception under oath* shows that the oath only has an effect in the presence of honesty costs ($p < .001$). Adjusted $R^2 = .286$. F-statistic: 8.393 on 12 and 209 DF, $p < .001$.

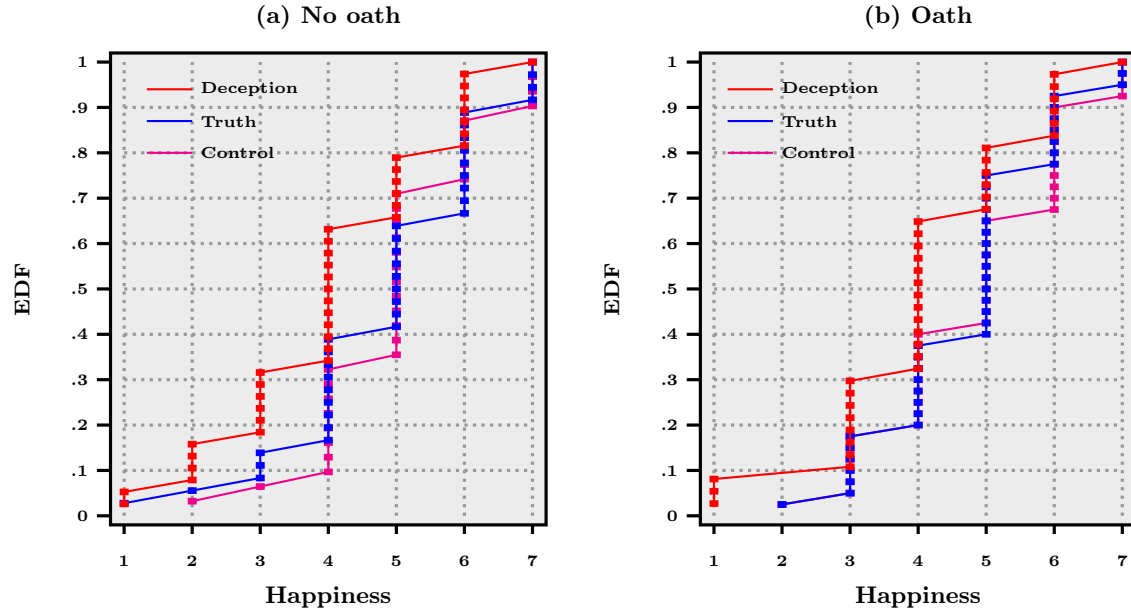
B.6 Poisson Regression Analysis

Dependent variable: Number of correct responses		
	Coefficient	<i>p</i> -value
Intercept	2.932	<.001
Male	-0.013	.685
Participated to previous experiments	0.052	.194
Age	-0.002	.328
Atheist (Ref. Christian and others)	0.023	.571
Muslim (Ref. Christian and others)	-0.003	.943
Strong believer (Religion)	-0.014	.753
Number of siblings	0.080	.307
Live in couple	0.037	.351
Condition		
<i>Truth</i>	0.006	.872
<i>Deception</i>	-0.407	<.001
<i>Truth under oath</i>	0.0007	.987
<i>Deception under oath</i>	0.273	<.001

Note. Poisson regression estimates ($N = 222$). The dependent variable (number of correct responses per respondent) is regressed on treatment dummies and individual characteristics. The referent individual is female, Christian, single and did not participate to previous experiments. Treatment dummies are defined so as to mimic the results presented in the paper — using the *Control* condition with no monetary incentive as a reference. The coefficient on *Truth* shows that monetary incentives for performance have no effect on accuracy ($p = .872$). The coefficient on *Deception* shows that honesty costs significantly decrease accuracy ($p < .001$). The coefficient on *Deception under oath* shows that the oath only has an effect in the presence of honesty costs ($p < .001$).

B.7 Subjective well-being

At the end of the experiment, subjects were asked how happy they feel at the moment (see MacKerron, 2012, for a review and the Appendix, Section A.5 for the question). The figure below reports the Cumulative Distribution Functions (CDF) of the happiness measure. We report the distribution of this variable separately for subjects in the control, truth and deception conditions in Panel (a), and subjects in the oath version of the same conditions in Panel (b).



