Big and Small Lies[*]

Diogo Geraldes[†] Franziska Heinicke[‡] Duk Gyoo Kim[§]

March 18, 2020

Abstract

Lying often involves many decisions that can yield big or small benefits. We investigate how asymmetry in the size of a lie affects lying behavior within a two-dimensional context, by conducting a laboratory experiment with three treatments. In the *Big and Small Lie* (BSL) treatment, participants can simultaneously tell a big and a small lie. In the *Big Lie* (BL) treatment, participants can misreport only a high-stakes outcome because a computer exogenously determines the low-stakes outcome. The *Small Lie* (SL) treatment reverses the BL treatment. We find that participants in the BSL treatment lie in both the low- and the high-stakes, but they lie more about the outcome of the high-stakes. Interestingly, we also show that participants behavior—either lying or truth-telling—is consistent between the big lie and small lie dimensions. Moreover, the level of lying on both the high-stakes (BSL vs. BL) and the low-stakes outcome (BSL vs. SL) does not differ between treatments. It means that the second outcome, regardless of being an exogenous prize or endogenously determined, does not affect participants' lying behavior regarding the first outcome. Notably, however, we find that repeatedly being lucky (unlucky) on a high-stakes prize leads to less (more) lying on the report of a low-stakes outcome.

JEL Classification: C91, D03

Keywords: Laboratory experiment, Lying

[*]We thank Collin Raymond and the participants of the HeiKaMaX Workshop and the ESA European Meeting 2019 for invaluable comments and suggestions. Duk Gyoo Kim gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)–Project-ID 139943784–SFB 884.

[†]Utrecht University School of Economics, 3584 EC Utrecht, the Netherlands. Email: diogogeraldes@gmail.com

[‡]Department of Economics, University of Mannheim, Germany. Email: f.heinicke@uni-mannheim.de

[§]Department of Economics, University of Mannheim, Germany. Email: d.kim@uni-mannheim.de

1 Introduction

Lying behavior is pervasive in social, political, and economic life. Nevertheless, the lies that people can tell are not all made equal but rather differ considerably regarding the consequences they have. Some lies do not cause much harm relative to the counterfactual situation under truth-telling. Other lies, however, have considerable consequences because they create a significant shift relative to the situation that would have occurred under truth-telling. For instance, whereas deceiving an employer regarding oversleeping is likely innocuous to a corporation, obtaining a job by misstating items in a resume will, in all likelihood, jeopardize it. In this paper, we investigate whether (and how) the size of a lie affects individual lying behavior.

Concerned with the adverse effect of self-serving dishonesty in economically relevant settings, a recent growing literature in experimental economics has been attempting to understand the determinants of lying behavior.[1] Traditionally, economists have considered lying to be inevitable as long as the material gains from a lie outweigh the risk and consequences of being detected (Lewicki, 1983). However, the experimental literature on lying behavior has come to a different conclusion. Two recent meta-studies (Abeler et al., 2019; Gerlach et al., 2019) indeed show the existence of liars. However, taking the impressively vast work conducted in the last few years altogether, we can also clearly identify a considerable proportion of people that hold preferences for honesty. In particular, many people not only fear the material consequences of lying but also suffer moral costs from lying. Even though this recent literature has largely enriched our understanding of lying behavior, hitherto we do not know much regarding the interaction of two lies that vary by size. This paper aims to study the latter element by examining people's lying behavior when allowed to provide lies with asymmetric consequences simultaneously.

To test the impact of the size of a lie on lying behavior, we conduct a laboratory experiment to compare big and small lies. To this end, we vary the size of a lie not only regarding the *stake component* (i.e., the payoff associated with each outcome) but also regarding the *outcome component*. Gneezy et al. (2018) define the outcome component of a lie as the distance between the true and the reported outcome. Thus, according to this component, lies are bigger the further away they are from the true outcome. By manipulating both the stake and the outcome components of a lie, we are confident to

[1] Abeler et al. (2014) and Arbel et al. (2014) investigate what individual characteristics shape the costs of lying. Charness et al. (2019) show how the moral cost of lying may prevent lying in a loss frame. Cohn and Maréchal (2018), Dai et al. (2018), and Potters and Stoop (2016) examine whether and to what extent the laboratory measure of lying predicts misconduct in real situations. Hurkens and Kartik (2009), Houser et al. (2012), and Cojoc and Stoian (2014) study the relationship between social preferences and lying behaviors.

induce participants to perceive the difference in size.

Importantly, we evaluate the interaction between big and small lies in a simultaneous two-decision setting. More specifically, in our experiment, participants toss a coin and roll a dice and receive a payoff based on the reports of both outcomes. The coin toss represents the *big lie* not only because we associate it with higher-stakes than the dice but also because lies on the coin are not scalable. That is, participants who decide to lie on the coin accept a larger deviation relative to the outcome under truthful reporting than is possible on the dice. Moreover, we compare behavior in this main treatment—which we call the *Big and Small Lie* (BLS) treatment—with two control treatments. In the *Big Lie* (BL) treatment, participants report only the coin while the dice is determined exogenously. In the *Small Lie* (SL) treatment, participants report only the dice while the coin is determined exogenously.

Often, lying does not involve a single decision but rather multiple decisions that asymmetrically affect one final outcome. In other words, the joint occurrence of big and small lying opportunities is common in real-world settings. A compelling example is tax declarations because people have to report private information regarding several items, and the combination of all self-reports will ultimately determine the tax return. Thus, if people intend to adjust the final outcome in their favor, they can misreport one, all, or just some of the individual items. However, misreports are not necessarily equal in size. That is, the consequence of a particular lie on the tax return depends on the item that is misreported. For example, misreporting the capital gains from assets held in a foreign country may lead to a considerable effect on the final tax payment, while overstating miscellaneous expenses will have a minor effect.

In the literature, the effect of the size of a lie on lying behavior has thus far been studied by varying the stake component of a lie, i.e., by varying the magnitude of the marginal payoff of lying. The two aforementioned meta-studies also included stake size as an independent variable to organize the literature on lying behavior. While Abeler et al. (2019) find no effect of stake size on lying behavior across 90 studies, Gerlach et al. (2019) report a positive effect of the maximum gain on lying but not in studies that compare different stake sizes directly. Importantly, we should highlight that previous studies that manipulate the size of a lie compare different stakes solely using a between-subjects design, i.e., in each treatment, a participant makes only a single-decision for a specific stakes level.

This paper contributes to a growing body of research on lying behavior by exploring the impact of the size of a lie in a two-dimensional setting. In particular, we extend the understanding of how the size of a lie affects lying behavior by analyzing the interaction of jointly told big and small lies while varying the size of a lie considering both the outcome and the payoff components. Our fundamental aim is to assess how people behave when they can simultaneously tell a big and a small lie, and both lies contribute jointly to an overall payoff. We deem this approach a more direct and realistic comparison of big and small lies. Many economically relevant real-world settings present simultaneous lying opportunities. Thus, for deriving accurate policy implications, it is of utmost importance to examine settings in the lab that preserve vital features of real-world settings.

Further, in this study, we elicit simultaneous decision-making in a two-dimensional context. Thus, we are not only interested in how size affects lying but also how two simultaneous lies interact with each other. Thereby, we are also contributing to understanding lying behavior in a multi-dimensional setting.

Our results show that lying behavior is complementary when both big and small lying opportunities are available. That is, participants who lie in the coin are also more likely to lie in the dice, and vice-versa. Moreover, in absolute terms, we observe more liars in dice than in the coin. However, taking into account the size of participants' lies, we observe significantly more lying in the coin than in the dice.

Interestingly, average reports on the coin do not differ between the main treatment and the BL control treatment. Nor do average reports on the dice differ between the main treatment and the SL control treatment. These findings indicate that lying is complementary when big and small lies can be reported jointly, but we do not observe more big-stake (small-stake) lies when only one type of lying opportunity is available. In short, collectively, our results support that participants consider the two lying opportunities largely as independent and behave consistently for both of them.

Besides analyzing big and small lies that are reported jointly, we also test the effect of observing a (un)favorable outcome on lying behavior. We find that the observation of an exogenous low-stakes outcome does not affect telling a big lie. Notably, however, we find that repeatedly observing a positive high-stakes exogenous outcome leads to decreased lying regarding the small lie.

The remainder of this paper is organized as follows. First, we provide an overview of the relevant literature and state the predictions for our research questions. Section 3 outlines the experimental design and procedure. In Section 4, we report the results of the experiment. Section 5 concludes.

2 Literature

Gneezy et al. (2018) divide the size of a lie into three different components. First, the size of a lie is determined by the payoff that can be gained from it, i.e., the *stake size* component of the lying task. According to this component, the size of a lie can be varied by increasing or decreasing the monetary payoff of different reports. Second, the size of a lie has an *outcome* component. While the stake size component describes how much can be gained from lying in absolute terms, the outcome component takes into account what the true state of the world is. A lie is greater in size when a false report differs further from the truth. Finally, the size of a lie also depends on how blatant a lie is to a possible audience, which is the *self-image concerns* component of a lie. People do not want to be seen as liars, and a report that makes them look like a liar makes it challenging to maintain a positive self-image. In this component, a bigger lie is a lie that is more likely to reveal a person as a liar. In the following, we discuss evidence on the effect of the first two components, which we manipulate in our experimental design.

