Does the die-under-the-cup device exaggerate cheating?

Abstract

Using a powered online experiment (774 subjects, 54% female, av. age = 24.27) under the die-under-the-cup paradigm, this paper shows that a minimal variation (reversing payoffs) increases participants' honesty. Dice numbers and monetary prizes are aligned in the control treatment (1→5€, 2→10€, …, 6→30€), while numbers and monetary prizes go in opposite directions in the reversed treatment (1→30€, 2→25€, …, 6→5€). Although this small variation has no theoretical consequences, it results in more honest behavior. Since the participants in the control and the treatment are identical, we conclude that the observed dishonesty is caused by the task. The effect is stronger for women and older participants.

Keywords: Honesty, order effects, reversed payoffs, die-under-the-cup.

Highlights

- This paper reports results from a powered online experiment about honesty (n = 774).
- We show that a reversed version of the die-under-the-cup task compared to the original devices achieves higher levels of honesty.
- We conjecture that the die-under-the-cup task exaggerates cheating.

1. Introduction

The die-under-the-cup paradigm (DUC) has been extensively used to measure honesty (see Shalvi et al., 2011a, 2011b; Fischbacher and Föllmi-Heusi, 2013; Gaetcher and Schulz, 2016). In the DUC task, each participant is asked to roll a 6-sided die and report the outcome. The outcome determines the payoff (money) the participants receive. It should be noted that experimentalists cannot know whether subjects are telling the truth and, importantly, subjects know that the experimentalists are blind to the outcome. Furthermore, the decision has no impact on other participants. Hence, there is no social cost or externality for dishonest behavior (only the participant's self-image) and participants have incentives to lie. The multi-country study by Gaetcher and Schulz (2016) showed that subjects lie substantially in this environment (see also Figure 2 in Abeler et al., 2019). However, Charness et al. (2019) found that subjects do not cheat in the absence of financial incentives. Hence, cheating is caused by financial incentives and not by the structure of the game.

This paper tests whether observed lying in this sort of experiments is artifactual, that is, partially motivated by the design. In other words, given a large enough sample of subjects randomly assigned to treatments A and B, we may say that the observed behavior is artefactual if $outcome^A$ is different to $outcome^B$. If samples are ex-ante identical, then the outcome is determined by the mechanism.

We test whether showing payoffs in the reversed order, that is, when numbers and monetary prizes go in opposite directions (1→30€, 2→25€, …, 6→5€), have any impact on honesty compared to the original order where dice numbers and monetary prizes are aligned (1→5€, 2→10€, …, 6→30€). Results show that this minimal variation decreases dishonesty although the *incentives to lie remain the same.* Therefore, this result suggests that the (online) DUC device "artificially" induces lying.

2. Experimental design

Our experiment uses the DUC paradigm with a variation where earnings are shown in reverse order. The task is introduced to participants in three subsequent phases:

- [dice problem] participants are asked to place an imaginary die in a cup, shake it, and see the resulting number (see Jiang, 2013)[1]
- [payoffs] participants receive information regarding payoffs: control [TO] or reversed treatment [Trev])
- [decision] subjects reveal the result by providing a number from 1 to 6.

The only difference between TO and TRev appears in the payoffs phase where the earnings are explained. The dice and decision phases are the same. Regarding the payoffs phase, the difference emerges from the link between the dice numbers and the payoffs. Table 1 illustrates the difference between treatments.

Table 1 Treatments.

TO (control)	TRev (reversed)
Number 1 → €5	Number 1 → €30
Number 2 → €10	Number 2 → €25
Number 3 → €15	Number 3 → €20
Number 4 → €20	Number 4 → €15
Number 5 → €25	Number 5 → €10
Number 6 → €30	Number 6 → €5

In TO, the dice numbers and monetary prizes are aligned (multiplication factor = 5), while in TRev the numbers and prizes go in opposite directions. Therefore, TRev is simply a variation in the order of TO. It is important to emphasize that the *incentives to lie remain the same.*

The question here is that a minimal change in order may have an effect on dishonesty. The purpose of this experiment is precisely to show whether a minimal

[1] Our design is similar to the "mind game" of Jiang (2013). The key difference is that Jian varies the order of the screens while ours varies the order of the payoffs.

variation in the order makes subjects cheat less (more) in an environment where economic costs are absent, there are no externalities, and the only potential cost is individual self-image.

