Experimental Economics
https://doi.org/10.1007/s10683-020-09644-0

ORIGINAL PAPER

(Not) alone in the world: Cheating in the presence of a virtual observer

Jantsje M. Mol[1,2] · Eline C. M. van der Heijden[1] · Jan J. M. Potters[1]

Received: 16 October 2018 / Revised: 14 January 2020 / Accepted: 18 January 2020

Abstract

We conducted an experiment in a high-immersive virtual reality environment to study the effect of the presence of a virtual observer on cheating behavior. Participants were placed in a virtual room and played 30 rounds of a cheating game without a chance of their cheating being detected. We varied whether or not a virtual observer (an avatar) was present in the room, and, if so, whether the avatar was actively staring at the decision maker or passively seated in a corner watching his smartphone. Results display significantly less cheating with an active than with a passive avatar, but not less cheating than in a control condition without an avatar. This suggests that an active (virtual) observer can intensify reputational concerns, but that the presence of someone passive and uninterested may actually alleviate such concerns.

Keywords Virtual reality experiment · Cheating · Watching eyes effect

JEL Classification B41 · C91 · C93

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10683-020-09644-0) contains supplementary material, which is available to authorized users.

✉ Jan J. M. Potters

Jantsje M. Mol
j.m.mol@vu.nl

[1] Tilburg School of Economics and Management, Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands

[2] Present Address: Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

Published online: 30 January 2020

1 Introduction

Studying cheating behavior is an advancing branch of experimental research in both economics and psychology.[1] It is motivated by the importance of cheating in many domains of life, and the advantages which experiments offer in terms of observation and control. Studies have shown that cheating behavior depends in important ways on context variables such as the payoff consequences to both self and others (Gneezy 2005), attention to moral standards (Mazar et al. 2008), procedural details (Jiang 2013), social information about others (Gino et al. 2009; Fosgaard et al. 2013), time pressure (Shalvi et al. 2012), and cognitive load (Mead et al. 2009). In the present paper, we examine one other context variable, namely the relevance of 'being watched'.

We study cheating behavior in a Virtual Reality (VR) environment.[2] Subjects play a version of the mind game, which is a variation of the die-under-cup paradigm, in which they have an incentive to be dishonest without any chance of getting caught (Jiang 2013). To investigate whether social control ('being watched') influences the prevalence of cheating, we vary the presence or absence of a virtual observer, that is, an avatar looking like a human. Moreover, we investigate whether the behavior of the virtual observer matters. In one treatment, the virtual observer is seated at some distance, watching his phone; in another treatment he stands close, actively observing the subject.

There is an extensive literature on the effect of observability and social control on pro-social behavior. A seminal study by Haley and Fessler (2005) reports that the mere display of stylized eyespots on the computer screen significantly increases giving in a dictator game. Although the robustness of the effect is a source of debate, the finding that cues of observation and control can affect pro-social behavior has been replicated many times in both field and lab.[3] In contrast to pro-social behavior, evidence on the effect of observability on cheating is scant. Cai et al. (2015) observe no effect of 'watching eyes' on cheating, while Oda et al. (2015) find that it reduces rates of pro-social lying. Oda et al. (2015) suggest that social context could be a possible explanation for the opposing findings of the studies; in their study subjects make decisions about donations to others, while in Cai et al. (2015) decisions concern only individual payoff. Another relevant and related study is Kroher and Wolbring (2015), who find that subjects are less likely to cheat when they conduct a

[1] For simplicity, we take cheating as more or less synonymous to lying, dishonesty, and deception. The literature is too vast for exhaustive referencing. Influential papers include Gneezy (2005), Mazar et al. (2008), and Fischbacher and Föllmi-Heusi (2013). For reviews, see Rosenbaum et al. (2014) and Abeler et al. (2019).

[2] Virtual Reality refers to 'a computer generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real way using special electronic equipment' (Oxford Dictionaries 2018).

[3] Sparks and Barclay (2013) mention ten lab studies that find an 'eyes effect' and five that do not. For evidence from field experiments see, e.g., Ekström (2012) and Ernest-Jones et al. (2011).

die-under-cup task in a booth with another subject compared to when doing the task alone.[4]

Based on the available evidence we hypothesize that the presence of a virtual observer will reduce the rate of cheating. The observer triggers feelings of being watched, which activates a concern for reputation and a desire to abide by the norm of honesty. Moreover, we hypothesize that this effect will be stronger in case the virtual observer is close by and actively staring at the decision maker as compared to the case that he is further away and minding his own business. This hypothesis is in line with Manesi et al. (2016) who find that the 'eyes effect' on prosocial behavior only occurs in case the eyes are really looking at the decision maker and not in case the eyes are closed or averted. We investigate whether a similar effect holds for cheating behavior.

Why do we conduct an experiment in a VR environment?[5] One reason is methodological. The use of avatars allows for more realistic and more immersive variations of observability than are possible with 'watching eyes' projected on a computer screen. Still, a VR environment offers as much control as a conventional lab experiment, and more control than a field experiment or the employment of confederates. The behavior and appearance of avatars can be controlled and varied at will, and the behavior of the decision maker can be observed and measured in all desired detail. A second reason is that VR settings are interesting in and of themselves. As technology continues to develop, artificial agents are expected to integrated into our daily lifes (Deng et al. 2019). We interact with these agents on-screen, but increasingly so also in virtual reality. Developments like Facebook Horizon will shift social interactions into social networks in VR. Also online shopping, customer services, and collaborative team environments are expected to be common in virtual reality (Lau and Lee 2019; Xi and Hamari 2019; Schroeder 2002; Pouliquen-Lardy et al. 2016). Also, there is a rapidly growing 'virtual world economy', in which participants of game platforms and social networks exchange virtual goods in real-world money. Transactional problems are bound to be prevalent in VR interactions, just as they are with transactions in cyberspace (Cohn et al. 2018; Papadopoulou et al. 2000). It is important to gain a better understanding of human behavior in VR settings, including the factors that affect propensenties to cheat and deceive. Ours is one of the first studies to contribute to this understanding.

