Journal of Economic Behavior & Organization 145 (2018) 24–33

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Research Paper

Allowing for reflection time does not change behavior in dictator and cheating games

Steffen Andersen[a], Uri Gneezy[b,c], Agne Kajackaite[d], Julie Marx[a,*]

[a] Copenhagen Business School, Department of Finance, Solbjerg Plads 3, 2000 Frederiksberg, Denmark
[b] University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States
[c] CREED, University of Amsterdam, Netherlands
[d] Berlin Social Science Center, Reichpietschufer 50, 10785 Berlin, Germany

ARTICLE INFO

Article history:
Received 1 April 2017
Received in revised form 6 October 2017
Accepted 20 October 2017
Available online 7 November 2017

JEL classification:
C91
D03
D63
D82

Keywords:
Cheating
Dictator
Delay
Reflection

ABSTRACT

Reaction time, usually measured in seconds, has been shown to be correlated with decisions in experimental games. In this paper, we study how allowing for a full day of "reflection time" alters behavior. We compare behavior in dictator and cheating games when participants make immediate choices with behavior when participants have an extra day to decide, and find that allowing for more time does not affect behavior.

1. Introduction

The revealed-preference approach in economics is focused on choice data. Yet, as Gul and Pesendorfer (2008) write: "In standard economics, an individual's decisions may improve when a constraint is relaxed. For example, an agent may make better decisions if he is given better information, more resources, or more time to make his decision." Experimental studies show that sometimes such non-choice data are important in understanding preferences. One example is the correlation found between reaction time of individuals when making their decisions and their choices.

Consider, for example, allocation choices in dictator or public-goods games. If participants have clear preferences over allocations, neither the amount of time they are given to decide nor their reaction time are predicted to correlate with their choices. However, according to the dual-system models of cognitive processes (Fudenberg and Levine, 2006; Loewenstein, 2005; Kahneman, 2003, 2011), time to decide might affect the cognitive process leading to the decision. This literature discusses the difference between System I, which is rapid and intuitive, and System II, which is slower and deliberative. As the dual-system models predict, recent papers testing the difference between "intuitive" and "deliberate" decisions find

* Corresponding author.
E-mail addresses: san.fi@cbs.dk (S. Andersen), ugneezy@ucsd.edu (U. Gneezy), agne.kajackaite@wzb.eu (A. Kajackaite), jma.fi@cbs.dk (J. Marx).

https://doi.org/10.1016/j.jebo.2017.10.012
0167-2681/

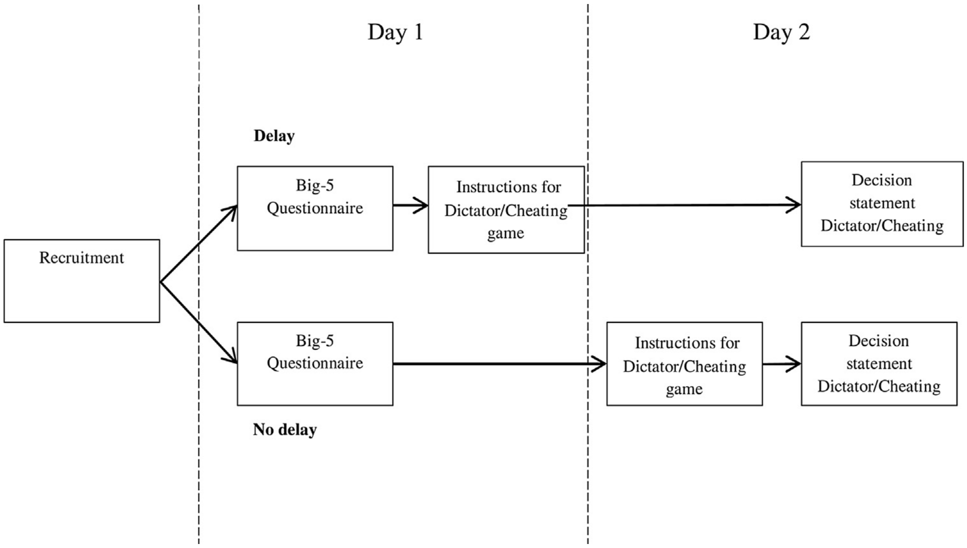

Fig. 1. The timeline of the experiment.

correlation between reaction time and decisions (e.g., Piovesan and Wengström, 2009; Grimm and MEngel, 2011; Jiang, 2013; Neo et al., 2013; Kocher et al., 2016; Rand et al., 2012, 2014; Nielsen et al., 2014; Lohse et al., 2016).

In the literature on reaction time, participants make decisions within an experimental session, and the reaction time is typically measured in seconds. In our everyday lives, however, we mostly have more than minutes to make decisions. In this study, we ask what happens when participants have a longer time period to reflect on social and ethical decisions. We do so by comparing decisions participants make in dictator and cheating games immediately during an experimental session with decisions they make after they have had a day to think about their choices. By giving participants time to "sleep on the decision," we allow them to reflect on the possible payoffs and moral costs, ask others for advice, and consider their decision while being home in familiar surroundings. We assume that in both conditions participants enter System II – participants are not pressured to decide fast and have some time to think about their decisions – therefore our interest is not giving more insights on dual-models, but understanding how a relatively long time to think affects decisions.