The question of stake size has been discussed in the lying literature from the start. In the experimental study of Mazar et al. (2008), participants can get a higher payment by overreporting their performance of a real-effort task. To analyze the effect of the size of payoffs on lying, the authors vary the payoff incentives of the real-effort task. They find no significant difference in lying when the stake sizes quadrupled. Similarly, Fischbacher and Föllmi-Heusi (2013) conduct a low-stakes baseline treatment and a high-stakes treatment where the stakes tripled as compared to the baseline. They find no significant effect of stake size on lying either.

Two recent meta-studies consider the issue of stake size in the lying literature systematically. Taking into account 72 studies in which participants could misreport a randomly generated outcome, Abeler et al. (2019) find that stake size affects neither the average report nor the patterns of lying. For all stake sizes, outcomes at the lower end of the distribution are under-reported, and high outcomes are over-reported. Gerlach et al. (2019) compiled data from 565 experiments, including lying about real-effort tasks and random outcomes, as well as lying in sender-receiver games. They find that a higher maximum gain increases lying in studies using a coin toss to generate a random outcome, but this effect is not present when reducing the analysis to those studies which compare different stake sizes directly. For all other individual lying tasks, stake size has no effect on reports. This said, we should highlight that even though stakes are varied in these studies, each treatment corresponds to a single stake level.

In line with the empirical evidence that lying is fairly unresponsive to stake size, theoretical arguments on the effect of stake size have been made for both directions of the effect. Arguably, stake size affects lying through two channels. First, a larger reward increases the marginal benefit of lying. The literature largely finds lying to be a trade-off between the monetary benefits from a lie and the psychological costs of lying, which include direct moral costs for breaking a moral norm and reputational costs for possibly being considered a liar by others (Gneezy et al., 2018; Abeler et al., 2019; Dufwenberg, 2018). Therefore, increasing the benefit of a lie while keeping all else equal will lead to more lying (Mazar et al., 2008; Hilbig and Thielmann, 2017).

The costs of lying are the second channel through which higher stakes affect lying. People try to maintain a positive self-image and are more likely to tell a lie if they can justify this lie (Shalvi et al., 2015). Mazar et al. (2008) interpret their result on the neutral effect of stakes as an indication that justification varies with the amount of money that is claimed. Specifically, the costs of lying increase for higher stakes due to the lack of justification. Also, telling a high-stakes lie makes it more likely to be seen as a liar by others. Kajackaite and Gneezy (2017) argue that the reputational cost of lying, incurred by the fear of being found out as a liar, is increasing with the size of a lie. Specifically, they show that participants only react to higher stakes with higher levels of lying in a treatment where the probability of detection is zero. Thus, stake sizes might not only affect the size of a lie directly but also indirectly through the likelihood dimension. These two opposing effects can explain the neutral effect of stake sizes on lying.

The evidence provided by Hilbig and Thielmann (2017) further suggests that the effect is confounded by the fact that stakes affect different types of liars differently. While people who are always lying or always answering truthfully are unaffected by stakes, people who tell partial lies can be divided into two groups. The first group is "corruptible" by high-stakes and reacts by lying more, while the second group is only willing to tell small lies and is becoming more honest under high-stakes.

The size of a lie due to the outcome dimension has received less attention in the literature. Gerlach et al. (2019) consider dice roll and coin flip experiments separately in their meta-study. They argue that with a binary task, people have no option to lie partially. If they decide to lie, they have to do so to the full extent. Thus, in a task with more possible outcomes, people can tell smaller lies in the sense that they are closer to the true state. Gerlach et al. (2019) find that coin and dice tasks do not differ with regard to the average level of lying but that more people lie in dice tasks than in coin tasks. The interpretation of this finding is that some people who tell small (partial) lies in dice tasks would not have lied in a coin task.

While the literature on the effect of multiple lies is still sparse, there is extensive literature on spillovers between different moral behaviors (Dolan and Galizzi, 2015; Blanken et al., 2015; Merritt et al., 2010). This literature has shown that in some situations, people balance an immoral act with a moral act to maintain a positive self-image. Moral spillovers have also been observed with regard to lying. Ploner and Regner (2013) show that people are more generous in a dictator game if they lied in a preceding lying task, and Gneezy et al. (2012) find people to lie more after having contributed to charity. However, Gneezy et al. (2012) further find that moral licensing only occurs if the generous act is not subtracted from the participant's payoff and, thus, is not costly. A costly generous act leads to moral consistency, i.e., was followed up with more truth-telling.

The closest work to our study is Chowdhury et al. (2018), Geraldes et al. (2019) and Barron (2019). Chowdhury et al. (2018) experimentally investigate high- and low-stakes lies, but in sequential order. That is, their focus is on isolated big and small lies for which the context might be very different. More specifically, they study the effect of a participant knowing (or not) about a follow-up second-round low-stakes lie on a first-round high-stakes lying opportunity. They show that people lie more in the first-round high-stakes opportunity than in the second-round low-stakes opportunity only when aware of the second-round. In other words, high-stakes lying is increased if participants could plan ahead. Using a dice roll task with two dice, Geraldes et al. (2019) investigate individual lying behavior in a two-dimensional context to test whether multi-dimensional decision-making affects lying behavior. Participants' decision-making in their experiment is also simultaneous, but the stake and outcome components of a lie are kept constant between dimensions. They find that participants over-report significantly more on the lower outcome dice than on the high outcome dice. Barron (2019) also asks participants to report two dice but the two reports differ in the size of the lie along the payoff dimension, i.e. there is a high stakes and a low stakes dice. The author finds that compared to a uniform distribution people over-report on the high stakes dice but under-report on the low stakes dice.

According to the two latter papers, we should expect people not to lie equally when given the opportunity to simultaneously tell a big- and low-stakes lie. In both Geraldes et al. (2019) and in Chowdhury et al. (2018), participants showed more extreme lying for the lying option on which they had to gain more. Further, the analysis by Gerlach et al. (2019) suggests that lying about the coin toss, which is not scalable and forces larger lies in the outcome dimension, is sensitive to stakes while lying about the outcome of a dice roll is not. Therefore, we not only predict an uneven spread of big and small lies but specifically, more lying on the big lying option. Also, if moral spillovers are present between the two lying decisions, we would further expect that big lies are more willingly told when paired with the decision not to tell a small lie than when told in isolation.

3 Experimental design

3.1 Treatments

We design laboratory experiments to observe how participants behave when two misreporting opportunities differ in size. In the *Big and Small Lie* (BSL) treatment, participants are asked to toss a coin, roll a dice, and self-report the outcomes separately on the computer interface. This elicitation is repeated over ten rounds.

A participant's report, (C,D), where $C \in \{head, tail\}$ and $D \in \{1,2,\ldots,6\}$, determines the points a participant earns in the following way:

$$\begin{cases} 15+D & \text{if } C = head \\ 7+D & \text{if } C = tail \end{cases}$$

At the end of the experiment, one round was selected randomly for actual payment (Azrieli et al., 2018). The conversion rate into Euros was 1 point = 0.50 EUR. Since the marginal benefit of lying on the outcome of the coin toss is eight points larger than that of the dice roll, the coin toss is associated with a big lie. More specifically, although the small lie is scalable, the maximum benefit from the small lie is 5 (reporting 6 when the outcome of the dice is 1), which is strictly smaller than the marginal benefit of the big lie (reporting heads when the outcome of the coin is tails). This treatment serves to capture the realm of real-life settings that involve big and small lies.

In the *Big Lie* (BL) treatment, participants are asked only to toss a coin and self-report its outcome and receive a low-stakes prize based on the outcome of a dice roll, which is exogenously determined by the server computer. This elicitation is also repeated ten times. The *Small Lie* (SL) treatment is the reverse of the BL treatment. That is, the participants self-report the outcome of a dice roll, and receive a high-stakes prize based on the outcome of a coin toss, which is exogenously determined by the server computer. These two treatments serve to examine whether, and to what extent, an exogenous random event (relative to a self-reported random event) affects lying behavior. Table 1 summarizes the key differences in the three treatments.[2]

3.2 Procedures

The experimental sessions were conducted in English at the Mannheim Laboratory for Experimental Economics (mLab) of the University of Mannheim. The participants

[2] The full experimental instructions for the BSL treatment are included in the Appendix.

Table 1: Experimental Treatments

	Coin ($\Delta_C = 8$)	Dice ($\Delta_D = 1$)
BSL	self-report	self-report
BL	self-report	exogenous
SL	exogenous	self-report

Δ_I is a marginal benefit of misreporting the outcome of $I \in \{Coin, Dice\}$.

were drawn from the mLab subject pool. Four sessions were conducted for each treatment, and a total of 152 subjects participated in one of the 12 (= 3 × 4) sessions. The number of participants per session varies from 8 to 17 due to no-shows, but the number of participants per treatment varies little (from 48 in SL treatment to 55 in BSL treatment). Python and its application Pygame were used to establish a server-client platform. After the participants were randomly assigned to separate computer cubicles, the experimenter read the general descriptions of the experiment out loud. Participants were asked to carefully read the instructions displayed on the monitor and to pass a comprehension quiz. Importantly, we do not track where participants are seated and emphasized to the participants that their decision would stay anonymous.