We find significantly less cheating with an Active than with a Passive avatar, but not less cheating than in a control condition where No avatar is present. This suggests that the presence of an avatar can affect a decision maker in more ways than merely as a cue of social control.

[4] However, as the authors indicate, the effect of observability in this case may be confounded with the effects of social information and social preferences since the subjects could observe each others reporting behavior.

[5] Whereas VR experiments are now commonly applied as a research tool by psychologists and engineers, few VR experiments have been carried out by economists. A notable exception is Gürerk et al. (2019) who study virtual peer effects in a real-effort task. See Innocenti (2017) and Mol (2019) for extensive discussions of the potential of VR experiments for economics.

Fig. 1 Slot machine in virtual pub environment

Fig. 2 The three steps in the adapted mind game

2 Experimental design and procedure

2.1 Task and measure of cheating

Participants in our experiment played a version of the mind game (Jiang 2013), adapted to a VR setting. They were positioned in front of a slot machine with two slots, in a virtual pub environment (see Fig. 1). Each of the slots contained the numbers 5, 10 or 15. The elements were randomly generated by the software. Participants played the slot machine for 30 rounds.

Each round involved three steps; see Fig. 2. In step 1 participants were asked to decide (in their minds) which of the two slots, Left or Right, they wanted to count

Fig. 3 Avatar used in the *Passive* (left) and *Active* (right) treatment

for their earnings. In step 2, they could virtually pull the handle by pressing a button on their hand-held device. In step 3, after seeing the outcomes of the two slots, participants were asked to indicate which of the two slots, Left or Right, they had chosen in step 1 by pointing at the chosen slot with their hand-held device and clicking the button.

Note that participants could cheat by 'changing their mind' in step 3 of the process. Suppose, for example, that in step 1 a participant has chosen Right. If after step 2 the left slot displays a 10 and the right slot displays a 5, then honesty would require the participant to click Right in step 3. However, it would be tempting to cheat (or change one's mind) and click Left instead to maximize payoff. The slot machine was programmed such that each slot always displayed 5, 10, or 15, and the slots never displayed the same number. Summarizing, in each round there was a financial incentive to click either one of the two slots, and the size of this incentive (stake) was either 5 (10 vs. 5, 15 vs. 10) or 10 (15 vs. 5). At the end of the experiment, one round was randomly chosen and the outcome of that round (5, 10, or 15) was paid out in Euro.

As our measure of cheating we take the fraction of rounds in which a participant clicked the slot with the highest number in step 3. As in Jiang (2013), foresight f takes value 1 if a respondent chose the highest number and 0 otherwise. We then define average foresight F as the average of f over the 30 rounds of the individual. If a participant is always honest, the expected foresight is $F = 0.5$; if a participant always picks the slot with the highest number we have $F = 1$. Simple comparison to a binomial distribution $B(n = 30, p = 0.5)$ indicates that the probability that an honest participant attains a foresight level $F \geq 20/30 = 0.67$ is less than 5%. As we will see below, 60% of our participants attain at least that level of foresight.[6]

We are mainly interested in how the level of foresight varies with the presence or absence of a virtual observer. For that purpose we have three treatments, which were varied between subjects. In the baseline treatment (labeled No avatar) the environment looks as in Fig. 1; participants were alone in the virtual pub. In the Passive and

[6] In view of the possibility of downward lying our foresight measure may actually represent a lower bound on the extent of self-interested lying.

Active treatment, one virtual observer (avatar) was also present in the pub. The same (male) avatar was used in all rounds for all participants. In the Passive treatment (see the left panel of Fig. 3), the avatar was passively seated in a corner of the pub, busy with his mobile phone, wobbling with his feet. In the Active treatment (see the right panel of Fig. 3), the avatar was standing much closer, actively gazing at the participant.[7]

2.2 Equipment and procedure

The experiment was conducted in the DAF Technology Lab at Tilburg University in February and March 2017.[8] This lab consists of a Research Room and an Experience Room. The Research Room contains a reception desk and a number of tables and chairs. The Experience Room is five by five meters wide and includes eight Short Throw Projectors, a high speed position tracker and radio frequency active 3D glasses. The task was programmed in C sharp via Unity 3D 5.16 by the development team of the DAF Technology Lab. The virtual pub scenario was constructed with Blender 3D by a visual artist of the VR lab at Maastricht University.[9] It was used before in a study about aggression assessment (Lobbestael 2015) and the original author agreed to its use in the current experiment. To control the eyes and eye lashes of the avatar in a natural way, the asset Realistic Eye Movements was purchased from the Unity 3D Asset Store.[10] Some elements (including the virtual slot machine) were added by the development team of the DAF Technology Lab. Basic animations for the avatar (walk, turn, sit down) were adapted from Adobe Mixamo[11] and combined by the development team to suit the experimental scenario.