3. Sample and recruitment

A total of 774 subjects (407 female, 54%, av. age = 24.27) participated in the online experiment. Subjects were assigned to TO (control) with $p = 0.5$.[2]

The experiment was programmed in LimeSurvey and conducted online. The three phases mentioned before were displayed on three subsequent screens: dice problem, payoffs, and decision. The subjects were informed that 1 out 10 participants would be paid real money. Indeed, 72 subjects received a monetary prize (average €21.875) for completing a task of 2 minutes and 30 seconds.

Participants were recruited by students from Loyola University who invited contacts to participate in the experiment (see Jorrat, 2000). The recruiters were informed that a prize of €100 would be given to the most successful recruiter. Recruiters had no impact on the assignment of participants to treatments. All participants signed an informed consent. The experiment was closed after 4 hours.

Table 2 summarizes the sample characteristics by treatments. We also compare whether the subsamples differ (*t*-test) in these characteristics.

Observe that we collected a large set of characteristics. Some of them are related to cheating, while others are completely unrelated. For instance, the question “In general, others consider me an honest person" refers to *honesty*, while *number chosen* refers to the number the participants guessed to win a prize in the lottery (last screen) and *mobile* refers to whether they prefer iPhone or Android. The sample is identical regarding all these characteristics.

[2] All data are available at Mendeley (see Alfonso et al., 2022).

Table 2 Sample: Balance.

	TO	*TRev*	*p-value*
Participants	404	370	0.249 [a]
Age [b]	24.20	24.34	0.999
Female	0.56	0.51	0.825
Honesty	4.40	4.40	0 .973
Happiness	3.90	3.97	0.232
Number chosen	5.26	5.05	0.188
Mobile: iPhone	0.56	0.54	0.587
Mobile: Android	0.44	0.45	0.705
Job: Sanitary	0.24	0.20	0.167
Job: Education	0.15	0.18	0.357
Job: Finance	0.10	0.12	0.219
Job: Industry	0.11	0.09	0.194
Job: Other	0.40	0.41	0.637
Team: Team1	0.35	0.32	0.438
Team: Team2	0.10	0.11	0.598
Team: Team3	0.09	0.10	0.788
Team: Team4	0.09	0.10	0.887
Team: Other	0.37	0.37	0.866
Place: Place1	0.25	0.21	0.128
Place: Place2	0.16	0.21	*0.078*
Place: Place3	0.15	0.15	0.940
Place: Place4	0.12	0.12	0.833
Place: Other	0.32	0.31	0.753

Notes: a) *p-value* refers to the *t*-test except the proportion of participants (Wald test), b) *min = 16* and *max = 85*.

Table 2 also shows the job distribution of both TO and Trev since we also asked about the participants' occupation (current job for those who were already working and desired job for those were not yet working).[3] Finally, we also included questions about personal tastes like *preferred football team* or place of *residence*. Again, no differences were found between subsamples.

The randomization worked properly as both samples were indistinguishable in a large array of personal characteristics, including self-reported honesty. Recall that if samples are ex-ante identical and the outcome differs, then the difference is caused by the mechanism.

4. Results

Let us begin by mentioning that the expected earnings should be €17.50 under truth-telling. The participants in our experiment (TO+TRev) earned €21.45 (SD = 7.62, SE = 0.273), thus indicating dishonest behavior.

Fig. 1 (left side) shows the cumulative distribution function (CDF) of the uniform distribution (in dots), the original treatment (in blue), and the reversed treatment (red dashes). As can be seen, the uniform distribution is always above the CDF of both the original and the reversed treatments, implying that the subjects in our experiment overreported beneficial numbers (larger payoffs). Hence, at least a fraction of the experimental subjects are not truth-tellers.