Consider a simple example of a cheating game: A participants is asked to roll a die in private and report the outcome to the experimenter. If the participant reports "5," s/he is paid 50 Euros; otherwise s/he is paid nothing. Cheating is inferred statistically by comparing the expected fraction of people who report a 5 with the actual fraction. In such a game, Kajackaite and Gneezy (2017) found the majority of participants do not lie. In a similar cheating game, we test whether giving participants a day to think about the decision and potentially discuss it with other people increases cheating. Other examples of experimental evidence on cheating include: Abeler et al. (2014), Abeler et al. (2017), Cohn et al. (2014), Dreber and Johannesson (2008), Erat and Gneezy (2012), Fischbacher and Föllmi-Heusi (2013), Gneezy (2005), Gneezy et al. (2017), Lundquist et al. (2009), Mazar et al. (2008), Shalvi et al. (2011) and Sutter (2009), amongst others. Similarly, we test whether time to think affects giving behavior in a dictator game.

We find no significant differences between decisions made immediately in the lab and decisions for which participants had a day to reflect – in neither dictator nor cheating games. From a methodological perspective, the results show the concern regarding participants not having enough time to consider their decision in the lab does not seem to be important for this kind of social and ethical decisions.

2. Experimental procedures, design, and sample

We employ a between-subjects design consisting of eight ($2 \times 2 \times 2$) treatments; we vary (i) the task (dictator game vs. cheating game), (ii) the given time to decide (no-delay vs. delay), and (iii) the size of the stake in the decision (low vs. high).

In all the treatments, participants were asked to come to the lab on two consecutive days. The timeline of the treatments is presented in Fig. 1. On the first day, participants read some initial instructions and completed a Big-5 questionnaire.[1] Depending on the treatment of reflection time, participants either (i) only completed the questionnaire and left the lab on day 1, or (ii) completed the questionnaire and received instructions for day 2 of the experiment. By giving participants

[1] All instructions are presented in Appendix B in the Supplementary data, along with the full questionnaire. Note that the use of the Big-5 questionnaire was completely instrumental. We collected the Big-5 data only because we needed to provide the participants in the no-delay treatments with some reason for coming to the experiment – otherwise they would have arrived just to pick up the show up fee, which is not conventional in experiments and might confuse the participants. To keep treatments consistent, we had all participants fill in the Big-5 questionnaire.

Table 1
Summary of treatments and number of participants.

Treatment	Stakes	Females	N
Dictator Game			
no delay	low 5 €	0.62	50
delay	low 5 €	0.70	55
no delay	high 50 €	0.60	51
delay	high 50 €	0.56	52
Cheating Game			
no delay	low 10 €	0.49	65
delay	low 10 €	0.69	67
no delay	high 50 €	0.71	41
delay	high 50 €	0.60	40

instructions for day 2 on day 1, we introduce the delay in decision-making—we provide participants with a day of reflection time to make their decision. On day 2, we gave instructions to the no-delay treatment group, and in both treatments, asked the participants to perform the experimental task. Lastly, participants completed a post-experiment questionnaire. Participants received a show-up fee of 6 Euros for each day.

2.1. The dictator game

The dictator tasks were based on the simple dictator game (Kahneman et al., 1986; Forsythe et al., 1994). Dictators received a cash endowment and were told that the amount they left for the recipient would be given to a matched participant from a different experiment.[2]

We used two levels of endowments for dictators. In the low-stakes treatment, we endowed dictators with 5 Euros (every transfer between 0 and 5 with an increment of 1 was feasible), and in the high-stakes treatment, we endowed dictators with 50 Euros (every transfer between 0 and 50 with an increment of 1was feasible).

Participants in the delay and the no-delay treatments received the same instructions; only the timing differed. Participants in the delay treatment received the instructions for the dictator game on day 1, but were asked to state their decision on day 2. On day 2, participants returned to the lab and stated their transfer decisions. In the no-delay treatment, on the other hand, the participants received the instructions and stated their decisions, both on day 2. By comparing the level of transfers in no-delay and delay, we can infer whether having one day to reflect affects the giving decision in a dictator game. The variation in the stakes allows us to infer whether the amount on the table alters these decisions.

2.2. The cheating game

In the second experiment, participants played a cheating game. We gave each participant a sealed envelope with 10 folded pieces of paper that carried the numbers from 1 to 10.[3] We asked participants to take out one piece of paper, look at the number they took out, put it back into the envelope, and then report it. The payments they received increased with the number they reported. Because the actual outcomes cannot be observed, participants had an incentive to cheat and report a higher number than actually observed.

In the low-stakes treatments, payments were equal to the reported number, in Euros. In the high-stakes treatments, payments were equal to the reported number multiplied by 5. In the high-stakes treatment, reporting a "1" resulted in 5 Euros, reporting a "2" resulted in 10 Euros, and reporting a "10" led to a payoff of 50 Euros.

We employ the same no-delay and delay treatments as we used in the dictator task. Participants in the delay treatments received the instructions for the cheating game on day 1. On day 1, they also received the envelope with numbers, were asked to take out one number from the envelope, put it back into the envelope, memorize it, and then report it the following day. We provided the no-delay participants with the instructions for the game on day 2 of the experiments, and they made their choice immediately after.

By comparing the distribution of numbers reported in the no-delay and delay treatments, we can infer whether having one day to reflect affects the lying behavior of the participants. By comparing the low- and high-stakes treatments, we can infer how reflection time and stakes matter for cheating behavior.

Table 1 presents the treatments and the number of participants in each treatment cell.