Participants were recruited by email, using the participant database of CentER-lab, one week in advance of the experiment. They were randomly assigned to one of the treatments. The invitation email did not mention VR or any other referral to a non-standard lab experiment, to prevent a selection bias of gamers. In total, 121 people participated. The responses of three participants could not be recorded due to a technical failure. This left 118 responses for the analysis.

Upon arrival in the Reception Room, participants were asked to carefully read and sign the informed consent form (see Online Appendix B). Participants with a high risk for simulator sickness, including migraine patients and epileptic patients, would be excluded from the experiment. All registered participants were qualified as having low risk for simulator sickness, so no participant was excluded from taking part. Next, participants were instructed about the experiment, including the random

[7] The avatar could follow and gaze at the participants because participants were wearing stereographic glasses with motion trackers.

[8] We ran a pilot experiment in May 2016 with 57 participants to check the set-up of the game and the expected length of the experiment. We were not confronted with any major technical issues or simulator sickness and participants understood the game.

[9] See http://www.richardbenning.nl/ for a portfolio.

[10] https://assetstore.unity.com/packages/tools/animation/realistic-eye-movements-29168.

[11] http://www.mixamo.com/.

Fig. 4 The field of view of subjects in the Active treatment

number generator of the virtual slot machine, by reading the instruction sheet (see Online Appendix C). After entering the Experience room, participants conducted four trial rounds to become familiar with the procedure and the equipment. To minimize simulator sickness and to decrease variability due to a different field of view, participants were seated in a chair. The chair was positioned under an angle such that the Passive avatar was far away and the Active avatar was impressively close to the participant (see Fig. 4).

After finishing the 30 rounds, participants left the Experience Room and were seated in the Reception Room to fill in a short post-experimental questionnaire, including questions on age, gender, simulator sickness and presence.[12] Lastly, the experimenter collected the completed questionnaire, randomly determined one round, and paid the participant the corresponding payoff in cash. Total session time including instructions, questionnaire and payment was approximately 20 minutes. On average, participants earned 12.05 Euro.

3 Results

Table 1 presents some descriptive experimental and questionnaire data: (1) time in minutes to complete the rounds in the Experience Room, (2) the proportions of rounds participants clicked the right (versus left) slot, (3) the proportions of times they faced a high stake of 10 Euro rather than a low stake of 5 Euro, (4) the total presence score, (5) the total sickness score, (6) age, (7) gender, (8) doing a major in economics, (9) being religious, and (10) pursuing a Masters degree (versus a Bachelors). The most important message to take from this, is that there are no differences

[12] To minimize demand effects, the simulator sickness questionnaire (SSQ) was not administered until after the experiment (Young et al. 2006). For the questions on presence, the Igroup Presence Questionnaire (IPQ) was used (Schubert et al. 2001).

Table 1 Descriptive statistics

	Active	Passive	No avatar	Total
Time in minutes	9.13	9.18	9.22	9.18
	(0.82)	(0.59)	(0.79)	(0.74)
Proportion right side chosen	0.52	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)
Proportion high stake	0.33	0.33	0.31	0.32
	(0.47)	(0.47)	(0.46)	(0.47)
Total presence score [0–84]	39.64	35.46	36.98	37.36
	(13.46)	(11.70)	(12.12)	(12.56)
Total sickness score [0–180]	26.56	28.00	30.76	28.46
	(30.11)	(25.66)	(29.97)	(28.71)
Age in years	23.40	23.26	21.88	22.83
	(3.04)	(4.72)	(2.80)	(3.68)
Gender (1 = female)	0.59	0.54	0.55	0.56
	(0.49)	(0.50)	(0.50)	(0.50)
Economics major	0.74	0.69	0.78	0.74
	(0.44)	(0.46)	(0.42)	(0.44)
Religious	0.67	0.54	0.70	0.64
	(0.47)	(0.50)	(0.46)	(0.48)
Master student	0.67	0.46	0.53	0.55
	(0.47)	(0.50)	(0.50)	(0.50)
N	39	39	40	118

Table displays means, stdev in parentheses

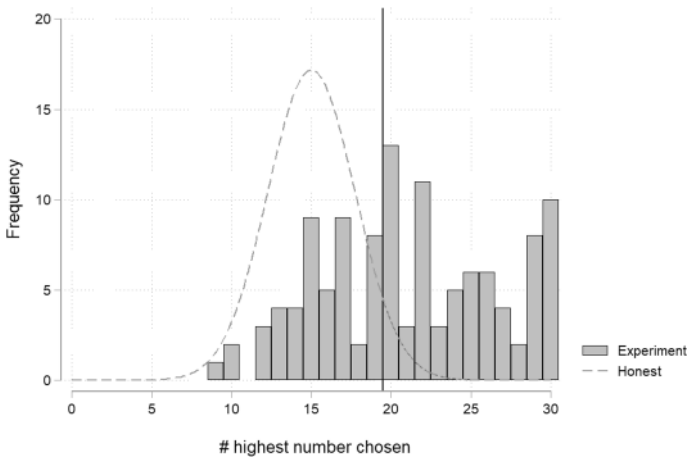

Fig. 5 Distribution of total foresight in experiment in relation to honest players *Note*: Reference line shows the threshold for statistically significant cheating ($F = 0.67$)