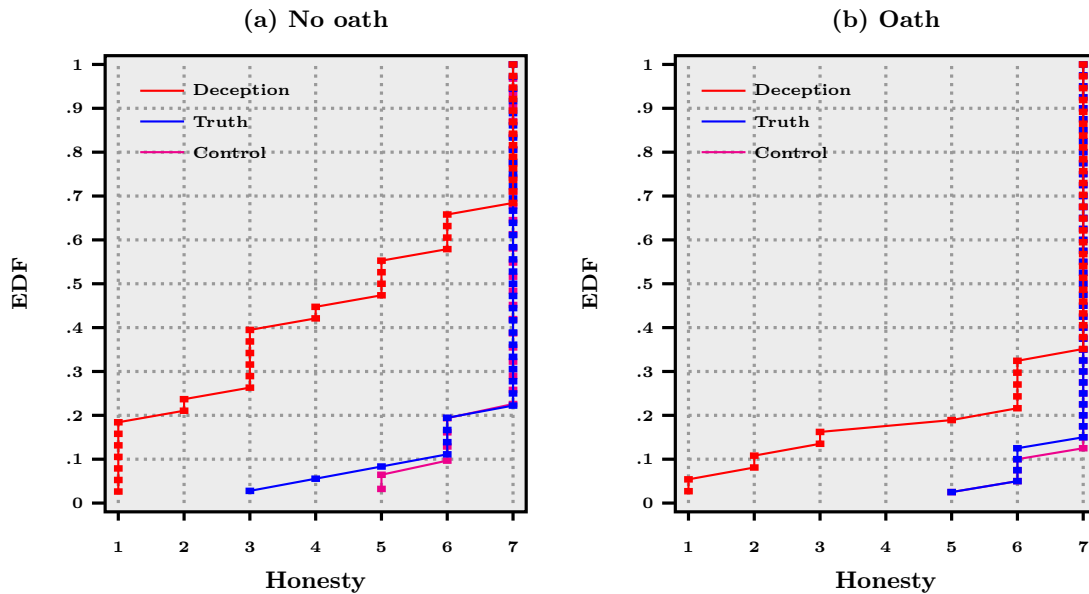
Note. The figures present subjective well-being at the individual level and show the Empirical Distribution Functions (EDF, i.e., the empirical cumulative distribution) of subjective well-being by experimental condition with and without the oath. **Panel (a)** presents the EDF of participants without oath whereas **Panel (b)** presents the EDF of participants under oath.

The results are three-fold. First, monetary incentives that encourage performance have no effect on subjective well being: the CDF in the *Truth* condition is similar to that in the *Control* ($p = .875$, KS bootstrap test). Second, the implementation of monetary incentives that encourage deception generates a small but significant decrease in subjective well-being. The CDF in the *Deception* condition without oath is first order dominated by the CDFs in the *baseline* ($p = .007$) and in the *Truth* ($p = .042$) conditions. Third, we find that the oath has no detrimental effect on well-being in the *Control* and *Truth* conditions ($p = .356$ and $p = .302$ respectively, KS bootstrap tests) nor does it restore the loss of well-being induced by monetary incentives in the *Deception* condition ($p = .755$, KS bootstrap test). This may suggest that the increase in honesty induced by the oath goes through a different channel than that induced by monetary incentives.

B.8 Self-reported honesty

To complement our behavioral results, the figure below reports the Cumulative Distribution Functions (CDF) of the self-reported measure of honesty collected at the end of the experiment (see the Appendix, Section A.5, question 17). We report the distribution of this variable separately for subjects in the control, truth and deception conditions in Panel (a), and subjects in the oath version of the same conditions in

Panel (b). The gap observed between the *Deception* and the *Truth* conditions (without oath) measures the degree of deception. The gap between *Deception under oath* and *truth* measures the extent of perjuries.



Note. The figures present self-reported honesty at the individual level and show the Empirical Distribution Functions (EDF, i.e., the empirical cumulative distribution) of self-reported honesty by experimental condition with and without the oath. **Panel (a)** presents the EDF of participants without oath whereas **Panel (b)** presents the EDF of participants under oath.

The comparison between the *Control* and *Truth* conditions provides an empirical benchmark of the validity of the self-reported measure of honesty. The CDF in the *Control* and *Truth* conditions first order dominates the CDF in *Deception* ($p < .001$ and $p < .001$, bootstrap KS test). This indicates that subjects in the *lie* condition are willingly reporting that they weren't honest in their testimony. Similarly, the CDF observed in the *Truth* condition closely mimic our behavioral findings. Self-reported honesty is similar in the *Control* and *Truth* conditions. The CDF indicates that a large proportion of subjects report that they were fully honest in their testimony (80.6% and 80.5% respectively). This is also the case when we consider the control and truth conditions under oath, although with a small increase in the proportion of subjects reporting full honesty (90.0% and 87.5% respectively).