The experiment was conducted in June-July 2015 at the Cologne Laboratory for Economic Research, University of Cologne. We recruited 434 students via ORSEE (Greiner, 2015) to participate in the experiment. All participants were invited for two sessions on two consecutive days with 24 h between the sessions. However, 13 did not come back on the second day,

[2] For the exact wording, see instructions in the appendix. We paid the transfers to randomly chosen participants in the cheating-game treatments at the end of their second sessions (see the last instructions in the appendix for the message the recipients received).

[3] We use the same game as in Gneezy et al. (2017).

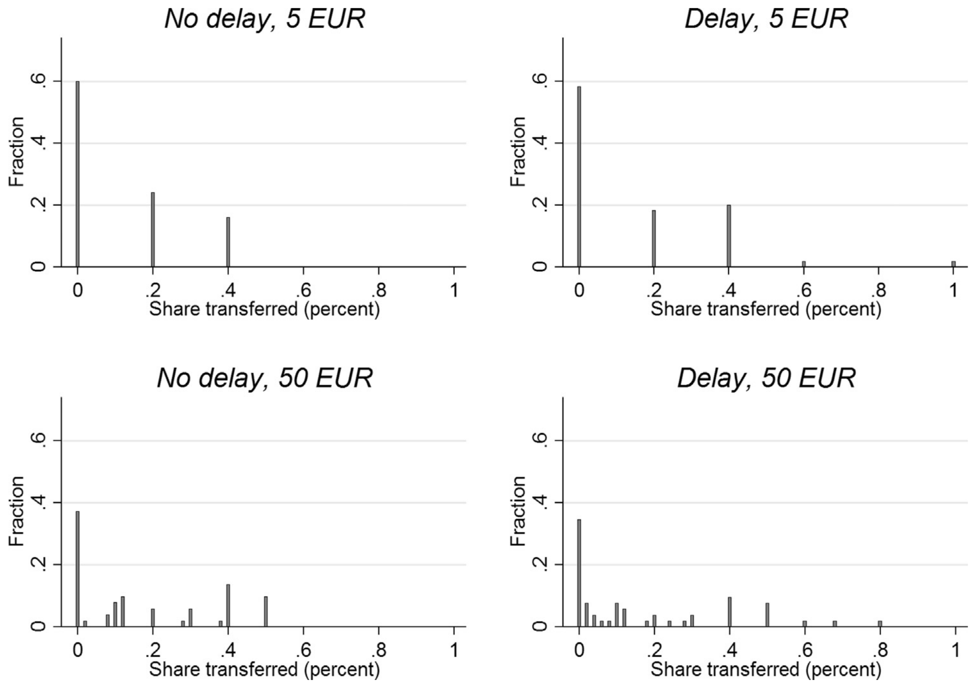

Fig. 2. Distribution of transfers in the dictator-game treatments.

Table 2
Average transfers in the dictator game.

	No delay	Delay	P-value
5 EUR	11.20% (0.56 EUR)	14.55% (0.73 EUR)	0.606
50 EUR	16.94% (8.47 EUR)	17.00% (8.50 EUR)	0.906
P-value	0.060	0.141	

leaving us with 421 participants (62% females) in the final sample (208 in the dictator game and 213 in the cheating game).[4] Participants took part only in this experiment in the experimental session, with a session lasting approximately 30 min.

In all the treatments, after stating their decisions, participants completed a post-experiment questionnaire that included questions on gender, age, field of study, motives behind their decisions, and timing of the decisions. At the end, participants privately received their payoffs in cash and left the laboratory.

3. Results

In what follows, we first present the results from the dictator-game treatments, then report the results from the cheating-game treatments, and end by relating to existing results for time effect in experiments.

3.1. The dictator game

Fig. 2 presents the distribution of transfers in the dictator game, and Table 2 shows statistical comparisons between the treatments. In the no-delay low-stakes treatment, dictators transfer 0.56 Euros (SD = 0.76) on average, out of 5 Euros (11.20% of the endowment, SD = 15.20). The majority (60%) transfers nothing to the recipient, whereas 24% transfer 1 Euro and 16% transfer 2 Euros. Providing the participants with one additional day to think about the transfer decision does not change their behavior. In the delay low-stakes treatment, dictators transfer 0.73 Euros (SD = 1.04) on average (14.55% of the endowment, SD = 20.89), which is not significantly different from the transfers in the treatment with no delay (p = 0.606, Mann–Whitney

[4] Ten out of 13 no-shows were participants in the no-delay condition (they did not have instructions for the day 2). Three no-shows were participants in the delay cheating treatment with low stakes.

Table 3
OLS regression models for dictator-game treatments.

	(1)	(2)
Delay	0.017 [0.026]	0.032 [0.037]
High Stakes	0.042 [0.027]	0.058 [0.038]
Delay X High Stakes		−0.031 [0.053]
Female	0.013 [0.027]	0.012 [0.028]
Constant	0.112*** [0.029]	0.105*** [0.032]
Observations	208	208
R-squared	0.014	0.016

Note: The models are OLS specifications with the share of the endowment transferred to the recipient (between 0 and 1) as the dependent variable. Standard errors are provided in the brackets.
*** Indicates coefficients that are significant at the 1% level by standard *t*-tests. Delay, High Stakes, and Gender are dichotomous variables equal to 1 if the participant is in a treatment with delay, the endowment is 50 Euros, and the participant is female, respectively.

U test).[5] Similarly to the no-delay treatment, the majority of dictators (58.18%) transfer nothing to the recipient and a high fraction of participants transfer 1 Euro (18.18%) or 2 Euros (20%).