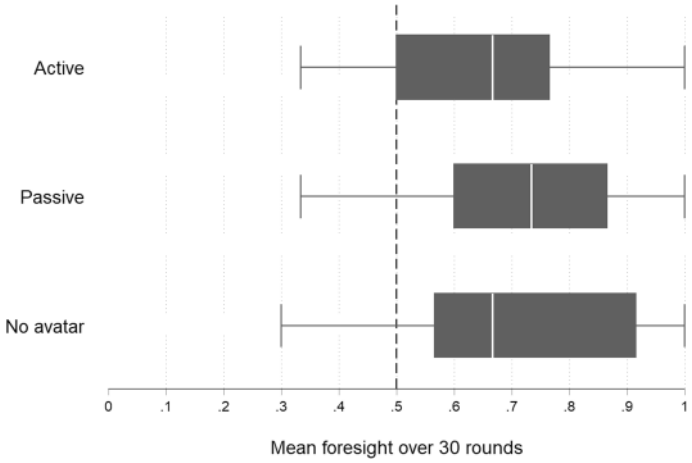

Fig. 6 Boxplot of mean foresight *F* over 30 rounds per treatment *Note*: White lines indicate medians, boxes indicate interquartile ranges, dashed reference line indicates perfect honesty

between the three treatments on any of these variables.[13] Later, in the regression analysis, we will also use these variables as control variables.

Before we analyze treatment differences, we first examine whether subjects do in fact cheat in our VR mind game. For that purpose we compare participants' observed levels of foresight to the distribution of foresight[14] levels we would expect if participants were honest. A normal approximation of the latter distribution is displayed in Fig. 5, alongside the histogram of individual foresight levels in the experiment.[15] The data indicate that 60% (71/118) of the participants have a foresight level of 0.67 or higher: in at least 20 of the 30 rounds they click on the slot with the highest number. If a participant is honest, the probability that s/he attains this level of foresight is less than 5%. So, we can safely conclude that there is massive cheating in our experiment. The average level of foresight in our experiment is even somewhat higher than the foresight level reported in Jiang (2013) for a desktop version of the mind game. It appears that a VR environment does not stimulate people to become honest (or to behave randomly, as that might also look like being honest).

Now we come to our main question: does the presence of the avatar affect the level of cheating? Fig. 6 displays boxplots of the foresight levels in the three

[13] Kruskall-Wallis tests indicated no significant differences for any of these variables, except for a small difference in age ($p = 0.085$).

[14] In Online Appendix A we reran the main analysis with an alternative dependent variable that captures both cheating gain, as well as strategic cheating loss. The results do not change.

[15] We present separate density plots for each of the three treatments in Figure A1 in Online Appendix A. They display some differences between the treatments. In No avatar there is a somewhat larger fraction of subjects (0.23) with almost perfect foresight than in Active (0.10) and Passive (0.13). The dispersion also seems to be a bit larger in No avatar than in Active and Passive. However, these differences are not statistically significant.

Table 2 Random effects GLS regression on panel data of foresight *f*

	Dependent variable: foresight *f*			
	(1)	(2)	(3)	(4)
Passive avatar	0.025	0.024	0.026	0.002
	(0.042)	(0.043)	(0.042)	(0.040)
Active avatar	− 0.054	− 0.055	− 0.057	− 0.066
	(0.043)	(0.043)	(0.043)	(0.043)
Round		0.002***	0.002***	0.002***
		(0.001)	(0.001)	(0.001)
High stake		0.052***	0.052***	0.052***
		(0.017)	(0.017)	(0.017)
Right slot chosen		− 0.022	− 0.022	− 0.022
		(0.014)	(0.014)	(0.014)
Time in lab in minutes			− 0.026	− 0.017
			(0.024)	(0.025)
Total presence score			0.001	0.001
			(0.001)	(0.001)
Total sickness score			0.001	0.001
			(0.001)	(0.001)
Age in years				0.011**
				(0.005)
Gender (1 = female)				0.064*
				(0.035)
Economics major				− 0.005
				(0.035)
Religious				− 0.029
				(0.038)
Master student				− 0.081*
				(0.044)
Observations	3540	3540	3540	3540
Nr of subjects	118	118	118	118

GLS with random effects. Standard errors clustered at individual level in parentheses (*(p < 0.1), **(p < 0.05), ***(p < 0.01))

treatments.[16] We find the following ordering for the mean levels of foresight: $F_{Passive} = 0.74 > F_{No\ avatar} = 0.71 > F_{Active} = 0.66$. The difference between $F_{Passive}$ and $F_{No\ avatar}$ is not statistically significant ($p = 0.267$, Wilcoxon ranksum test), nor is the difference between $F_{No\ avatar}$ and F_{Active} ($p = 0.562$). The only significant

[16] In Online Appendix A we show that an analysis of the cumulative density functions displays the same pattern. Foresight levels of the Passive treatment first-order stochastically dominate those of the Active treatment (and are significantly different with a Kolmogornov-Smirnov test), while the CDF of the No avatar treatment is in between the Passive and the Active treatment (and not significantly different from either one).