Fig. 1 (right side) complements the analysis and shows the average earnings for participants in the original and reversed treatments. Figure 1 (bottom) shows the requested payoffs by treatment. Participants in the original treatment requested more money that those in the reversed treatment, thus indicating that subjects in the original treatment are less likely to tell the truth, i.e., they lie more.[4] Hence, when payoffs are shown in reverse order, subjects lie less. Given the balance shown in

[3] We did not ask whether they are currently working or not.

[4] The sample size yields an effect from 0.207 times the standard deviation with a power of 0.8 and an alpha of 0.05.

the previous section, it is hard to believe that the samples are different, i.e., that the proportion of honest people differs by treatment.

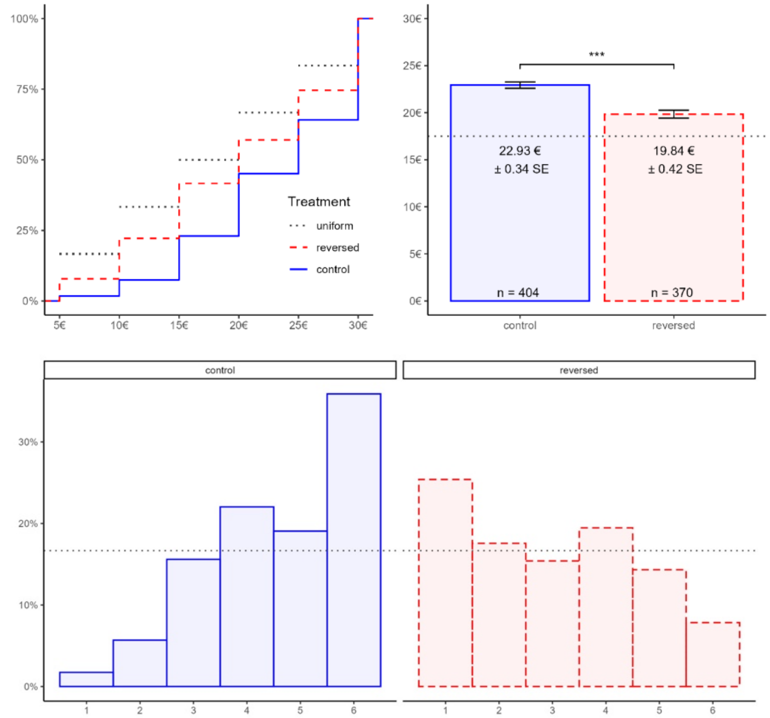

Fig. 1. Requested money by treatment.

Top left: CDF for the uniform, control, and reversed distribution. Top right: Mean (SE) of control (TO) and reversed (TRev) treatments. Bottom: Histogram by treatment.

Therefore, if the reversed version (where payoffs decrease with the die numbers) provides different results although the incentives to cheat are identical, then either *i*) the original DUC is an accurate measurement of honesty and the reversed DUC is under-reporting cheating or *ii*) the reversed DUC is accurate and the original one is over reporting cheating.

Now we show the distribution of self-reported honesty. It is important to mention that the subjects were asked if they were honest *after* the decision. Hence, they replied after being honest or after cheating in screen #3. Fig. 2 shows the results.

First, we observe that there are no differences between treatments, implying that the participants are identical. Subjects in both subsamples reported the same average value of 4.4 out of 5 (see Table 2): subjects consider themselves fairly honest and this self-valuation is independent of the treatment. Hence, the subjects did not change their self-image due to the treatment. Interestingly, their self-image did not change due to asking a beneficial number and the subjects did not have the feeling they were cheating even if they were in fact doing so.

We therefore conclude that subjects in both subsamples are identically honest (average 4.4) and that *the original DUC exaggerates cheating*.