The giving behavior is similar in the high-stakes treatments. In the no-delay high-stakes treatment, dictators transfer an average of 8.47 Euros (SD = 9.07), which amounts to 16.94% (SD = 18.13) of the endowment. The fraction who transfer zero is 37.25%, and other participants transfer up to a half the endowment. In the delay high-stakes treatment, dictators behave in a similar way—they transfer 8.50 Euros (SD = 10.59) on average (17.00% of the endowment, SD = 21.17), which is not significantly different from the transferring in the no-delay high-stakes treatment ($p = 0.906$, Mann–Whitney *U* test).[6] Here, similarly, to the no-delay high-stakes treatment, a high fraction of dictators (34.62%) transfers nothing, others transfer up to a half the endowment, and only a few participants transfer more than half the endowment.[7]

To examine if stakes matter for giving decisions, we compare low- and high-stakes treatments, and find the stakes have no significant effect on transfers in the dictator game, on conventional levels. Whereas in the no-delay low-stakes treatment, dictators transfer 11.2% of the endowment to the recipient, in the no-delay high-stakes treatment, the transfers increase to 16.94%, but only marginally significantly so: $p = 0.060$, Mann–Whitney *U* test. Similarly, the stakes have no effect on giving in the delay treatments, where the transferred share increases from 14.55% for low stakes to 17.00% for high stakes, but not significantly so ($p = 0.141$, Mann–Whitney *U* test).

Table 3 reports the results using a simple ordinary least-squares regression analysis, where we examine the robustness of the null effect, allowing for controls on treatments and gender.[8] We run two regressions to test for the effect of delay and stakes. The dependent variable in the regressions is the share of the endowment transferred to the recipient (a scaled variable with an interval from 0 to 1). The independent indicator variables are delay (0 = no delay, 1 = delay), high (0 = low stakes, 1 = high stakes treatment), and female (0 = male, 1 = female).

From Table 3, we see that neither of our treatments nor the simple demographic control has any effect on the transferring decision in the dictator game. Similarly, the regression results support the null effect of stakes.

3.2. The cheating game

Fig. 3 presents the distribution of numbers reported in cheating-game treatments, and Table 4 summarizes the results. In the no-delay low-stakes treatment, participants report an average number of 9.15 (SD = 0.138), with the majority of 53.85% reporting a 10 and a high fraction (30.77%) reporting a 9. In expectation, all numbers are equally probable, as the dashed horizontal line in the figure indicates. On average, participants clearly lie when reporting their numbers in the no-delay

[5] A power calculation using a two-sided *t*-test for simplicity shows that we would need a sample size of 470 individuals per treatment cell to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects.

[6] A simple power calculation using a two-sided *t*-test shows that a sample size of 1,762.692 per treatment cell is required to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects.

[7] In the post-questionnaire, we asked participants in delay treatments to self-report when they made their decisions (in the lab yesterday, between the sessions, or in the lab today). We group the data according to these answers, and report the results in Appendix A. The results in the dictator game do not change. We find that cheating behavior is affected in the low-stakes treatments—delay makes participants lie less. However, the result is driven by a very low fraction of participants (N = 11 for participants who delay their decision in the low-stakes cheating game).

[8] A two-censored Tobit model generates similar results on delay significance, but makes interpreting the coefficients harder. We therefore choose to report simple OLS results.

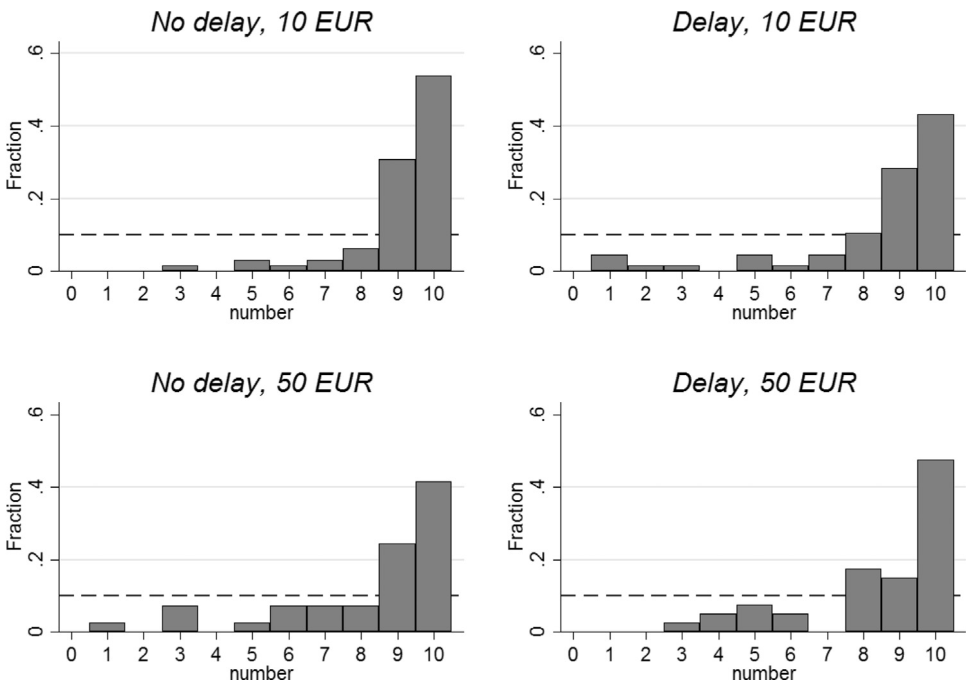

Fig. 3. Distribution of reported numbers in the cheating-game treatments.