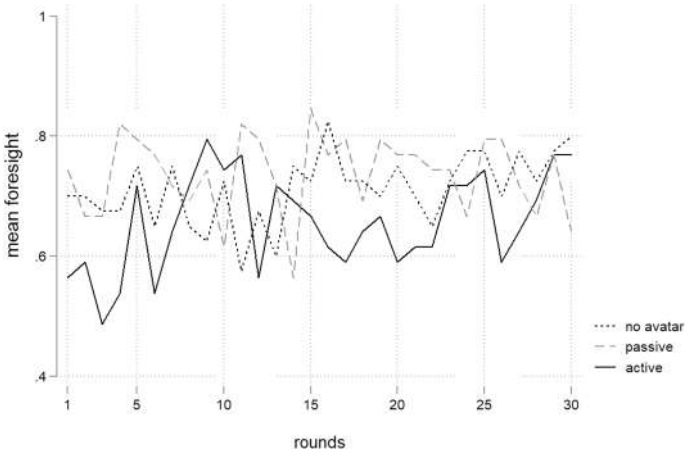

Fig. 7 Trend of foresight F over 30 rounds, by treatment

treatment difference is between $F_{Passive}$ and F_{Active} ($p = 0.029$). We can conclude that the presence of the avatar by itself does not encourage honesty and reduces cheating, contrary to our hypothesis. Foresight levels with an avatar, Active or Passive, are not significantly different from those without an avatar. At the same time, the presence of an Active avatar significantly decreases the level of foresight relative to the presence of a Passive avatar. So, even though his mere presence does not seem to matter, the behavior of the avatar does matter. Below we discuss the interpretation of these results. First we look at the development of foresight over the rounds, and examine the robustness of the treatment effect with a parametric analysis in which we control for several covariates.

Table 2 presents the results of linear panel regressions in which the binary foresight f of a participant in a round is taken as the dependent variable. We use Generalized Least Squares (GLS) regressions with standard errors clustered at subject level. Model (1) only uses dummy variables for the Active and the Passive treatment, using the No avatar treatment as the benchmark. The coefficients for Active and Passive are not significantly different from zero. However, the two coefficients are significantly different from each other ($p = 0.036$, Wald test), in line with the results from the non-parametric test discussed above. Model (2) adds several round-specific experimental control variables to the regression: the round number, the stakes in the round, and whether the right or the left slot was chosen. We find that the round number has a significant effect on foresight, suggesting more cheating in later rounds. Foresight also increases significantly with the money at stake, and is higher with a stake of 10 Euro rather than 5 Euro. Adding these variables, however, hardly has an effect on the estimated treatment effects as the estimated coefficients for Active and Passive remain significantly different ($p = 0.036$, Wald test). Model (3) adds three individual-specific experimental variables: time in the lab, presence score, and sickness score. None of these turn out to have an effect on foresight. A Wald test for equality of coefficients shows that the difference between Active and

Passive is still significant ($p = 0.024$). Finally, model (4) adds demographic variables. We observe that foresight correlates positively with age and being female, and negatively with being a Master student rather than a Bachelor student. We do not wish to put too much value on these coefficients. Most important for our purpose is that controlling for these variables does not much change the overall picture regarding the treatment effects ($p = 0.023$, Wald test).

Figure 7 depicts the development of the average foresight levels over the rounds. We can discern a slightly increasing trend, especially in No avatar and Active. Such an increasing trend is in line with the evidence reported in Abeler et al. (2019). An interesting question is whether such an increase can be the result of a 'slippery slope' of cheating. Does it become easier to cheat if one has done it before? To examine this we estimate a dynamic panel model in which foresight in a round is allowed to depend on foresight in the previous round. Because of the endogenous nature of the lagged foresight variable, we rely on the GMM estimator proposed by Arellano and Bond (1991). The results show a negative coefficient for the lagged foresight level (see Table A2 in Online Appendix A). Having cheated in the previous round (or having been lucky) reduces the propensity to cheat in the current round. This points toward moral licensing according to which 'good' behavior in the past is compensated by 'bad' behavior now (Blanken et al. 2015; Clot et al. 2014), and moral balancing, according to which 'bad' behavior in the past is compensated by 'good' behavior now (Mazar and Zhong 2010).[17]

We find that foresight levels are significantly higher when the stakes are higher. This finding is in line with models assuming that individuals make a trade-off between the benefits of cheating and a preference for being seen as honest. In line with these models, a natural explanation for the impact of observational cues is that they strengthen the preference for being seen as honest. This is why foresight levels are lower with an Active avatar. It is not immediately clear though how stakes would interact with this treatment effect. If the two utility components—payoffs and reputation for honesty—are separable then they are not predicted to interact. Such separability is assumed in most models (e.g., Khalmetski and Sliwka 2019; Gneezy et al. 2018). In line with these models, we do not find a significant interaction effect between stakes and the treatment variables (see Table A1 in Online Appendix A.[18]

4 Concluding discussion

We carried out one of the first economic experiments in a virtual reality lab. In particular, we have examined the effect of the presence of a virtual observer on cheating behavior. In the experiment, subjects played 30 rounds of an adapted version of the

[17] The estimated effect of lagged foresight on current foresight is almost the same for each of the three treatments. This suggests that the cognitive process underlying these dynamics does not differ substantially between the treatments.

[18] Another variable that significantly affects foresight levels is gender; women seem to cheat more than men. It is noticeable that the avatar is male, so it is conceivable that the effect of the avatar is gender-specific. However, we find no evidence for such an interaction effect.

mind game in a virtual pub environment. The treatments concerned the presence of a virtual observer: not present, passively seated in the corner or actively staring at the participant. We hypothesized that the presence of a virtual observer would reduce the rate of cheating, and that the effect would be stronger with an active observer. Our experimental results show that the presence of the observer does not have a significant effect on the propensity of individuals to cheat in comparison to an environment without an avatar. However, given that the avatar is present, it matters significantly how the avatar behaves. Cheating is lower in the Active than in the Passive treatment.