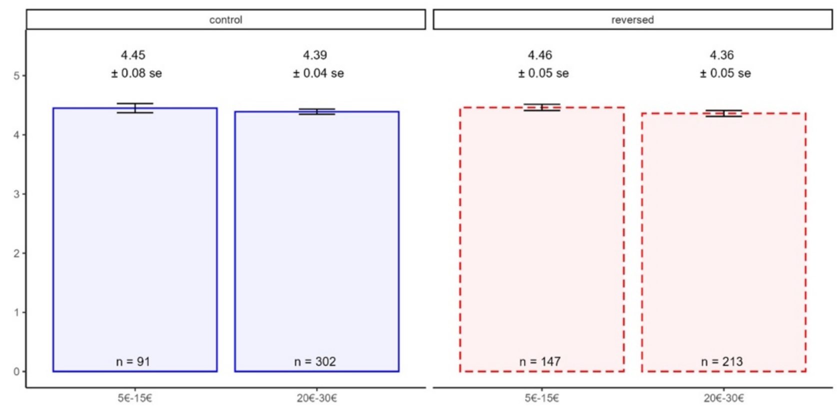

Fig. 2. Self-reported honesty by choice and treatment.

Note: €5€-€15 shows the proportion of subject who chose €5, €10, and €15; €20-€30 represents €20, €25, and €30.

A possible explanation is that reversing the payoff (i.e., decreasing numbers) complicates cheating by making computation harder or more cognitively demanding (increasing “lying costs” in Abeler et al. 2019, Shalvi et al., 2012). In the DUC set-up, people easily associate increasing high numbers (die roll outcomes) with increasing large payments. In fact, in most card games, high numbers are associated with large earnings. Hence, subjects may pick high numbers in the DUC not because they explicitly wish to cheat, but because it is easy and probably comes naturally to them. In other words, it would be the usual response.

Hence, reversing the numbers might make the task harder as it would require more cognitive resources, but be unusual to the natural response (picking the highest number). These arguments support the idea that DUC artificially exaggerates cheating.

We now discuss some additional results. Regarding gender (Figure 3 left), we see that women lie more than men in the control group ($p = 0.003$), whereas they behave identically to men under the reversed order ($p = 0.989$). Therefore, gender differences vanish in the reversed order treatment. This result is consistent with Ezquerra et al. (2018). Hence, our results indicate that women are not more dishonest than men but more likely to overreport cheating when using the standard DUC.

Participants 22 years old or younger (age $\leq$ 22) requested more money (Fig. 3 right) than older participants (age > 22; $p = 0.030$). When comparing treatment effects by age groups, we see that the effect is significant for both younger ($p = 0.001$) and older participants ($p = 0.009$), that is, reversing the order reduces dishonesty in both sub-samples. In fact, older participants do not lie in the reversed order. The latter result is in line with the meta-analytic review of Gerlach at al. (2019) and implies that both old and young people are sensitive to the DUC device. However, we found that older people are even more sensitive as they were more likely to overreport large numbers in our experiment.