Table 4
Average reported numbers in the cheating game.

	No delay	Delay	P-value
10 EUR	9.15	8.46	0.096
50 EUR	8.24	8.45	0.720
P-value	0.060	0.935	

treatment. Providing the participants with an additional day to think about which number they report does not change the reporting.[9] In the delay low-stakes treatment, participants report an average number of 8.46 (SD = 2.34), which is not significantly different from the treatment with no delay, on conventional levels ($p = 0.096$, Mann–Whitney *U* test).[10] Similarly to the no-delay treatment, a high fraction of participants reports a 9 (28.36%) and 10 (43.28%).

Cheating behavior does not change when moving to high stakes. In the no-delay high-stakes treatment, participants report an average number of 8.24 (SD = 2.36) with a high fraction reporting a 9 (24.39%) and 10 (41.46%). In the delay high-stakes treatment, the average number reported is 8.45 (SD = 2.06) and a high fraction reports a 10 (47.5%), whereas in this treatment, the fraction of 9's decreases to 15%. Overall, the reported number distributions are not significantly different between the two high-stakes treatments: $p = 0.720$, Mann–Whitney *U* test.[11] That is, from the comparisons of reporting between the delay and no-delay treatments, it follows that an additional day to think about whether to lie does not change behavior.

When we compare low- and high-stakes treatments, we find that stakes have no significant effect on lying, on conventional levels. Whereas in the no-delay low-stakes treatment, participants report an average number of 9.15 and we observe 8.24 in the no-delay high-stakes treatment, but this finding is only marginally significant ($p = 0.060$, Mann–Whitney *U* test). In the delay treatments, stakes have no effect on reporting; the average number reported is 8.46 in the delay low-stakes treatment and 8.45 in the delay high-stakes treatment ($p = 0.935$, Mann–Whitney *U* test).

In Table 5, we provide the simple ordinary least square regression, in which we check for the robustness of the null effect of the *delay* treatment on cheating.[12] We run equivalent regressions to the ones we conducted in the dictator game

[9] Since cheating is already high in the no-delay treatment, the null-result could be driven by a ceiling effect. However, since cheating decreases slightly in the low stakes treatment, this does not seem to be the case.

[10] A power calculation using a two-sided *t*-test for simplicity shows that a sample size of 123 per treatment cell is required to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects

[11] A power calculation using a two-sided *t*-test for simplicity shows that a sample size of 1820 per treatment cell is required to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects.

[12] Using an ordered logit does not change the significance of delay coefficients, and because interpreting the simple coefficients is easier, we report OLS.

Table 5
OLS regression models for cheating-game treatments.

	(1)	(2)
Delay	−0.327 [0.282]	−0.647* [0.361]
High Stakes	−0.443 [0.290]	−0.861** [0.413]
Delay X High Stakes		0.829 [0.585]
Gender	−0.290 [0.291]	−0.227 [0.293]
Constant	9.140*** [0.279]	9.266*** [0.292]
Observations	213	213
R-squared	0.024	0.033

Note: The models are OLS specifications with the share of the reported number (between 1 and 10) as the dependent variable. Standard errors are provided in the brackets. ***, **, and * indicate coefficients that are significant at the 1%, 5%, and 10% levels, respectively, by standard *t*-tests. Delay, High Stakes, and Gender are dichotomous variables equal to 1 if the participant is in a treatment with delay, the payoffs are five times the number reported, and the participant is female, respectively.

treatments with the dependent variable “Number reported”. Table 5 shows that the results from non-parametric tests are confirmed – *delay* has no significant effect on cheating.[13] Overall, regression analyses suggest no effect of stakes on cheating. However, as can be seen from model 2, the OLS analyses suggest that participants cheat significantly less for high stakes than low stakes in the no delay treatments.

3.3. *Position in the literature*

Our null results are in line with several studies of time delays in other experimental games. Oechssler et al. (2015) test how a 24 h delay affects rejections rates in an ultimatum games. They find only a weakly significant decrease in rejection rates. In a Stackelberg duopoly game, Cardella and Chiu (2012) find no effect of a ten minutes cooling-off period on second-mover responses.

In contrast to our results for the dictator and cheating game, Grimm and Mengel (2011) finds a large positive effect of ten minutes delay on acceptance rates of small offers in the ultimatum game. Neo et al. (2013) also finds a decrease in responder rejections in the ultimatum game when subject are given a break of 15 min. Cooling-off periods in emotional decisions have also been studied in the field. Lee (2013) finds that a forced waiting period of between one week and 3 months after divorce filing has significantly reduced divorce rates in South Korea. In contrast, Cramton et al. (1999) finds no effect of cooling-off periods on strike incidence and duration in Canadian labor market negotiations. Note, however, that unlike the studies described above, we do not study emotional responses.

Results on dictator games are mixed. Controlling for cognitive ability and general swiftness in the dictator game, Cappelen et al. (2016) finds a strong correlation between long response times and selfishness. In contrast, in a modified dictator game, Piovesan and Wengström (2009) find that faster subjects are more selfish.

Only one study – as far as we know – links response time and cheating. Jiang (2013) studies two variants of the mind game and find that honest subjects spend more time responding. She also finds that the easier it is to cheat, the longer is the response time of the honest participants.