In a thorough review, Abeler et al. (2019) show that one of the main motivations for truth-telling is a preference for being seen as honest. Individuals care about their reputation for being a truthful person.[19] This conclusion is partly based on the finding that there is more truth-telling in case the experimenter can observe the true state than in case the true state is private information of the subject. We extend this result in two important directions. First, cues of being observed can matter even in case the true state is private information. In our experiment, the state is in the subject's mind and cannot be observed by others. Second, the observer does not have to be a real person; even a virtual observer can affect behavior. This suggests that a preference for being seen as honest can be prompted by subtle observational cues, which may partly operate at the subconscious level (Conty et al. 2016). Our results provide a step in exploring the boundary conditions of reputational concerns for being honest. They signify that the weight of this concern is not a zero-one variable that is either off or on; it is more likely to be a continuous variable that is affected by contextual cues which vary in shape and scope.

Our finding that the presence of an Active avatar reduces cheating relative to the presence of a Passive avatar is in line with the hypothesis that stronger cues trigger stronger responses. In particular, direct gaze signals someone's attention which initiates a sense of being watched and activates a reputational concern, while averted eyes indicate inattention and indifference (Hietanen et al. 2018; Manesi et al. 2016; Vaish et al. 2017). This may also explain why there is more cheating in the Passive treatment than in the No avatar treatment (even though the difference is not significant). The avatar is focused on his phone and does not show any interest in his surroundings. The fact that someone is present who could in principle observe you and interact with you, but clearly does not, may activate a feeling of not being watched. This sense of privacy and anonymity can reduce a concern for reputation and may lead to more cheating relative to a setting in which no-one is present (Ayal et al. 2015).[20]

Another potential explanation for why a Passive avatar induces more cheating relative to No avatar is that the avatar distracts the participants. Some studies

[19] Many recent models feature such a preference, including Frankel and Kartik (2019), Khalmetski and Sliwka (2019) and Gneezy et al. (2018).

[20] This kind of counterfactual thinking is not novel (Miller et al. 2005). It is similar to the 'license to cheat effect' in Gino et al. (2013), where a setting with voluntary regulation leads to more misreporting than a setting without regulation. Not being regulated, whereas you could be, may prompt a stronger sense of anonymity than regulation not being an issue in the first place. See also Keizer et al. (2011) for a reverse effect of ineffective prohibition signs.

suggest that depletion of attention may lead to more cheating, as overcoming the impulse to cheat requires attention, and self-control (Mead et al. 2009; Gino et al. 2009; Pitesa et al. 2013). The effect is highly contested though (Greene et al. 2004; Capraro 2017; Lohse et al. 2018; Wibral et al. 2012; Suchotzki et al. 2017). Some of our results also speak against this explanation. Cognitive load and lack of attention are usually associated with a reduced sensitivity to other pieces of information (Gilbert 1998; Hirshleifer et al. 2009; Deck and Jahedi 2015). We find no evidence for this in our experiment. For example, the impact of the stake size on the level of foresight and the dynamics of foresight over the rounds do not differ across the treatments. So, we find no indication that the presence of an avatar, Passive or Active, distracts attention and affects the role of other cues and characteristics that impact participants' decisions.

We used VR to study how visual cues of being watched affect dishonesty. We believe this could not have been done as effectively with conventional methods. One alternative would be to use peer participants as observers. This, however, may come with additional nuisance variables such as the physical appearances, facial expressions, or verbal utterances of the observers (Kroher and Wolbring 2015; van de Ven and Villeval 2015). Moreover, the behavior of peer observers cannot be varied in the controlled way that is possible with avatars. One may use confederates to allow for this, but it is questionable whether they can behave consistently. Moreover, some subjects may suspect that they interact with a confederate which can also impede control (Hietanen et al. 2018). An alternative method is to rely on the display of prerecorded videos, pictures, or pictorials of watching eyes (see e.g. Nasiopoulos et al. 2015, for a discussion of different stimuli on attention). This ensures control, but comes at the expense of immersiveness and realism (Reader and Holmes 2016). Therefore, it is questionable whether static observational cues are strong enough to consistently affect behavior (Fehr and Schneider 2010; Dear et al. 2019; Pfattheicher et al. 2019; Oda et al. 2015; Cai et al. 2015). Indeed, evidence indicates that the behavioral, physiological, and neurological responses to the presence of an avatar are more representative for the responses to human presence than are those of other simulated social presences such as videotapes, images or pictorials (Hartmann et al. 2010; Oh et al. 2018; Pan and Hamilton 2018; Rubo and Gamer 2018; Yaremych and Persky 2019). This offers an important advantage of our VR setting. In sum, our study on the impact of social observation combines experimental control and ecological validity to a degree that cannot be obtained with other methods. This combination of control and realism may offer unique advantages to other economic research areas as well. Controlled variations in identity, appearance and proximity create new possibilities in studies of discrimination (Peck et al. 2013) and the use of virtual humans avoids the reflection problem in studies of peer effects (Gürerk et al. 2019). Realistic but controllable environments further allow for strong emotion induction in studies on moral judgements (Kugler et al. 2019) and for measurement

of responses that would be impossible or unethical to obtain in any other way, such as evacuation behavior of non-experts (Kinateder et al. 2014) and the trolley problem (Navarrete et al. 2012). For a more extensive review of the possibilities and drawbacks for high-immersive VR experiments in economics the reader is referred to Mol (2019).

Besides the methodological advantages, we believe that VR settings are valuable economic environments in which humans are expected to interact more frequently in the near future. Our study is one of the first to contribute to a better understanding of human behavior in response to avatars in VR. We find significantly less cheating with an Active than with a Passive avatar, but not less cheating than in a control condition where No avatar is present. This suggests that an active (virtual) observer can intensify reputational concerns, but that the presence of a passive (virtual) person may mitigate these concerns.