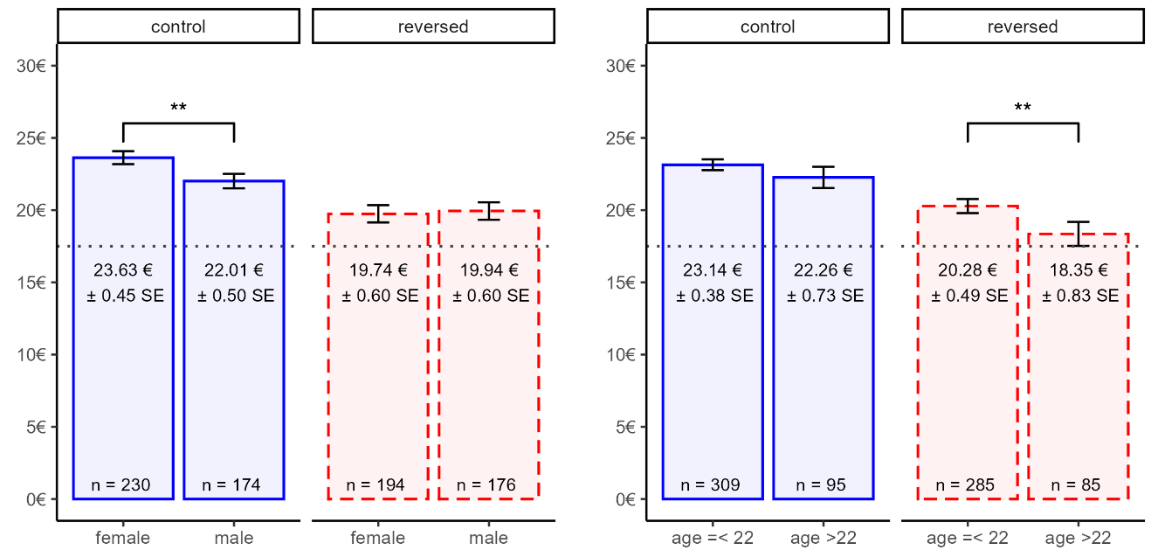

Fig. 3. Requested money by treatment: Gender and age.

5. Conclusions

Using a powered experiment, this paper shows that subjects lie less often if a minimal variation of the DUC task is introduced: a change in the order. In our original treatment, dice numbers and monetary prizes were aligned, while in the reversed treatment, numbers and monetary prizes went in opposite directions. This small variation resulted in more honest behavior. Since *the incentives to lie (and the theoretical prediction) remain the same*, this paper poses the question of whether the observed dishonest behavior might be artefactual, that is, caused by the task itself.

Given that individuals' honesty is unobservable, we cannot assess whether the DUC is indeed exaggerating cheating or the reversed DUC is underreporting cheating.

To shed some light on this issue, we used self-reported honesty. Since the majority of the sample stated that they were honest regardless of the treatment, we conclude that the DUC exaggerates cheating.

According to Abeler et al. (2019), dishonesty might be explained by costs, conformity, and reputation. Since reputation has no role in the DUC, both costs and conformity may explain our results. Indeed, lying costs might be perceived as higher in the reversed DUC, but subjects may also expect less lying in the reversed version. However, our results show that lying cannot being explained by moral costs since most of the participants in our experiment did not feel guilty. Moreover, no differences were found in self-reported honesty due to beneficial choices. Therefore, costs other than moral ones (or social norms) are driving these differences.

Additionally, we also observed that the treatment effect was stronger for women and older participants, indicating that these subsamples cheat more when using the DUC. However, we need to consider that our experiment was run online and we cannot ensure that the observed results would also be replicated in the “real” experiment.

6. References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truth-telling. Econometrica, 87(4), 1115-1153.

Alfonso, A., Branas-Garza, P., López-Martín, M.C. (2022). Dataset: Does the die-under-the-cup device exaggerate cheating?, Mendeley Data, V2, doi: 10.17632/mtxcrx5j52.2

Ayal, S., Gino, F. (2012). Honest rationales for dishonest behavior. *The Social Psychology of Morality: Exploring the Causes of Good and Evil,* 149-166.

Charness, G., Blanco-Jimenez, C., Ezquerra, L., Rodriguez-Lara, I. (2019). Cheating, incentives, and money manipulation. *Experimental Economics*, 22(1), 155-177.

Ezquerra, L., Kolev, G. I., Rodriguez-Lara, I. (2018). Gender differences in cheating: Loss vs. gain framing. *Economics Letters*, 163, 46-49.

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

Gaetcher, S., Schulz, J. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531(7595), 496-499

Gerlach, P., Teodorescu, K., Hertwig, R. (2019). The truth about lies: A meta-analysis on dishonest behavior. *Psychological Bulletin*, 145, 1.

Jorrat, D. (2021). Recruiting experimental subjects using WhatsApp. *Journal of Behavioral and Experimental Economics,* 90, 101644.

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

Shalvi, S., Eldar, O., Bereby-Meyer, Y. (2012). Honesty requires time (and lack of justifications). *Psychological Science*, 23(10), 1264-1270.

Shalvi, S., Dana, J., Handgraaf, M. J., De Dreu, C.K. (2011a). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes*, 115(2), 181-190.

Shalvi, S., Handgraaf, M., De Dreu, C. K. (2011b). Ethical manoeuvring: Why people avoid both major and minor lies. *British Journal of Management*, 22, S16-S27.