In all treatments, participants were subsequently asked to fill out a survey asking their basic demographic characteristics, risk preferences, and degree of familiarity with the experiment. The participants' risk preferences were measured by the dynamically optimized sequential experimentation (DOSE) method (Chapman et al., 2018; Imai and Camerer, 2018).

The average payment per participant was 8.31 EUR. The payments were made in private, and participants were asked not to share their payment information. Each session lasted less than 25 minutes.

4 Results

In each treatment, participants completed ten rounds of the same task, which means that we have ten observations per participant. We call each of these observations as a single report. In some parts of the analysis, we consider the aggregated observation at the participant level, which aggregates the single observations across the ten rounds. Unless indicated otherwise, the following analysis considers single reports, or reports for short.

4.1 Big lie vs. Small lie

Table 2: Overview of results in the BSL treatment

	Single reports Mean (SD)	N
Head	0.6727 (0.4696)	550
Head for dice≤ 4	0.6085 (0.4889)	281
Head for dice≥ 5	0.7398 (0.4396)	269
Dice	4.08 (1.6849)	550
Dice for Tail	3.75 (1.6507)	180
Dice for Head	4.2405 (1.6801)	370

Note: Under truth-telling, the expected average of Head is 0.5 and the expected average of Dice is 3.5.

4.1.1 Over-reporting on coin and dice

Table 2 summarizes reports in the BSL treatment. In this treatment, reports show that *head* was reported significantly more frequently than the 50% expected under truth-telling ($p<0.001$, binomial test (BT)).

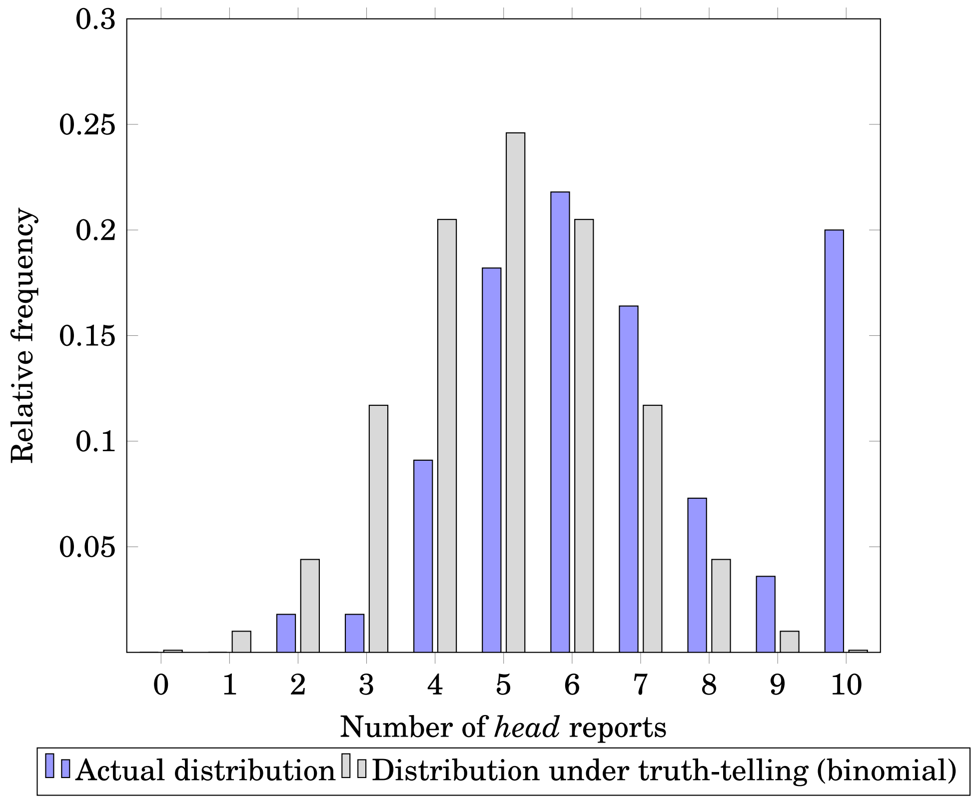
Considering observations at the participant level, Figure 1 shows the distribution of coin reports over the ten rounds. Compared to a binomial distribution with ten draws and probability 0.5, reports are significantly shifted towards higher numbers of *head* (two-sided Kolmogorov-Smirnov test (KS), $p<0.001$). This shift is particularly salient regarding the (too high) share of participants reporting ten rounds of *head*.

The distribution of reports on the dice roll in the BSL treatment is significantly shifted from the value expected under truth-telling of 3.5 ($p<0.001$, two-sided Wilcoxon signed-rank test (WT)). Figure 2 shows that reports are shifted towards higher outcomes, which results in a distribution significantly different from the uniform distribution expected from a fair dice ($p<0.001$, KS). More specifically, lying resulted in over-reporting of the outcomes 5 and 6 on the dice. Both of these outcomes were reported significantly more frequently than expected under truth-telling ($p=0.005$ and $p<0.001$, respectively, BT).

Moreover, from Figure 2, we can estimate that the percentage of truthful reports regarding the die roll is 57.6%.[3] Regarding the coin, from Table 2, we can estimate that

[3]Assuming that participants do not report a lower number than observed, the percentage of reports of 1 is an adequate estimate of true reports for each reported number (i.e., 9.6%x 6).

Figure 1: Distribution of the number of *head* reported in the BSL treatment



the percentage of truth-tellers is 65.4%.[4] These two estimates indicate that there were slightly more truth-tellers regarding the coin toss. However, that does not necessarily mean that the level of lying—which besides the *number* of lies, takes into account the *size* of the lies—is lower in the coin toss. We carefully test the latter aspect in the following subsection.

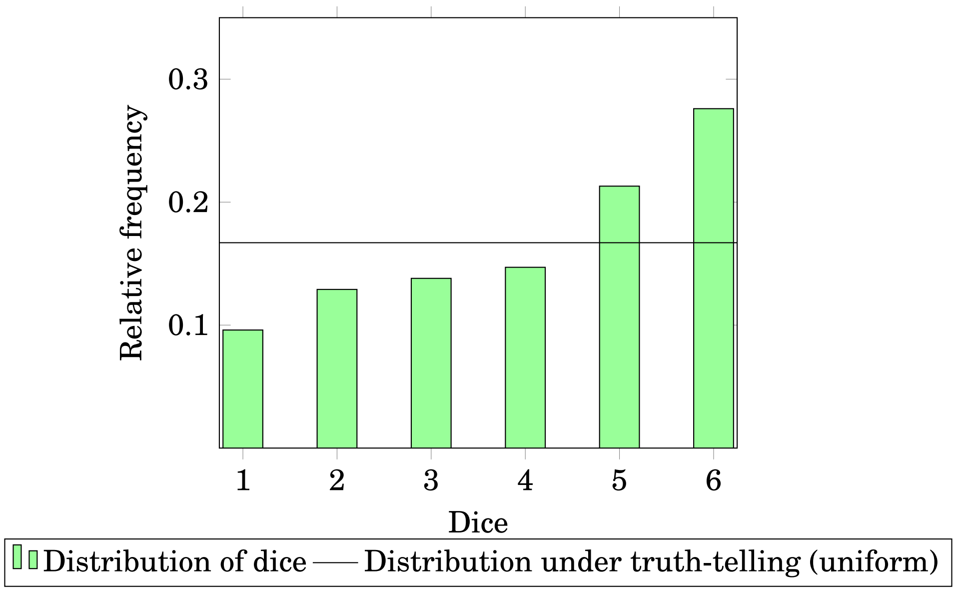
4.1.2 Comparing lying on the coin and the dice

The report on the coin is binary, while the report on the dice can take six different outcomes. Importantly, the latter report also has a lower marginal contribution to the overall payoff.

To compare lying behavior between the two dimensions, we standardize the reports (Abeler et al., 2019). Specifically, the reports are standardized so that the lowest possible report takes the value -1, and the highest possible report takes the value 1. For a dice roll with linear payments the reports $[1,2,3,4,5,6]$ become $[-1,-0.6,-0.2,0.2,0.6,1]$. For

[4]Assuming that participants who report *tail* are not lying, the percentage of *tail* reports is an adequate estimate of true reports for each reported outcome (i.e., 32.7%x2).

Figure 2: Distribution of single dice reports in the BSL treatment



a coin toss with *head* paying the higher amount, *head* will be evaluated at 1 and *tail* at -1. The average standardized report is 0.3455 (sd 0.9393) for the coin and 0.232 (sd 0.6739) for the dice. The standardized single reports on the coin are significantly higher than on the dice, which indicates a higher level of lying on the coin (p=0.001, WT).