4. Conclusion

In laboratory experiments, participants have little time to make their decisions, because an experimental session does not last more than a couple of hours. In this study, we test whether the restricted decision time drives participants’ decisions on giving and cheating. We find that having an additional day to think about the decision affects neither giving nor cheating.

This finding suggests decisions made about cheating and giving in the relatively short lab experiments are robust to variables such as long decision time, the possibility of discussing the decision with peers, and having the decision process outside of the lab environment. In the future research it would be interesting to see how a relatively long delay of decisions affects behavior in other games, as, for example, in more complex games.

[13] The delay effect is marginally significant when we include more control variables than in the simple OLS regression. Using a Tobit or ordered logit, the effect is not significant, not even marginally.

Acknowledgements

We thank student research assistants at the University of Cologne for their help while conducting the experiment Andersen is grateful to the Danish Social Science Research Council for financial support through research grant 11-104456, and the European Research Council through research grant 639383.

Appendix A. Delay according to the post-questionnaire answers

Although the participants had a day to think about their decisions in the delay treatments, they might not have made use of the extra time. In this section, we repeat the analysis using stated decision time instead of given decision time.

In the post-questionnaire, we asked the participants when they made their decisions. Participants could choose one of the following answers: “in the lab yesterday,” “between the sessions,” or “in the lab today.” In the following, we split the data with respect to when the participants said they had made their decisions. We pooled participants who answered with “in the lab yesterday” with the participants in the no-delay treatments and call the group “no stated delay” (see Table A1 and Fig. A1). If the participants answered with “between the sessions” or “in the lab today,” we assigned them to the stated-delay group.

13. a. The dictator game

In the dictator-game treatments, 49.53% of participants said they had made their decision in the lab on day 1, 30.84% between the sessions, and 19.63% in the lab on day 2. From this finding follows that roughly half the participants in the dictator-game treatments did not delay their decisions when given the opportunity to do so.

Table A1
Average transfers in the dictator game (stated delay).

	No stated delay	Stated delay	P-value
5 EUR	13.66% (0.68 EUR)	10.43% (0.52 EUR)	0.629
50 EUR	16.64% (8.32 EUR)	17.74% (8.87 EUR)	0.857
P-value	0.073	0.123	

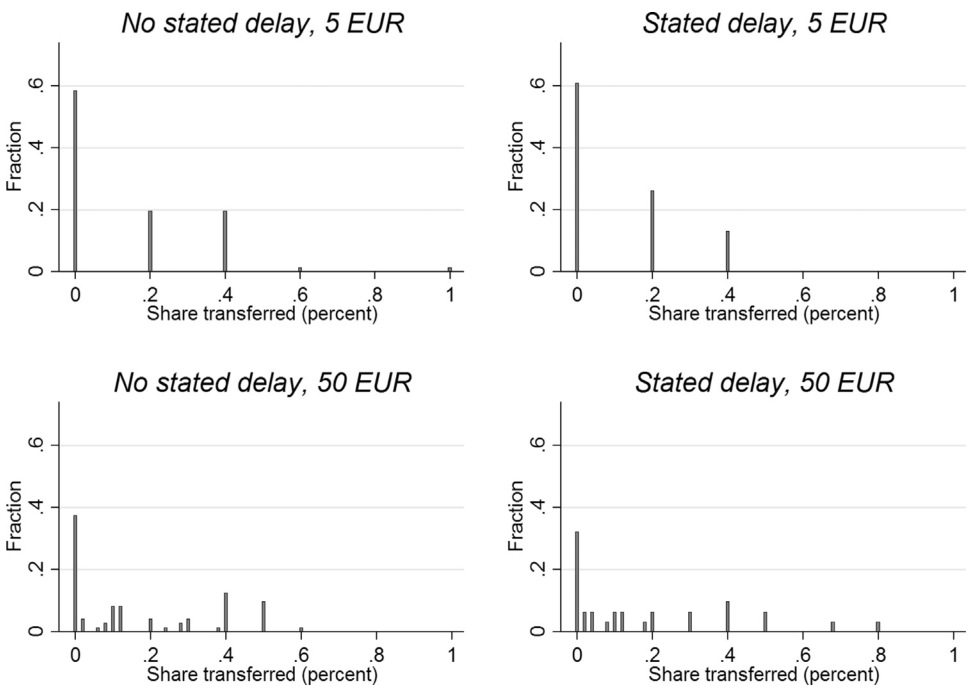

Fig. A1. Distribution of transfers in the dictator game (stated delay).
Notes: Number of observations amount to 82, 23, 72, and 31 in the no-stated-delay low-stakes, stated-delay low-stakes, no-stated-delay high-stakes, and stated-delay high-stakes treatments, respectively.

The results from the alternative definition of delay are in line with the conclusions in the main analysis: delaying a decision has no effect on behavior in a dictator game.[14] Also, stakes have no significant effect on transfers in the dictator game (see Table A1).

14. b. The cheating game

In the cheating-game treatments, we pool and analyze the data according to the participants' answers to the question, "When did you decide whether you would report truthfully?" We found that 84.11% of participants answered that they had made their decision in the lab the day before, 10.28% between the sessions, and only 5.61% in the lab that day. It follows that a vast majority of the participants did not delay their decisions in the cheating games, and they delayed significantly less than participants in dictator games ($p < 0.001$, Fisher exact test).

