Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2020/18

The Influence of Self and Social Image Concerns on Lying

Zvonimir Bašić
Simone Quercia

MAX PLANCK
SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods 2020/18

The Influence of Self and Social Image Concerns on Lying

Zvonimir Bašić / Simone Quercia

August 2020

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
https://www.coll.mpg.de

The Influence of Self and Social Image Concerns on Lying*

Zvonimir Bašić and Simone Quercia†

July 29, 2020

Abstract

We investigate the influence of self and social image concerns as potential sources of lying costs. In a standard die-rolling experiment, we exogenously manipulate self-awareness and observability, which mediate the focus of a person on their private and public selves, respectively. First, we show that an increase in self-awareness has no effect on reporting private information. This suggests that self-image concerns may be less important than previously hypothesized in the literature on lying costs. Second, we show that increasing subjects' observability, while still maintaining private information, significantly decreases the subjects' reports. We finally show in a survey experiment that respondents believe that the likelihood of a lie increases with the reported outcome and attribute negative traits to people who make high reports. This further supports reputational concerns as the explanation behind the results of our social image treatment.

Keywords: honesty, truth-telling, lying, private information, self-image concerns, social image concerns, reputation

JEL Classification: C91, D63, D82, D91

*We thank Johannes Abeler, Stefania Bortolotti, Armin Falk, Agne Kajackaite, Sebastian Kube, Daniele Nosenzo, Matthias Praxmarer, Stefan Trautmann, and Matthias Wibral as well as workshop and conference participants at Lund University, Maastricht University, University of East Anglia, Humboldt University, and the Vienna University of Economics and Business for helpful comments. We gratefully acknowledge financial support from Leibniz Programme of the German Research Foundation (DFG) and Institute on Behavior and Inequality (briq).

†Zvonimir Bašić: Max Planck Institute for Research on Collective Goods, Bonn (e-mail: basic@coll.mpg.de); Simone Quercia: University of Verona (e-mail: simone.quercia@univr.it).

1 Introduction

In recent years, a considerable number of studies have shown that people experience psychological lying costs as they refrain from lying even when this increases their payoff (e.g., Gneezy, 2005; Mazar et al., 2008; Shalvi et al., 2011a; Fischbacher and Föllmi-Heusi, 2013; Abeler et al., 2014; Gächter and Schulz, 2016). While several explanations have been proposed to characterize lying costs, recent papers show that combining a preference for being honest (intrinsic lying costs) with a preference for being seen as honest (reputation or social image costs) can reconcile the existing empirical findings (Abeler et al., 2019; Gneezy et al., 2018; Khalmetski and Sliwka, 2019).[1] While social image costs arise from the person's desire to appear honest in the eyes of others, the underlying psychological motives for intrinsic costs are debated in the literature. One important view is that these costs might originate from self-image concerns, that is, the desire to think of oneself as an honest person (Mazar et al., 2008; Shalvi et al., 2011a; Fischbacher and Föllmi-Heusi, 2013).

In this paper, we investigate the influence of these two notions of image concerns on lying behavior, using the die-rolling paradigm introduced by Fischbacher and Föllmi-Heusi (2013). In this setup, subjects are given a six-sided die, they are asked to roll it in private, and to report the outcome to the experimenter. Payoffs are generally increasing in the report. While lies are not detectable at the individual level, they can be inferred at the group level comparing the distribution of reports with the expected distribution of die rolls.

To make self and social image concerns salient, we exogenously manipulate self-awareness and observability, which direct the subjects' focus on their private and public selves, respectively. To manipulate self-awareness, we expose subjects to a real-time video of their face on the computer screen, i.e., we expose them to their "self-image", as in Falk (2018). To manipulate observability, we expose subjects to a real-time video of another subject sitting in the lab while they take their

[1] We use the terminology used in Abeler et al. (2019); Gneezy et al. (2018) use the term "direct costs" instead of "intrinsic costs", while Khalmetski and Sliwka (2019) use the term "fixed costs". Also, for the reputation component, Gneezy et al. (2018) use the concept of social identity. While social identity and reputation (i.e., social image) might not indicate the same constructs in general, in this context they are both used to refer to the willingness to appear honest to external observers. One exception to this modeling approach is Dufwenberg and Dufwenberg (2018), who assume that people suffer only a reputation cost, but in contrast to the previous models this cost does not depend on the probability of being seen as a liar by an external observer, but on the inference that the observer makes on the extent of the lie.

decisions. This other subject also sees the decision-maker's face and his or her computer screen in real time, but does not observe his or her die-roll outcome. We compare these two treatments to a *Control* treatment where subjects see a neutral pre-recorded video of another person.