As a robustness check, we further use the Bayesian method proposed by Hugh-Jones (2019)[5] to compare lying on the coin and the dice. Since this method is designed for a binary event, we consider a report of 5 or 6 on the dice as the high outcome because Figure 2 shows that lying led to over-reporting of these two outcomes. Figure 3 shows that the method by Hugh-Jones (2019) corroborates that there is more lying on the coin than on the dice in the BSL treatment.[6]

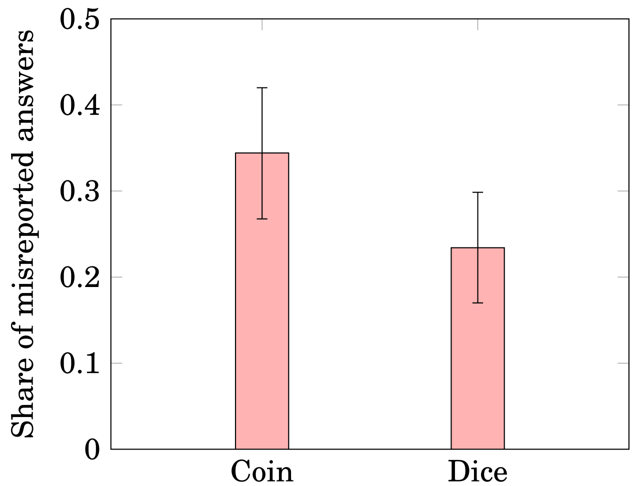
4.1.3 Lying conditional on the second dimension

To understand how reporting on one dimension is related to that on the other dimension, we analyze the connection of a coin report to the dice report that was submitted

[5] The measure uses the total number of reports, the number of reports that indicate the high outcome, and the probability of receiving the low payoff outcome under truth-telling to update an initial prior and to calculate a point estimate of the share of misreported answers as well as corresponding confidence intervals.

[6] Alternatively, we can consider the confidence intervals resulting from the method by Garbarino et al. (2018). This method provides narrower confidence intervals than the method of Hugh-Jones (2019), which is more conservative in the estimation of confidence intervals and is more reliable for small sample sizes. This alternative method estimates a 95% confidence interval of 0.287–0.397 for the coin and of 0.191–0.280 for the dice.

Figure 3: Estimates of lying and 95% confidence intervals based on Hugh-Jones (2019)



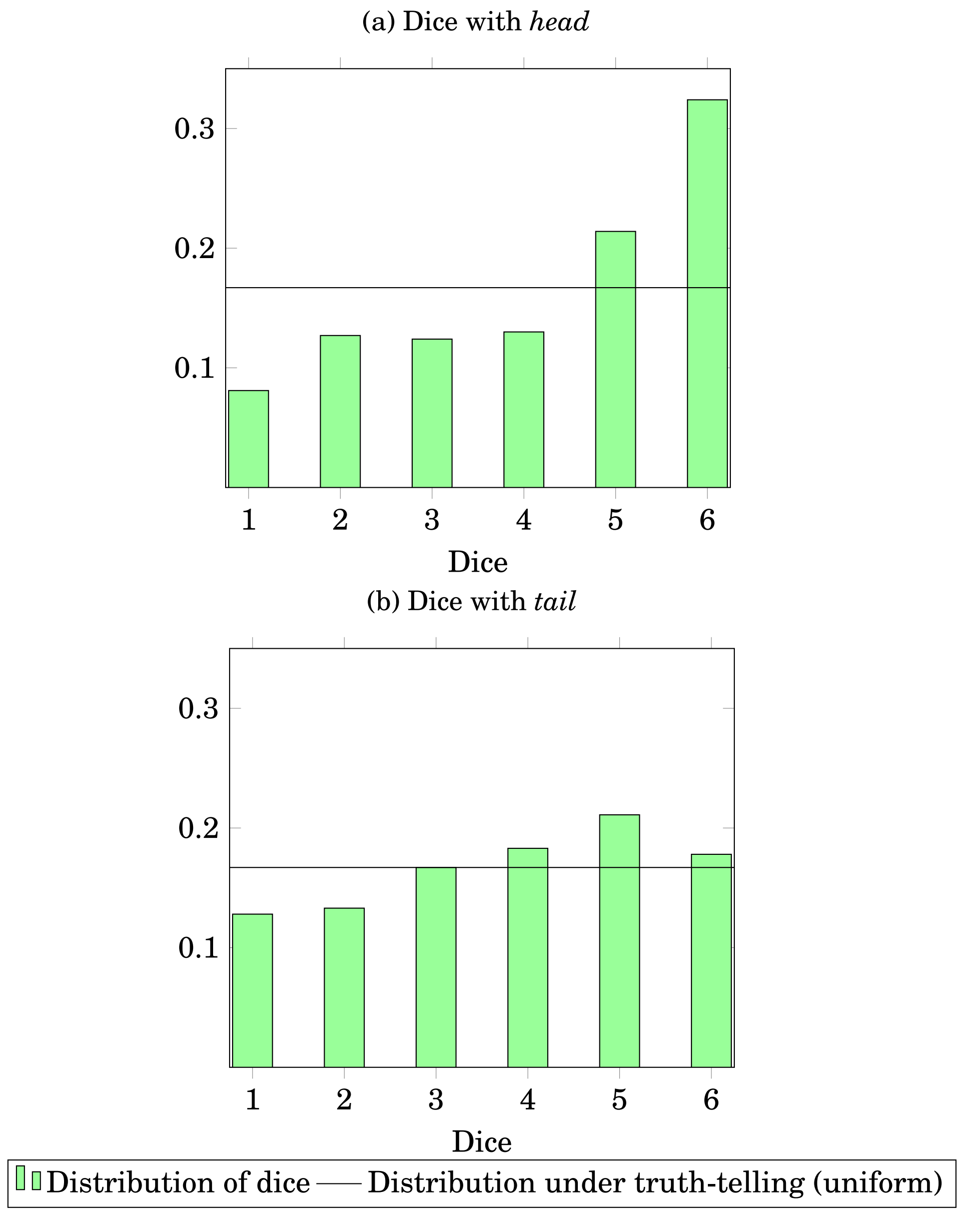
on the same screen (i.e., in the same round) and vice versa. Reports conditional on the other dimension are included in the overview in Table 2.

Since lying on the dice resulted in over-reporting of 5 and 6, we analyze whether *head* was reported significantly more often when the dice report was 5 or 6. Table 2 indicates that higher shares of *head* were reported for higher reports on the dice. The difference is significant (p=0.001, Fisher's exact test (FE)), which shows that participants reporting behavior (either lying or truth-telling) is consistent between the coin and the dice dimensions.

We find the same connection between reports when analyzing the reports on the dice conditional on the report of the coin. When *head* was reported, the average report on the dice was higher than when *tail* was reported. The difference is significant (p=0.001, two-sided Mann-Whitney U test (MW)), which further shows that participants are consistent on their reports, i.e., high reports on the coin correspond to high reports also on the dice. Figure 4, which shows the distributions of dice reports for *tail* and *head*, corroborates the latter finding. Conditional on reporting *head*, we see significant over-reporting of the outcomes 5 and 6 (p=0.018 and p<0.001, respectively, BT). Conditional on reporting *tail*, we observe that reporting of 5 and 6 is more than 17%, respectively, but these inclinations are not significant (p=0.110 and p=0.689, respectively, BT). Moreover, the two distributions shown in Figure 4 are significantly different (p=0.009, KS). Notably, we cannot reject that the distribution of the dice reports conditional on reporting *tail* is different from the expected distribution under truth-telling (p=0.304, KS).

To account for the fact that reports are clustered at the participant level, we conduct a regression analysis with the dice report as the dependent variable and the coin as an

Figure 4: Distribution of dice single reports for each coin outcome in the BSL treatment



independent variable.[7] In Model 1a in Table 3, we report a regression analysis with standard errors clustered at the participant level. The regression results of Model 1a support the positive correlation of coin and dice described above. However, when we include participant fixed effects (Models 1b and 1c), there is no significant correlation between the reports. The change caused by introducing fixed effects indicates that the correlation between single reports is driven by some participants always reporting high outcomes on both dimensions, while others always report lower outcomes on both dimensions. In other words, the differences in the report level are due to between-participant differences rather than within-participant differences.

Table 3: Linear regression analysis of dice reports in the BSL treatment

Outcome variable:	Single dice report			Average dice report	
	Model 1a[a]	Model 1b[b]	Model 1c[b]	Model 2a[b]	Model 2b[b]
head	0.491*	−0.118	−0.136		
	(0.190)	(0.157)	(0.159)		
sum_head				0.289***	0.287***
				(0.0367)	(0.0387)
constant	3.750***	4.159***	4.344***	2.135***	2.475***
	(0.127)	(0.125)	(0.233)	(0.259)	(0.407)
Participant FE	No	Yes	Yes	–	–
Round FE	No	No	Yes	–	–
Controls	–	–	–	No	Yes
N	550	550	550	55	55
adj. R^2	0.017	−0.110	−0.111	0.531	0.536

[a] Standard errors clustered on participant level in parentheses.

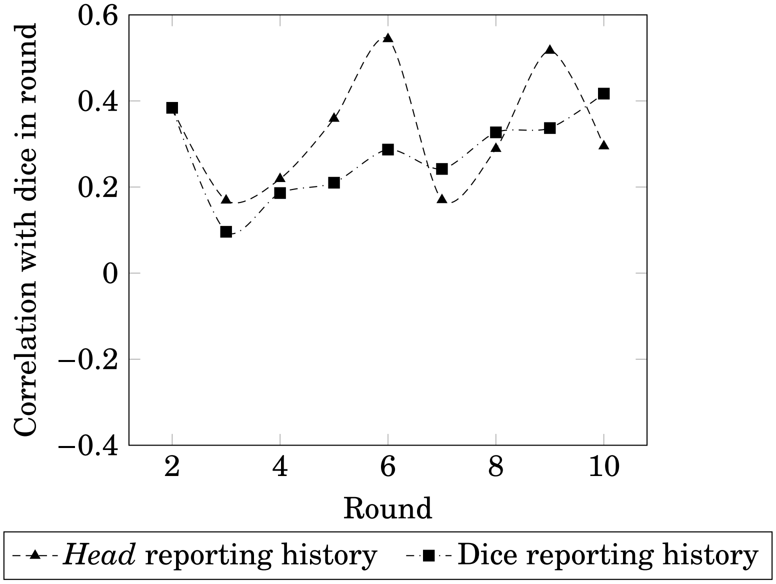
[b] Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As a robustness check, we conduct a regression analysis at the participant level. The dependent variable in the new specification represents the average dice report across the ten rounds per participant, and the independent variable represents the sum of *head* reports across the ten rounds per participant. Models 2a and 2b in Table 3 show that participants who made many *head* reports also made higher reports on the dice.