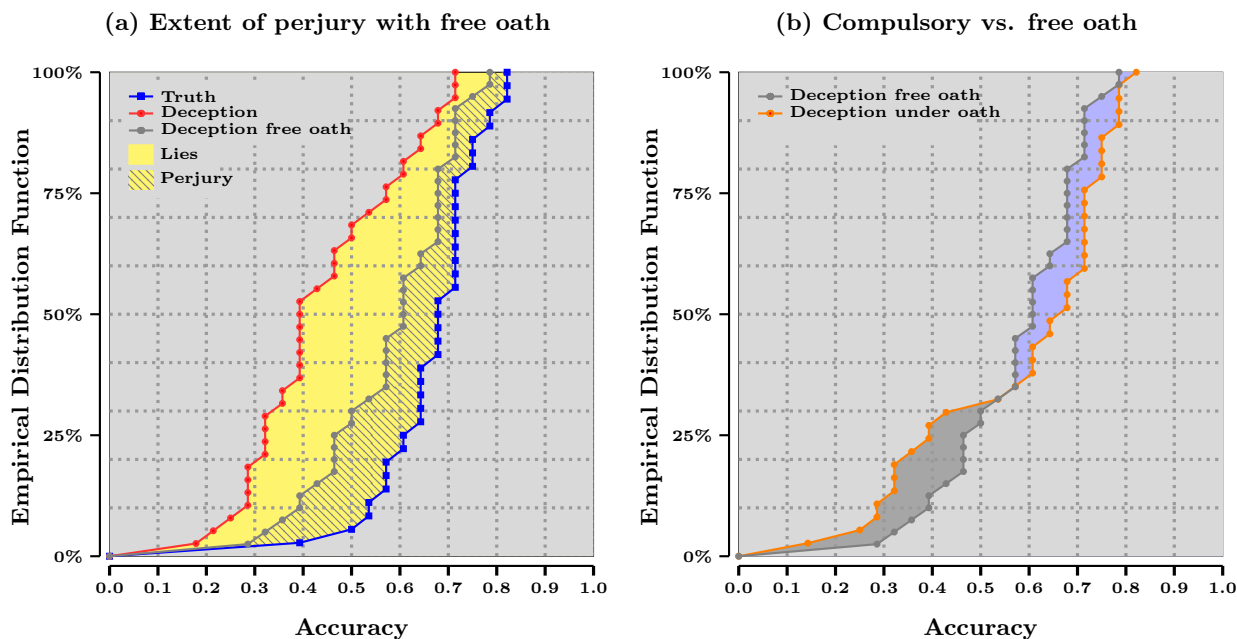
We now turn to the effect of the oath, by comparing the distributions between Panel (a) and (b). In both the *Deception* and the *Truth* conditions, the EDF of self-reported honesty moves to the right under oath. It first order dominates the one observed with no oath in the *Deception* condition ($p = .008$, bootstrap KS test). Subjects under oath declare themselves more honest than subjects without the oath do. Interestingly, around two-third of subjects under oath in the *Deception* condition under oath (67.6%) report that they were fully honest.

B.9 Mandatory oath and free compliance

Herein, we provide a robustness check of the effect of the oath on participants behavior by departing from the ecology of eyewitness testimony. We relax the mandatory nature of the oath, typical to its application in the court-room, and instead consider free compliance.

This *Free oath* condition replicates the *Oath* treatment except that subjects are free to sign or not the oath. Before entering the laboratory, participants enter one by one a separate room close to the laboratory

Figure 4: Accuracy by treatment at the aggregate and individual levels



Note. Panel (a) Empirical Distribution Functions (EDF) of the proportion of recognized items computed at the subject level in the *Deception free oath* condition (in grey) along with a reminder of the distributions observed in the *Truth* (blue) and *Deception* (red, both with no oath) conditions. The gap observed between the *Deception* and the *Truth* conditions measures the degree of deception. The gap between *Deception free oath* and *Truth* measures the extent of perjuries — i.e., lying behavior under a free oath. Since we find that the EDF of the *Truth under oath* condition is very similar to that of the *Truth* condition, we omit it from the Figure to ease the reading (see Appendix B, Sections B.2 and B.3). **Panel (b)** Comparison of the EDF of the proportion of recognized items between the *Deception free oath* and the *Deception under oath* conditions.