When pooling the data with respect to stated delay, we find different results than in the main analysis. Table A2 and Figure A2 report the results. Whereas the no-stated-delay low-stakes group reports an average number of 8.95 (SD = 1.77), participants in stated-delay low-stakes group report an average number of 7.18 (SD = 3.06). Fig. A2 shows a high fraction

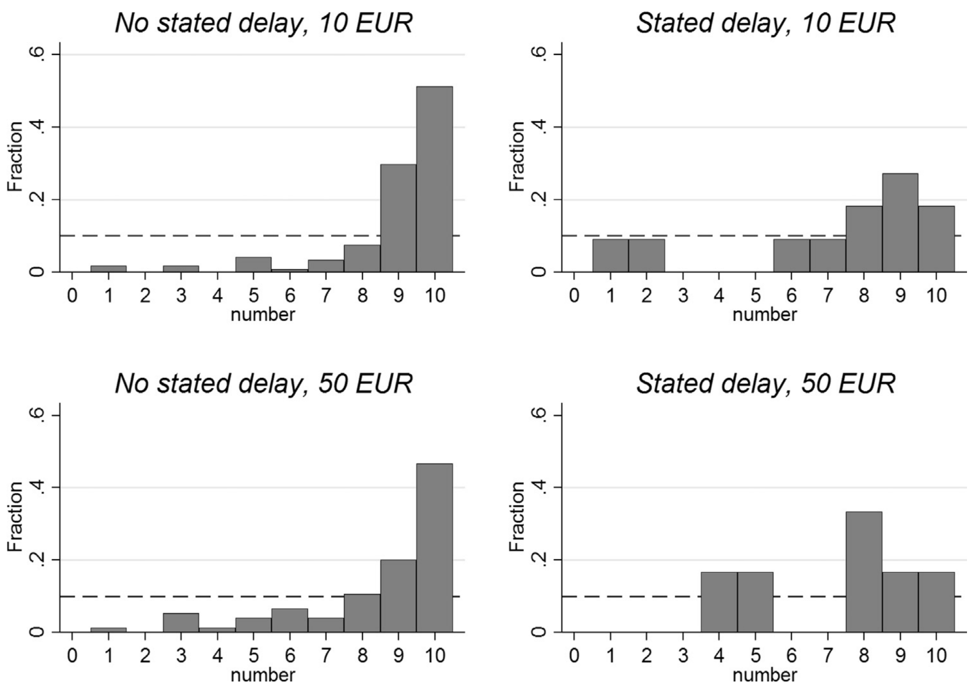

Fig. A2. Distribution of reported numbers in the cheating game (stated delay).
Notes: Number of observations amount to 121, 11, 75, and 6 in the no-stated-delay low-stakes, stated-delay low-stakes, no-stated-delay high-stakes, and stated-delay high-stakes groups, respectively.

(51.24%) report a 10 when the decision is not delayed. However, the maximal reporting drops to 18.18% after delaying the reporting decision. Overall, the reporting is significantly lower when participants delay their decision in the low-stakes treatment, with $p = 0.008$, Mann–Whitney U test.[15] However, note that only 11 participants delay their decision in the low-stakes cheating game, and thus a very low fraction of participants drive the result.

In the high-stakes treatments, on the other hand, delaying a decision has no significant effect on the level of cheating. Participants in the no-stated-delay high-stakes group report an average number of 8.43 (SD = 2.19), and in the stated-delay high-stakes group, the average number reported amounts to 7.33 (SD = 2.34). The distributions are not significantly different in a Mann–Whitney U test: $p = 0.129$.[16] Note the stated-delay high-stakes group contains only six participants.

[14] For the low-stakes dictator game, simple power calculation using a two-sided t-test for simplicity shows that we would need a sample size of 445 individuals per treatment cell to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects. For the high-stakes dictator game, the power shows that we would need a sample size of 5381 individuals per treatment cell.

[15] A simple power calculation using a two-sided t-test for simplicity shows that we need a sample size of 33 individuals per treatment cell to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects.

[16] A simple power calculation using a two-sided t-test for simplicity shows that we would need a sample size of 69 individuals per treatment cell to detect a significant effect if power is set to 80%, significance level is set to 5%, and given the responses of our subjects.

We also compare cheating in the low- and high-stakes treatments and find stakes have no significant effect on cheating, on conventional levels (see Table A2).

Table A2
Average reported numbers in the cheating game (stated delay).

	No stated delay	Stated delay	P-value
10 EUR	8.95	7.18	0.008
50 EUR	8.43	7.33	0.129
P-value	0.159	0.878	

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jebo.2017.10.012.