[7] Since participants report two outcomes simultaneously, this analysis does not capture any causal relationships. It only captures the correlation between the two reports.

Figure 5: Correlation of dice report with coin and dice history in the BSL treatment



Individual characteristics have no effect on reporting.[8]

Finally, the result found in Models 2a and 2b that reporting a high number of *head* is positively correlated with high reports on the dice suggests that the reports in earlier rounds are predictive of later rounds. In Figure 5, we show that this is indeed the case. In particular, we see that the history of reporting is positively correlated with reports in a given round. The latter observation further supports the finding that participants behaved consistently in the BSL treatment.

4.2 Big lie under an exogenous low-stakes prize

Table 4: Overview of results in the BL treatment

	Single reports Mean (SD)	N
Head	0.6633 (0.4731)	490
Head for dice≤ 4	0.6752 (0.4691)	314
Head for dice≥ 5	0.6420 (0.4808)	176

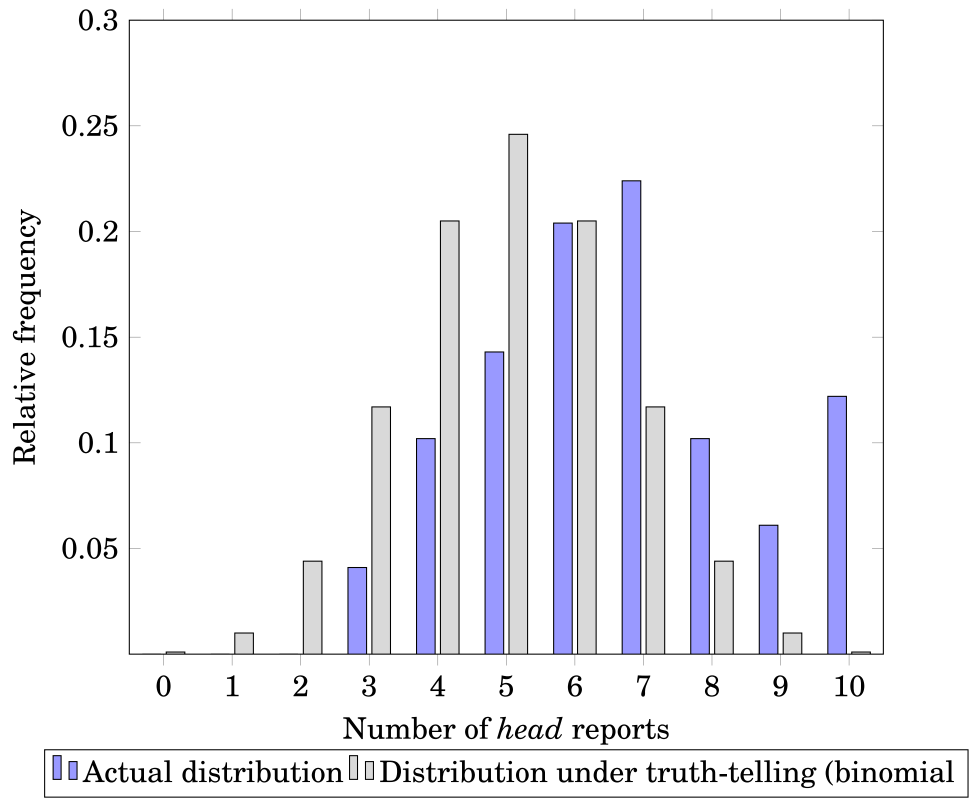
Note: Under truth-telling, the expected average of Head is 0.5.

Table 4 gives an overview of the results in the BL treatment, where participants

[8] This result also holds when excluding the sum of *head*, which might already contain the effect of individual characteristics, from the model.

reported the outcome of a coin toss while the outcome of a dice was exogenously randomly determined. In this treatment, we also find significant over-reporting on the coin ($p<0.001$, BT). In Figure 6, we look at the distribution of participant's coin reports in the ten rounds of the BL treatment. The distribution is significantly different from a binomial distribution with ten draws and probability 0.5 ($p<0.001$, KS). The difference is driven by an increase in the frequency of 7,8,9, and, in particular, of 10 rounds of *head*.

Figure 6: Distribution of the number of *head* reported in the BL treatment



Regarding the effect of receiving an exogenously randomly drawn low-stakes prize, we observe that there is less lying on the coin for participants who observed a high outcome on the dice, but the difference is not significant ($p=0.259$, FE). The linear regression analysis of single coin reports shown in Table 5 provides further support that coin reports in the BL treatment were not affected by the dice outcome that determines the exogenous low-stakes prize.[9] The coefficient for the dice is negative but insignificant, and this result is robust when including participant- and round-fixed effects.

Finally, we also conduct a regression analysis for the BL treatment at the participant

[9]The results are qualitatively the same when using logistic regression instead of the linear probability model.

Table 5: Linear regression analysis of coin reports in the BL treatment

Outcome variable:	=1 if *head* reported			Sum of *head*	
	Model 1a[a]	Model 1b[b]	Model 1c[b]	Model 2a[b]	Model 2b[b]
dice	−0.0149	−0.0180	−0.0198		
	(0.0140)	(0.0123)	(0.0124)		
avg_dice				0.153	0.0755
				(0.521)	0.551)
constant	0.716***	0.727***	0.761***	6.091**	7.506**
	(0.0544)	(0.0481)	(0.0775)	(1.861)	(2.400)
Participant FE	No	Yes	Yes	–	–
Round FE	No	No	Yes	–	–
Controls	–	–	–	No	Yes
N	490	430	430	49	49
adj R^2	0.001	−0.106	−0.102	−0.019	−0.079

[a] Standard errors clustered on participant level in parentheses.

[b] Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

level. Specifically, we use a linear regression model in which the dependent variable is the sum of *head* that a participant reported over all the ten rounds, and the independent variable is the average dice report across the ten rounds per participant. Models 2a and 2b in Table 5 indicate that neither the average observed dice outcome nor individual characteristics affect the number of rounds in which *head* was reported. Figure 7 illustrates how this result on the aggregate reports of participants developed over rounds. The correlations between the coin report and the history of observed dice are small and are not consistently above or below 0. Reporting history on the coin has a small correlation with reports in a given round, but correlations increase over time.

4.3 Small lie under an exogenous high-stakes prize

Table 6 summarizes reports in the SL treatment. Recall that participants in this treatment only reported the outcome of a dice while the outcome of the coin was exogenously randomly determined. Reports on the dice are significantly shifted from the expected value under truth-telling of 3.5 (p<0.001, WT). As shown in Figure 8, reports are shifted away from smaller outcomes towards the outcomes of 5 and 6, which are both reported significantly more frequently than expected (p=0.005 and p<0.001, respectively,

Figure 7: Correlation of *head* report with coin and dice history in the BL treatment

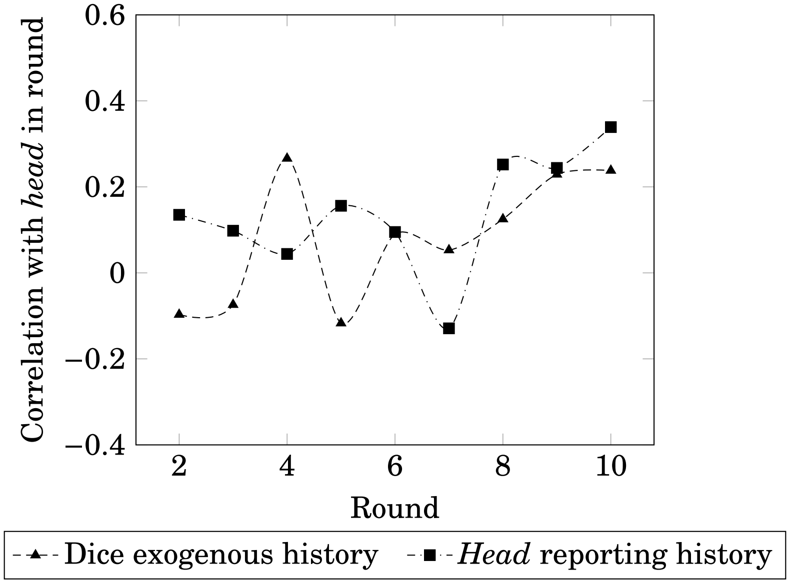


Table 6: Overview of results in the SL treatment

	Single reports Mean (SD)	N
Dice	4.1292 (1.6474)	480
Dice for Tail	4.1745 (1.5850)	235
Dice for Head	4.0857 (1.7073)	245