Acknowledgements We would like to thank conference participants at TIBER 2017, VR Workshop Aachen 2017 and M-BEES 2018 for valuable comments. We would also like to thank the development team of the DAF Technology lab: Anton Sluijtman, Rob Vossen, Hans van den Dool, Erwin Peters and Dineke Kolen. The DAF Technology Lab is partially funded by a philanthropic donation from the PACCAR foundation/DAF Trucks, and with funding from the European Union, OP Zuid, the Ministry of Economic Affairs, the Province of Noord-Brabant and the municipalities of Tilburg and Gilze Rijen (PROJ-00076) awarded to Tilburg University. The usual exculpations apply.

References

Abeler, J., Nosenzo, D., & Raymond, C. (2019). Preferences for truthtelling. *Econometrica*, *87*(4), 1115–1153.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, *58*(2), 277.

Ayal, S., Gino, F., Barkan, R., & Ariely, D. (2015). Three principles to REVISE peoples unethical behavior. *Perspectives on Psychological Science*, *10*(6), 738–741.

Blanken, I., van de Ven, N., & Zeelenberg, M. (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin*, *41*(4), 540–558.

Cai, W., Huang, X., Wu, S., & Kou, Y. (2015). Dishonest behavior is not affected by an image of watching eyes. *Evolution and Human Behavior*, *36*(2), 110–116.

Capraro, V. (2017). Does the truth come naturally? Time pressure increases honesty in one-shot deception games. *Economics Letters*, *158*, 54–57.

Clot, S., Grolleau, G., & Ibanez, L. (2014). Smug alert! Exploring self-licensing behavior in a cheating game. *Economics Letters*, *123*(2), 191–194.

Cohn, A., Gesche, T., & Maréchal, M. (2018). Honesty in the digital age. *University of Zurich Working Paper Series*, *280*, 1–25.

Conty, L., George, N., & Hietanen, J. K. (2016). Watching eyes effects: When others meet the self. *Consciousness and Cognition*, *45*(October 2016), 184–197.

Dear, K., Dutton, K., & Fox, E. (2019). Do watching eyes influence antisocial behavior? A systematic review & meta-analysis. *Evolution and Human Behavior*, *40*(3), 269–280.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review*, *78*, 97–119.

Deng, E., Mutlu, B., & Mataric, M. J. (2019). Embodiment in socially interactive robots. *Foundations and Trends in Robotics*, *7*(4), 251–356.

Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*(3), 530–546.

Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: A field experiment. *Evolution and Human Behavior*, *32*(3), 172–178.

Fehr, E., & Schneider, F. (2010). Eyes are on us, but nobody cares: Are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, *277*(1686), 1315–1323.

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

Fosgaard, T. R., Hansen, L. G., & Piovesan, M. (2013). Separating Will from Grace: An experiment on conformity and awareness in cheating. *Journal of Economic Behavior and Organization*, *93*, 279–284.

Frankel, A., & Kartik, N. (2019). Muddled information. *Journal of Political Economy*, *127*(4), 1739–1776.

Gilbert, D. T. (1998). Speeding with Ned: A personal view of the correspondence bias. In J. M. D. J. Cooper (Ed.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 5–66). Washington, DC: APA Press.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior. *Psychological Science*, *20*(3), 393–398.

Gino, F., Krupka, E. L., & Weber, R. A. (2013). License to cheat: Voluntary regulation and ethical behavior. *Management Science*, *59*(10), 2187–2203.

Gneezy, U. (2005). Deception: The role of consequences. *The American Economic Review*, *95*(1), 384–394.

Gneezy, U., Kajackaite, A., & Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, *108*(2), 419–453.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Gürerk, Ö., Bönsch, A., Kittsteiner, T., & Staffeldt, A. (2019). Virtual humans as co-workers: A novel methodology to study peer effects. *Journal of Behavioral and Experimental Economics*, *78*, 17–29.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? *Evolution and Human Behavior*, *26*(3), 245–256.

Hartmann, T., Toz, E., & Brandon, M. (2010). Just a game? Unjustified virtual violence produces guilt in empathetic players. *Media Psychology*, *13*(4), 339–363.

Hietanen, J. O., Syrjämäki, A. H., Zilliacus, P. K., & Hietanen, J. K. (2018). Eye contact reduces lying. *Consciousness and Cognition*, *66*(September), 65–73.

Hirshleifer, D., Lim, S. S., & Teoh, S. H. (2009). Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance*, *64*(5), 2289–2325.

Innocenti, A. (2017). Virtual reality experiments in economics. *Journal of Behavioral and Experimental Economics*, *69*(49), 71–77.

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

Keizer, K., Lindenberg, S., & Steg, L. (2011). The reversal effect of prohibition signs. *Group Processes and Intergroup Relations*, *14*(5), 681–688.

Khalmetski, K., & Sliwka, D. (2019). *Disguising lies: Image concerns and partial lying in cheating games* (pp. 1–38). American Economic Journal: Microeconomics.

Kinateder, M., Müller, M., Jost, M., Mühlberger, A., & Pauli, P. (2014). Social influence in a virtual tunnel fire: Influence of conflicting information on evacuation behavior. *Applied Ergonomics*, *45*(6), 1649–1659.

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

Kugler, T., Ye, B., Motro, D., & Noussair, C.N. (2019). On trust and disgust: Evidence from face reading and virtual reality. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550619856302.