References

Abeler, J., Becker, A., Falk, A., 2014. Representative evidence on lying costs. J. Public Econ. 113, 96–104.
Abeler, J., Nosenzo, D., Raymond, C., 2017. Preferences for Truth-Telling. Technical Report. Oxford University.
Cappelen, A.W., Nielsen, U.H., Tungodden, B., Tyran, J.-R., Wengström, E., 2016. Fairness is intuitive. Exp. Econ. 19, 727–740, http://dx.doi.org/10.1007/s10683-015-9463-y.
Cardella, E., Chiu, R., 2012. Stackelberg in the lab: the effect of group decision making and cooling-off periods. J. Econ. Psychol. 33, 1070–1083, http://dx.doi.org/10.1016/j.joep.2012.07.004.
Cohn, A., Fehr, E., Marechal, A.M., 2014. Business culture and dishonesty in the banking industry. Nature 516 (7529), 86–089.
Cramton, P., Gunderson, M., Tracy, J., 1999. The effect of collective bargaining legislation on strikes and wages. Rev. Econ. Stat. 81, 475–487, http://dx.doi.org/10.1162/003465399558238.
Dreber, A., Johannesson, M., 2008. Gender differences in deception. Econ. Lett. 99, 197–199.
Erat, S., Gneezy, U., 2012. White lies. Manage. Sci. 58 (4), 723–733.
Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguise—an experimental study on cheating. J. Eur. Econ. Assoc. 11 (3), 525–547.
Forsythe, Robert, Horowitz, Joel L., Savin, N.E., Sefton, Martin, 1994. Fairness in simple bargaining experiments. Games Econ. Behav. 6 (3), 347–369, http://dx.doi.org/10.1006/game.1994.1021.
Fudenberg, Drew, Levine, David K., 2006. A dual-self model of impulse control. Am. Econ. Rev. 96 (5), 1449–1476.
Gneezy, Uri, Kajackaite, Agne, Sobel, Joel, 2017. Lying Aversion and the Size of the Lie. Mimeo.
Gneezy, U., 2005. Deception: the role of consequences. Am. Econ. Rev. 95 (1), 384–394.
Greiner, B., 2015. Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. J. Econ. Sci. Assoc. 1 (1), 114–125.
Grimm, Veronica, Mengel, Friederike, 2011. Let me sleep on it: delay reduces rejection rates in ultimatum games. Econ. Lett. 111 (2), 113–115.
Gul, Faruk, Pesendorfer, Wolfgang, 2008. The case for mindless economics. In: Caplin, Andrew, Shotter, Andrew (Eds.), The Foundations of Positive and Normative Economics. Oxford University Press.
Jiang, Ting., 2013. Cheating in mind games: the subtlety of rules matters. J. Econ. Behav. Organ. 93, 328–336.
Kahneman, Daniel, Knetsch, Jack L., Thaler, Richard H., 1986. Fairness and the assumptions of economics. J. Bus., S285–S300.
Kahneman, Daniel, 2003. Maps of bounded rationality: psychology for behavioral economics. Am. Econ. Rev. 93 (5), 1449–1475.
Kahneman, Daniel, 2011. Thinking, Fast and Slow. Farrar Straus Giroux http://scholar.google.com/scholar?cluster=3255681708785115121&hl=en&oi=scholarr.
Kajackaite, Agne, Gneezy, Uri, 2017. Incentives and cheating. Games Econ. Behav. 102, 433–444.
Kocher, Martin G., Martinsson, Peter, Myrseth, Kristian Ove R., Wollbrant, Conny E., 2016. Strong, bold, and kind: self-control and cooperation in social dilemmas. Exp. Econ., 1–26, http://dx.doi.org/10.1007/s10683-015-9475-7.
Lee, J., 2013. The impact of a mandatory cooling-off period on divorce. J. Law Econ. 56, 227–243, http://dx.doi.org/10.1086/667710.
Loewenstein, George., 2005. Hot-Cold empathy gaps and medical decision making. Health Psychol. 24 (4 Suppl), S49–S56, http://dx.doi.org/10.1037/0278-6133.24.4.S49.
Lohse, Johannes, Goeschl, Timo, Diederich, Johannes H., 2016. Giving is a question of time: response times and contributions to an environmental public good. Environ. Resour. Econ. (May), 1–23, http://dx.doi.org/10.1007/s10640-016-0029-z.
Lundquist, T., Ellingsen, T., Gribbe, E., Johannesson, M., 2009. The aversion to lying. J. Econ. Behav. Organ. 70 (1), 81–92.
Mazar, N., Amir, O., Ariely, D., 2008. The dishonesty of honest people: a theory of self-concept maintenance. J. Mark. Res. 45 (6), 633–644.
Neo, Siong, Yu, Michael, Weber, Roberto, 2013. The effects of time delay in reciprocity games. J. Econ. Psychol. 34, 20–35.
Nielsen, Ulrik H., Tyran, Jean-Robert, Wengström, Erik, 2014. Second thoughts on free riding. Econ. Lett. 122 (2), 136–139, http://dx.doi.org/10.1016/j.econlet.2013.11.021.
Oechssler, J., Roider, A., Schmitz, P.W., 2015. Cooling off in negotiations: does it work? J. Inst. Theor. Econ. JITE 171, 565–588, http://dx.doi.org/10.1628/093245615X14307212950056.
Piovesan, Marco, Wengström, Erik, 2009. Fast or fair? A study of response times. Econ. Lett. 105 (2), 193–196, http://dx.doi.org/10.1016/j.econlet.2009.07.017.
Rand, David G., Greene, Joshua D., Nowak, Martin A., 2012. Spontaneous giving and calculated greed. Nature 489 (7416), 427–430, http://dx.doi.org/10.1038/nature11467.
Rand, David G., Peysakhovich, Alexander, Kraft-Todd, Gordon T., Newman, George E., Wurzbacher, Owen, Nowak, Martin A., Greene, Joshua D., 2014. Social heuristics shape intuitive cooperation. Nat. Commun. 5 (April), 3677, http://dx.doi.org/10.1038/ncomms4677.
Shalvi, S., Dana, J., Handgraaf, M.J.J., De Dreu, C.K.W., 2011. Justifed ethicality: observing desired counterfactuals modifies ethical perceptions and behavior. Organ. Behav. Hum. Decis. Process. 115 (2), 181–190.
Sutter, M., 2009. Deception through telling the truth? Experimental evidence from individuals and teams. Econ. J. 119 (534), 47–60.