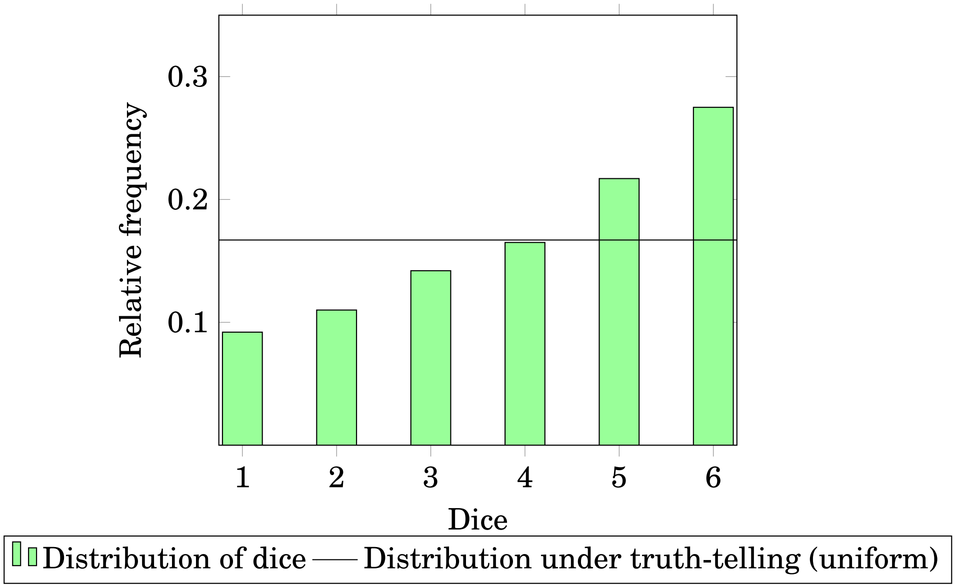
Note: Under truth-telling, the expected average of Dice is 3.5.

BT).

Regarding the effect of receiving an exogenously randomly drawn high-stakes prize, we observe that there is less lying on the dice for participants who observed *head* on the coin, but the difference is not significant (p=0.711, MW). A graphical analysis supports the latter finding. Figure 9 shows the distribution of dice reports for each outcome of the coin. We observe that both distributions have a clear shift towards high outcomes and, most importantly, the two distributions are not significantly different (p=0.963, KS), which indicates that observing an exogenously randomly determined outcome for the coin does not affect reporting on the dice. The regression analysis in Table 7 corroborates that dice reports in the SL treatment were not affected by the participants' observation of the coin outcome. The coefficient for the *head* is not significant, and this result is robust when including participant and round fixed effects.

In Models 2a and 2b of Table 7, we report the results of a linear regression analysis

Figure 8: Distribution of single dice reports in the SL treatment



at the participant level where the independent variable is the average dice report a participant made over ten rounds. Notably, these models' results show that the number of rounds of *head* is negatively correlated with the average dice report. In other words, this analysis unveils that participants who observed more rounds of *head* (i.e., had more luck) reported on average lower outcomes for the dice. Individual characteristics have no significant effect on reporting.

Figure 10 shows how participant reports in a given round are correlated with the history of reporting and the history of the exogenous coin. The largely negative correlations with the history of coin observations reflect the result of Models 2a and 2b in Table 7. After being (repeatedly) lucky on the exogenous outcome, there is a lower tendency to make high reports on the dice. Moreover, the positive correlations with the dice reporting history show that participants were consistent over rounds, and that previous reports are predictive of reporting in a given round.

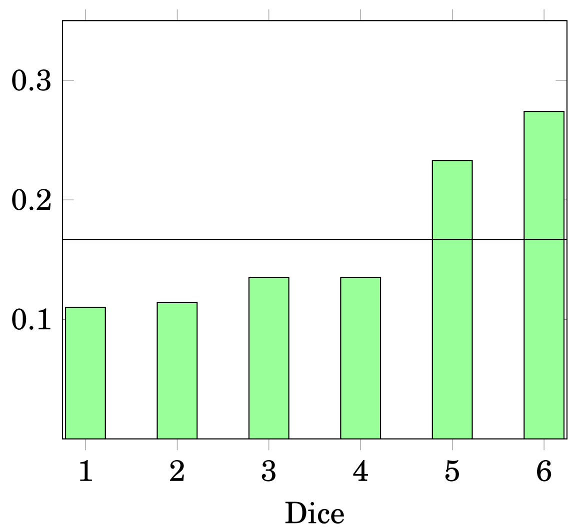
4.4 Comparison of treatments

4.4.1 BSL vs. BL: Comparison of self-reports on the coin

First, we test whether receiving a low-stakes prize had a significant effect on reports on the coin outcome, i.e., whether the overall level of reporting on the coin is different between the BSL and BL treatments. We find that the share of *head* does not differ significantly between the BSL and BL treatments (p=0.792, FE). Thus, the overall level

Figure 9: Distribution of single dice reports for each coin outcome in the SL treatment

(a) Dice with *head*



(b) Dice with *tail*

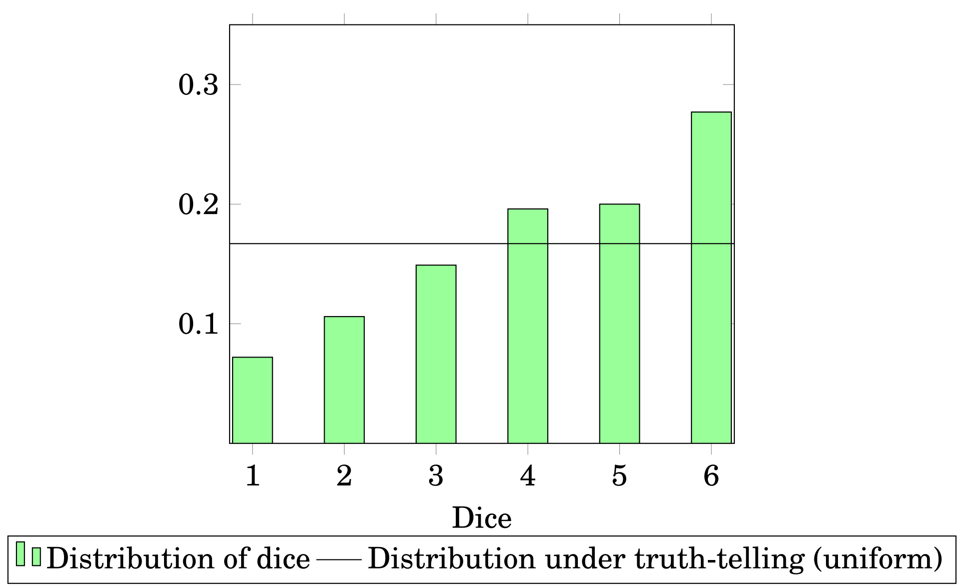


Table 7: Linear regression analysis of dice reports in the SL treatment

Outcome variable:	Single dice report			Average dice report	
	Model 1a[a]	Model 1b[b]	Model 1c[b]	Model 2a[b]	Model 2b[b]
head	−0.0888	0.111	0.117		
	(0.128)	(0.138)	(0.138)		
sum_head				−0.248***	−0.317***
				(0.0889)	(0.0943)
constant	4.174***	4.072***	4.151***	5.396***	5.010***
	(0.143)	(0.0966)	(0.226)	(0.470)	(0.655)
Participant FE	No	Yes	Yes	–	–
Round FE	No	No	Yes	–	–
Controls	–	–	–	No	Yes
N	480	480	480	48	48
adj. R^2	−0.001	−0.110	−0.077	0.126	0.111

[a] Standard errors clustered on participant level in parentheses.

[b] Standard errors in parentheses.

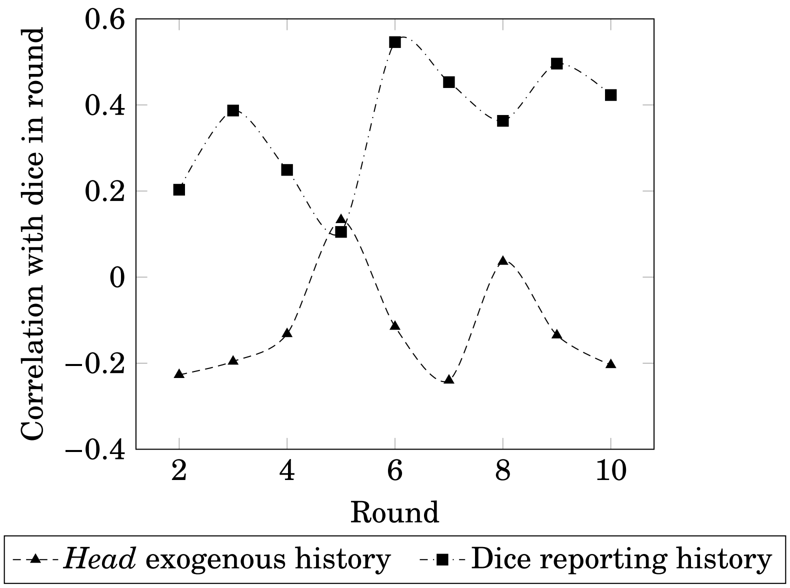
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

of lying on the coin is similar, regardless of whether the outcome of the dice is reported by the participant or exogenously determined.