where they receive the oath form. Participants are told that signing the oath is not mandatory and that their earnings in the experiment are not conditional on signing or not the oath (see, e.g., Jacquemet et al., 2013, 2019, for an experiment using the exact same procedure). We carried out two additional sessions of this *Deception free oath* condition. Each session involved 20 subjects. All subjects but one took an oath. This high compliance rate (equal to 97.5%) rules out any selection effect in the behavioral response to the oath. In the following, we apply an intention-to-treat strategy and keep all 40 subjects to conduct the statistical analysis.

This new data confirms the overall effect of the oath on deception. Mean accuracy in the *Deception free oath* condition is 58.5%. The increase in truth-telling as compared to *Deception* (with no oath) is highly significant ($p < .001$, two-sample t-test), and of the same order of magnitude as the one observed in the *Deception under oath* (in which signing the oath is mandatory) described in the main text (58.7%, $p = .957$, two-sample t-test).

Figure 4 complements the analysis with individual data. Panel (a) compares the EDF of the proportion of recognized items observed in the *Deception free oath* to the *Deception* (with no oath) and the *Truth* (with no oath) conditions. This figure shows that the free oath induces the same change in behavior as the one observed under a mandatory oath: the whole EDF of recognized items is moved to the right when compared to that observed in the *Deception* condition ($p < .001$, KS bootstrap test). This indicates that

all subjects respond to the free oath. Second, as compared to the behavioral benchmark provided by the *Truth* condition, the free oath fails to restore full honesty for subjects at the top of the distribution.

Panel (b) compares the behavior elicited in the *Deception Free Oath* and *Deception Under Oath* conditions. While the two distributions are similar overall—confirming the results observed for average accuracy—they intersect around the first tercile of the distribution. This suggests that a freely taken oath is more effective at committing participants who respond most strongly to monetary incentives under a mandatory oath, resulting in particularly low levels of accuracy. However, this benefit appears to come at a cost: participants at the upper end of the distribution exhibit slightly lower accuracy under a free oath compared to a mandatory oath. This finding aligns with long-standing insights from social psychology, which emphasize that volitional or freely chosen acts lead to stronger commitment (Cialdini, 2007; Joule and Beauvois, 2010). What is novel here, however, is the identification of a social trade-off between the two procedures.

B.10 Shares of lying aversion and the oath on deceptive behavior

The average performance of participants in the *Truth* condition (without oath) is 67.2%. If the same respondents had instead faced a honesty cost and fully responded to these incentives, we would have observed an average performance equal to $100 - 67.2 = 32.7\%$. This means that, given the memory performance observed in the *Truth* condition, the range for deceptive behavior is $67.2 - 32.7 = 34.6\%$. This gives us a theoretical lower bound for performance in the *Deception* condition. We observe that participants do not fully respond to the honesty cost in this condition as the mean accuracy is equal to 44.8%. The share of potential deceptive behavior that is addressed by intrinsic motivation, i.e., the intrinsic willingness to tell the truth, is therefore: $\frac{44.8-32.7}{34.6} = 35.0\%$.

We can replicate the same reasoning on the oath conditions. The theoretical lower bound for performance is obtained by reversing the answer of participants in the *Truth under oath* condition. The range for deceptive behavior in the oath conditions is $67.8 - 32.2 = 35.6\%$. This is not significantly different from the range calculated in the no oath conditions, since memory performances are very similar between the two *Truth* conditions (given that they are also not different from the *Control* conditions with and without the oath, we can use either of these as a benchmark to measure the true memory performance; our results are only minimally affected by the choice: the results are available upon request). In the *Deception under oath* condition, mean performance is 58.7%. This means that the combined effect of intrinsic motivation and the oath is: $\frac{58.7-32.2}{35.6} = 74.4\%$.

The specific effect of the oath on deceptive behavior is obtained by deducing the share of what is achieved by intrinsic motivation alone in the no oath conditions: $74.4 - 35.0 = 39.4\%$. The remaining share, $100 - 74.4 = 25.6\%$, is the magnitude of perjury, i.e., the remaining deceptive behavior under oath.