Lau, K. W., & Lee, P. Y. (2019). Shopping in virtual reality: A study on consumers shopping experience in a stereoscopic virtual reality. *Virtual Reality*, *23*(3), 255–268.

Lobbestael, J. (2015). Challenges in aggression assessment: The gap between self-report and behavior, and a call for new valid behavioral paradigms. *Journal of Socialomics*, *05*(01), 5–6.

Lohse, T., Simon, S. A., & Konrad, K. A. (2018). Deception under time pressure: Conscious decision or a problem of awareness? *Journal of Economic Behavior and Organization*, *146*, 31–42.

Manesi, Z., Van Lange, P. A. M., & Pollet, T. V. (2016). Eyes wide open: Only eyes that pay attention promote prosocial behavior. *Evolutionary Psychology*, *14*(2), 1–15.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, *45*(6), 633–644.

Mazar, N., & Zhong, C. B. (2010). Do green products make us better people? *Psychological Science*, *21*(4), 494–498.

Mead, N. L., Baumeister, R. F., Gino, F., Schweitzer, M. E., & Ariely, D. (2009). Too tired to tell the truth: Self-control resource depletion and dishonesty. *Journal of Experimental Social Psychology*, *45*(3), 594–597.

Miller, D. T., Visser, P. S., & Staub, B. D. (2005). How surveillance begets perceptions of dishonesty: The case of the counterfactual sinner. *Journal of Personality and Social Psychology*, *89*(2), 117–128.

Mol, J. M. (2019). Goggles in the lab: Economic experiments in immersive virtual environments. *Journal of Behavioral and Experimental Economics*, *79*(C), 155–164.

Nasiopoulos, E., Risko, E. F., & Kingstone, A. (2015). Social attention, social presence, and the dual function of gaze. In A. Puce & B. I. Bertenthal (Eds.), *The many faces of social attention: Behavioral and neural measures* (pp. 129–155). Cham: Springer.

Navarrete, C. D., Mcdonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional trolley problem. *Emotion*, *12*(2), 364–370.

Oda, R., Kato, Y., & Hiraishi, K. (2015). The watching-eye effect on prosocial lying. *Evolutionary Psychology*, *13*(3), 1–5.

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers Robotics AI*, *5*(OCT), 1–35.

Oxford Dictionaries,. (2018). *Virtual reality*. https://www.oxforddictionaries.com/definition/english/virtual-reality. Accessed September 01, 2018.

Pan, X., & Hamilton, A. F. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417.

Papadopoulou, P., Andreou, A., Kanellis, P., & Martakos, D. (2000). Enabling trust formation within agent-mediated virtual environments. *Virtual Reality*, *5*(3), 149–159.

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–87.

Pfattheicher, S., Schindler, S., & Nockur, L. (2019). On the impact of honesty-humility and a cue of being watched on cheating behavior. *Journal of Economic Psychology*, *71*(October 2017), 159–174.

Pitesa, M., Thau, S., & Pillutla, M. M. (2013). Cognitive control and socially desirable behavior: The role of interpersonal impact. *Organizational Behavior and Human Decision Processes*, *122*(2), 232–243.

Pouliquen-Lardy, L., Milleville-Pennel, I., Guillaume, F., & Mars, F. (2016). Remote collaboration in virtual reality: asymmetrical effects of task distribution on spatial processing and mental workload. *Virtual Reality*, *20*(4), 213–220.

Reader, A. T., & Holmes, N. P. (2016). Examining ecological validity in social interaction: Problems of visual fidelity, gaze, and social potential. *Culture and Brain*, *4*(2), 134–146.

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

Rubo, M., & Gamer, M. (2018). Virtual reality as a proxy for real-life social attention? In *Proceedings of the 2018 ACM symposium on eye tracking research & applications - ETRA '18*, (pp. 1–2). New York: ACM Press.

Schroeder, R. (2002). *The social life of avatars: Presence and interaction in shared virtual environments*. Berlin: Springer.

Schubert, T. A., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence*, *10*(3), 266–281.

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

Sparks, A., & Barclay, P. (2013). Eye images increase generosity, but not for long: The limited effect of a false cue. *Evolution and Human Behavior*, *34*(5), 317–322.

Suchotzki, K., Verschuere, B., Bockstaele, B. V., Ben-Shakhar, G., & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin*, *143*(4), 428–453.

Vaish, A., Kelsey, C. M., Tripathi, A., & Grossmann, T. (2017). Attentiveness to eyes predicts generosity in a reputation-relevant context. *Evolution and Human Behavior*, *38*(6), 729–733.

van de Ven, J., & Villeval, M. C. (2015). Dishonesty under scrutiny. *Journal of the Economic Science Association*, *1*(1), 86–99.

Wibral, M., Dohmen, T., Klingmüller, D., Weber, B., & Falk, A. (2012). Testosterone administration reduces lying in men. *PLoS ONE*, *7*(10), e46774.

Xi, N., & Hamari, J. (2019). VR shopping: A review of literature. In *Twenty-fifth Americas conference on information systems, Cancun, 2019 1*, (p. 10).

Yaremych, H. E., & Persky, S. (2019). Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. *Journal of Experimental Social Psychology*, *85*(April), 103845.

Young, S. D., Adelstein, B. D., & Ellis, S. R. (2006). Demand characteristics of a questionnaire used to assess motion sickness in a virtual environment. *Proceedings IEEE Virtual Reality*, *2006*, 13.