Second, we assess how the coin reports depend on the dice outcome, which is either self-reported or exogenously determined. While in the BSL treatment *head* is reported more frequently together with high reports on the dice, the pattern is reversed in the BL treatment, i.e., *head* is reported less frequently when participants observe a high dice outcome. The difference in patterns for coin reports conditional on reporting (BSL treatment) or observing (BL treatment) high outcomes on the dice is significant (p=0.018, FE). The same holds for coin reports, conditional either on reporting (BSL treatment) or observing (BL treatment) low outcomes on the dice (p=0.054, FE).

While Models 2a and 2b in Table 3 show a positive correlation between the number of *head* a participant reports and the average dice report, Models 2a and 2b in Table 5 show no significant effect of the observed average die roll on the amount of *head* reports. Figure 11 corroborates this difference between treatments graphically. Regarding the BSL treatment, we observe a positive correlation between the two reports. In the BL treatment, however, we observe evidence for no correlation between the observed dice and the reported coin.

Figure 10: Correlation of dice report with coin and dice history in the SL treatment



To further test the difference between treatments, we conduct a joint regression analysis of reports at the participant level. Results are presented in Table 8. In Model 1, the positive coefficient for the average dice indicates that reporting (or observing) high dice outcomes is positively correlated with reporting many *head* rounds in the combined data set of the BSL and BL treatments.[10] Model 1 also indicates that the average participant aggregate report on the coin does not differ between the two treatments. Model 2 includes an interaction of the average dice and the BL treatment. In this model, the average dice coefficient now captures only the correlation between coin and the dice report at the participant level in the BSL treatment. The large significant coefficient for the BL treatment has to be interpreted jointly with the interaction term. For an average dice of 3.5, participant reports on the coin in the BL treatment were, on average, 6.98+3.5*(-1.713)=0.9845 points higher than in the BSL treatment. However, the difference between the coin in the BSL and the BL treatment decreases for increasing dice reports, which in combination with higher dice reports in BSL explains why we observe no significant difference in coin reports between the two treatments.

[10]In this regression, the dice variable is endogenous in the BSL treatment but exogenous in the BL treatment. Thus, we do not claim a causal relationship. We merely intend to identify possible differences in the two treatments.

Figure 11: Relation of sum of head and average dice in the BSL and BL treatments

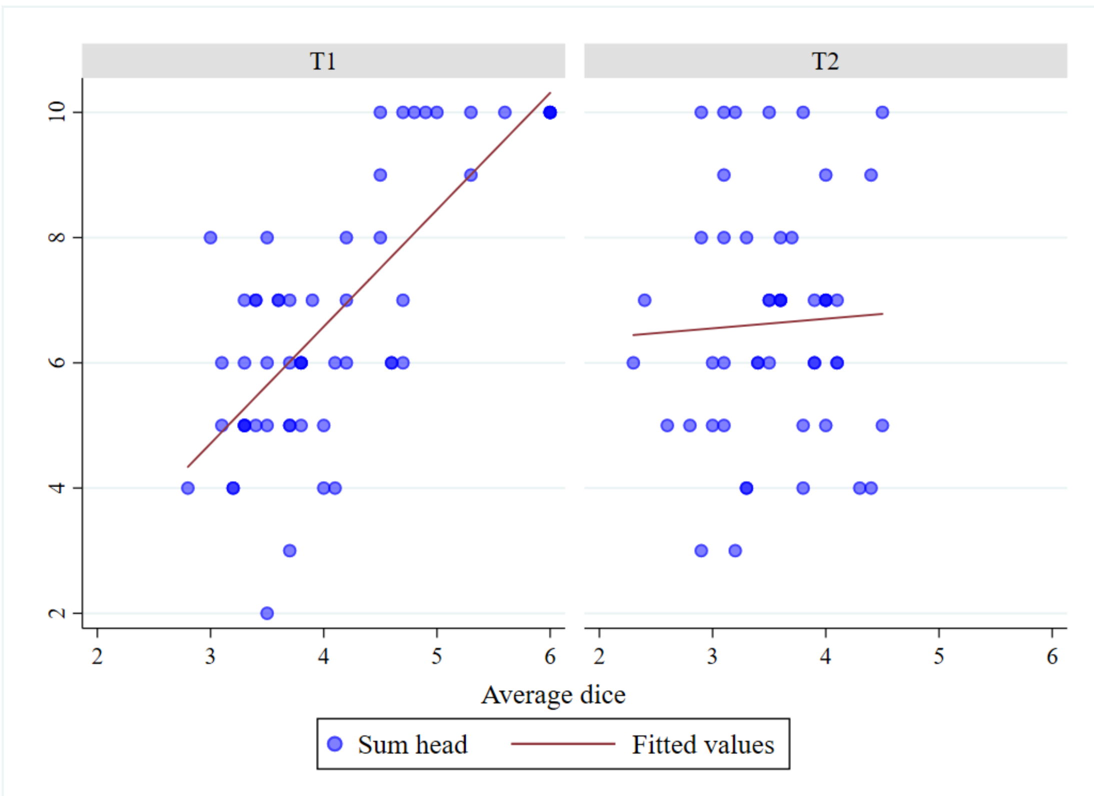


Table 8: Linear regression analysis of coin reports at the participant level in the BSL and BL treatments

Outcome variable: Participant aggregate coin report (sum of head)	Model 1	Model 2
avg_dice	1.404^{***}	1.867^{***}
	(0.248)	(0.277)
BL	0.677	6.980^{***}
	(0.376)	(1.997)
avg_dice*BL		-1.713^{**}
		(0.534)
constant	0.998	-0.889
	(1.039)	(1.155)
N	104	104
adj. R^2	0.058	0.292

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

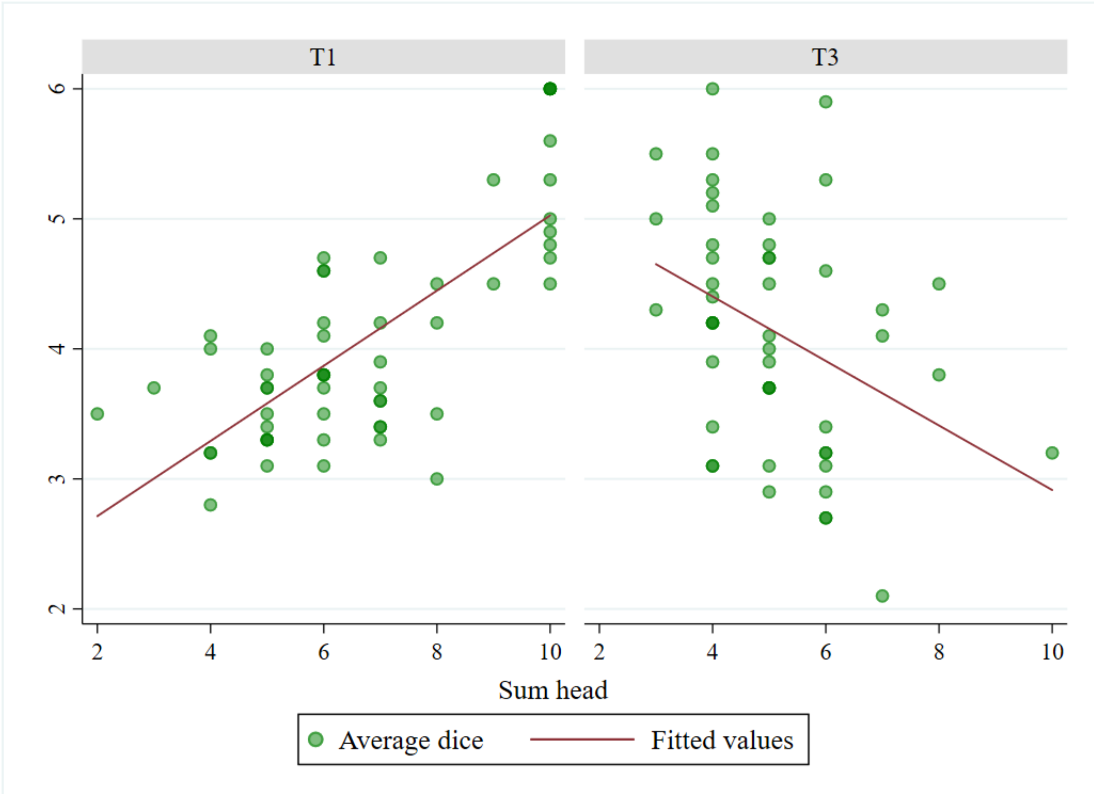
4.4.2 BSL vs. SL: Comparison of self-reports on the dice

First, we test whether receiving a high-stakes prize had a significant effect on reports on the dice outcome, i.e., whether the overall level of reporting on the dice is different between the BSL and SL treatments. We find that lying on the dice is not significantly different between the BSL and SL treatments (p=0.999, KS; p=0.708, MW). Thus, whether there was a possibility to lie on the coin, or not, did not significantly affect the dice report.

For participants who self-report (BSL) or observe (SL) *head*, the reports on the dice do not differ significantly (p=0.840, KS; p=0.235, MW). After reporting or observing *tail*, the distribution of reports in the SL treatment is shifted more towards high outcomes than reports in the BSL treatment (p=0.009, MW), but this shift is not significant (p=0.259, KS).

From Models 2a and 2b in Table 3 and Table 7, respectively, we see that the effect of the total number of *head* reported, or observed, over the ten rounds on the average dice report is different in the two treatments. In the BSL treatment, participants who reported a higher number of *head* made higher dice reports. In the SL treatment, participants who observed a higher number of *head* drawn by the computer made on average

Figure 12: Relation of average dice and sum of head in the BSL and SL treatments



lower reports on the dice. Figure 12 shows the difference between treatments graphically. We see a positive correlation between participant reports on the coin and the dice in the BSL treatment, whereas the correlation between the observed coin and dice reports is negative in the SL treatment.

As a final analysis, we conduct a joint regression of the BSL and SL treatments. In Model 1 in Table 9, we show that there is a positive correlation between the number of *head* rounds and the average dice report in the combined data of the BSL and SL treatments. Also, overall reporting on the dice is not significantly different in the BSL and SL treatments. Regarding Model 2 in Table 9, we see that the coefficient of the number of *head* rounds is the same as in Model 2a in Table 3 because it only captures the relation of coin and dice in the BSL treatment. The coefficient of the SL treatment has to be interpreted together with the interaction term. For five rounds of *head*, the average dice in the SL treatment is 3.26+5*(-0.537)=0.575 points higher than in the BSL treatment.

Table 9: Linear regression analysis of dice reports at the participant level in the BSL and SL treatments

Outcome variable: Participant aggregate dice report	Model 1	Model 2
sum_head	0.142^{**}	0.289^{***}
	(0.0500)	(0.0399)
SL	0.279	3.260^{***}
	(0.181)	(0.481)
sum_head*SL		-0.537^{***}
		(0.0885)
constant	3.126^{***}	2.135^{***}
	(0.313)	(0.263)
N	103	103
adj. R^2	0.070	0.319

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Conclusion

Overall, our study of jointly-reported big and small lies shows that people lie more on the big lying option than on the small lying option. Further, we find that people who are more willing to tell a big lie are also more willing to tell a small lie. However, the differences in lying do not cause different behavior than in a setting where only one lie can be told. Given this evidence, we conclude that people consider the two lying options largely independently and behave consistently across them.

Our design also allows us to consider the effect of observing an exogenous, payoff-relevant outcome on lying behavior. Our results indicate that being lucky about this exogenous income component decreases lying. However, this is only the case when the satisfaction derived by being lucky is big enough. When the small outcome was exogenous, and it was possible to tell a big lie, the observation of the exogenous outcome did not factor into people's decision to lie. Observing the big exogenous outcome did affect lying behavior, but only if the satisfaction was building up, i.e., multiple successful rounds were observed.

References

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

— **, Daniele Nosenzo, and Collin Raymond**, "Preferences for truth-telling," *Econometrica*, 2019, *87* (4), 1115–1153.

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

Azrieli, Yaron, Christopher Chambers, and Paul J. Healy, "Incentives in Experiments: A Theoretical Analysis," *Journal of Political Economy*, 2018, *126* (4), 1472–1503.

Barron, Kai, "Lying to appear honest," WZB Discussion Paper SP II 2019–307 2019.

Blanken, Irene, Niels van de Ven, and Marcel Zeelenberg, "A Meta-Analytic Review of Moral Licensing," *Personality and Social Psychology Bulletin*, April 2015, *41* (4), 540–558.

Chapman, Jonathan, Erik Snowberg, Stephanie Wang, and Colin Camerer, "Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)," 2018. CESifo Working Paper, No. 7262, Center for Economic Studies and Ifo Institute (CESifo), Munich.

Charness, Gary, Celia Blanco-Jimenez, Lara Ezquerra, and Ismael Rodriguez-Lara, "Cheating, incentives, and money manipulation," *Experimental Economics*, 2019, *22*, 155–177.

Chowdhury, Subhasish M., Chulyoung Kim, and Sang-Hyun Kim, "Pre-planning and the Continuation of Dishonest Behavior: An Experiment," Working Paper 2018.

Cohn, Alain and Michel André Maréchal, "Laboratory Measure of Cheating Predicts School Misconduct," *The Economic Journal*, 2018, *128* (615), 2743–2754.

Cojoc, Doru and Adrian Stoian, "Dishonesty and charitable behavior," *Experimental Economics*, 2014, *17*, 717–732.

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

Dolan, Paul and Matteo M. Galizzi, "Like ripples on a pond: Behavioral spillovers and their implications for research and policy," *Journal of Economic Psychology*, April 2015, *47* (Supplement C), 1–16.

Dufwenberg, Martin, "Lies in Disguise - A Theoretical Analysis of Cheating," *Journal of Economic Theory*, December 2018, *175*, 248–264.

Fischbacher, Urs and Franziska Föllmi-Heusi, "Lies in Disguise—an Experimental Study on Cheating," *Journal of the European Economic Association*, June 2013, *11* (3), 525–547.

Garbarino, Ellen, Robert Slonim, and Marie Claire Villeval, "Loss Aversion and lying behavior: Theory, estimation and empirical evidence," *Journal of the Economic Science Association*, 2018, *4*, 136–150.

Geraldes, Diogo, Franziska Heinicke, and Stephanie Rosenkranz, "Lying on two dimensions and moral spillovers," Working Paper 2019.

Gerlach, Philipp, Kinneret Teodorescu, and Ralph Hertwig, "The truth about lies: A meta-analysis on dishonest behavior.," *Psychological bulletin*, 2019, *145* (1), 1.

Gneezy, Ayelet, Alex Imas, Amber Brown, Leif D. Nelson, and Michael I. Norton, "Paying to be nice: Consistency and costly prosocial behavior," *Management Science*, 2012, *58* (1), 179–187.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel, "Lying Aversion and the Size of the Lie," *American Economic Review*, 2018, *108* (2), 419–53.

Hilbig, Benjamin E and Isabel Thielmann, "Does everyone have a price? On the role of payoff magnitude for ethical decision making," *Cognition*, 2017, *163*, 15–25.

Houser, Daniel, Stefan Vetter, and Joachim Winter, "Fairness and cheating," *European Economic Review*, 2012, *56* (8), 1645–1655.

Hugh-Jones, David, "True Lies: Comment on Garbarino, Slonim and Villeval (2018)," *Journal of the Economic Science Association*, 2019.

Hurkens, Sjaak and Navin Kartik, "Would I lie to you? On social preferences and lying aversion," *Experimental Economics*, 2009, *12*, 180–192.

Imai, Taisuke and Colin Camerer, "Estimating Time Preferences from Budget Set Choices Using Optimal Adaptive Design," Working Paper 2018.

Kajackaite, Agne and Uri Gneezy, "Incentives and cheating," *Games and Economic Behavior*, 2017, *102*, 433–444.

Lewicki, Roy J., "Lying and Deception: A Behavioral Model," in M. H. Bazerman and R. J. Lewicki, eds., *Negotiation in Organizations*, Sage Publications 1983, pp. 68–90.

Mazar, Nina, On Amir, and Dan Ariely, "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance," *Journal of Marketing Research*, December 2008, *45* (6), 633–644.

Merritt, Anna C., Daniel A. Effron, and Benoît Monin, "Moral self-licensing: When being good frees us to be bad," *Social and personality psychology compass*, 2010, *4* (5), 344–357.

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

Potters, Jan and Jan Stoop, "Do cheaters in the lab also cheat in the field?," *European Economic Review*, 2016, *87*, 26–33.

Shalvi, Shaul, Francesca Gino, Rachel Barkan, and Shahar Ayal, "Self-Serving Justifications: Doing Wrong and Feeling Moral," *Current Directions in Psychological Science*, April 2015, *24* (2), 125–130.

Appendix - Experimental Instructions (BSL treatment)

Welcome to this experiment. Please read these instructions carefully.

Overview

The experiment consists of 10 rounds. In each round, your task is to toss a coin, roll a dice, and report the outcomes. Your cash payment will be based on your reports. The details follow.

Your task

You can find one coin and one dice in front of you on the table. Please inspect them to verify that they are fair. In each round, toss a coin and roll a dice. Report the outcomes on the computer. You repeat this procedure for 10 rounds. In each round, your points will be determined as follows:

- 15+[outcome of the dice] if the coin lands Head
- 7+[outcome of the dice] if the coin lands Tail

For example, if you report (Head, 4), your points will be 19 (=15+4). If you report (Tail, 6), your points will be 13 (=7+6).

Payment

The server computer will randomly select one round, and your points in that round will be paid. This means that each round has an equal chance to be selected for the final cash payment. Thus, it is in your best interest to take each round equally seriously. Your points will be converted into Euros at the exchange rate of 2 points = 1 euro.

Anonymity

Your choices and answers will be linked with a computer number of your seat. We will never link your identity with your responses in any way. Your personal information provided for your payments will never be stored nor used for any research. In addition, since we do not track where you seat, we cannot match you with your reports, although we match the reports with the computer number.

Quiz

To ensure your understanding of the instructions, we will provide you with a quiz. If you have one or more wrong answers, you have to re-take the quiz. This quiz is only intended to check your understanding of the instructions. It will not affect your earnings.

Q1 If the coin lands head, and the outcome of the dice is 4, how many points do you receive?

Q2 If the coin lands tail, and the outcome of the dice is 1, how many points do you receive?

Q3 If the coin lands head, and the outcome of the dice is 6, how many points do you receive?

Q4 If the coin lands tail, and the outcome of the dice is 5, how many points do you receive?