We find that the increase of self-awareness has no significant effect on the average reported die-roll outcome. This suggests that self-image concerns may be less important than previously hypothesized to explain lying behavior, and that intrinsic lying costs might need to incorporate other psychological mechanisms. On the other hand, we show that the increase of observability decreases the average reported outcome even when information about the die-roll outcome is held private. To complement this finding, we conduct a survey experiment where we show that the likelihood of being perceived as a liar increases monotonically with the reported outcome. Moreover, we find that survey respondents associate high reports with the likelihood of having undesirable traits in several other dimensions. This further suggests that our effect in the *Social image* treatment indeed stems from the concern that decision-makers have about adverse inferences observers could make from the observation of high reports.

Our paper contributes to the literature on lying costs in several ways. In particular, numerous studies have suggested self-image or closely related concepts as drivers of intrinsic lying costs. Most closely to our paper, Mazar et al. (2008) investigate a tightly connected notion, one's self-concept. They show that subjects behave more honestly if primed with religious reminders and honor codes that increase attention to moral standards. While using such priming techniques might have its benefits, it has two main drawbacks. First, it primes everybody towards honest behavior by reminding people of specific moral standards. This, however, does not necessarily imply that these moral standards are the ones congruent with the inner standards of individuals. Second, reminding about moral standards might conflate individual with collective standards of behavior. For these reasons, our design abstracts from reminding subjects of a specific set of morals, but purely emphasizes the salience of inner standards, whatever they might be. Other studies have also hypothesized self-image concerns as determinants of intrinsic lying costs, and have gathered indirect evidence. For example, in their seminal study on the die-rolling paradigm, Fischbacher and Föllmi-Heusi (2013) find that reporting the second-highest-paying outcome is perceived as much less dishonest than reporting the highest-paying outcome. The authors suggest

that for the subjects who overreport partially, maintaining a favorable self-image might be one of the driving behavioral motives. Another example is Shalvi et al. (2011a), who manipulate the number of instructed die rolls, while holding fixed the first roll as payoff relevant. They find that subjects lie more if more die rolls are instructed, and argue that high reports are easier to justify to oneself if observed in any die roll after the first.

In this paper, we design an *exogenous* manipulation where we increase the salience of self-image concerns. In particular, our manipulation builds on the recent work of Bénabou et al. (2020), which implies that self-image concerns arise from the awareness of discrepancy between internal standards of behavior and the self (e.g., in light of current behavior). To make self-image concerns salient, we manipulate one's self-awareness, increasing the awareness of the aforementioned discrepancy and, ceteris paribus, self-image costs.[2] This reasoning dates back to the objective self-awareness theory (Duval and Wicklund, 1972), which posits that high levels of self-awareness induce behavior driven by salient moral standards. To test this theory, the most common manipulation has been to place a mirror in front of the subjects during the decision phase. It was used to show that increased self-awareness, for example, decreases simple transgressions (Beaman et al., 1979), increases the attribution of causality for a specific consequence to oneself (Duval and Wicklund, 1973), and can induce the use of corporal punishment depending on the subjects' inner attitudes towards it (Carver, 1975).

Our *Social image* treatment connects to a large body of literature which posits and shows that when being observed, people favor societal standards of behavior. More specifically, it investigates the effect of increased observability on honesty. Previous studies (see Gneezy et al., 2018; Abeler et al., 2019), in order to test predictions from their theoretical models, design experiments similar to the standard die-rolling paradigm, except that subjects see the outcome of a randomizing device on their screens. This allows the experimenter to map the observed outcome to the report for each individual and observe lying behavior at the individual level. Hence, these experiments increase observability by removing private information vis-à-vis the

[2]Theoretically this can be related to intrinsic lying costs under perfect information, assuming that the motive for this cost is to see oneself as honest (see, e.g., Abeler et al., 2019), or alternatively to imperfect information settings where the agent is assumed to forget his or her "type" and makes inferences about it given the actions taken. This notion is suggested by Dufwenberg and Dufwenberg (2018) as one possible interpretation of their model of lying costs and is in line also with other more general models (e.g., Bénabou and Tirole, 2006, 2011; Bodner and Prelec, 2003).

experimenter.[3] While this is reasonable in order to test the aforementioned theories, it comes at a high cost. Specifically, it removes one of the defining characteristics of lying situations, that is, the possibility of deception. In a usual lying situation, a sender of a message is trying to deceive the recipient regarding the state of the world privately known only to the sender. Since in the aforementioned studies the experimenter can observe both the state of the world (the outcome of the randomizing device) and the report, uncertainty about the state is absent; thus, there is no possibility of deception. In contrast to these studies, our aim is to investigate the effect of social image concerns in standard lying situations; hence, our observer is informed about the die-roller's report and can link the report to his or her identity, but cannot observe the state of the world.

From the perspective of recent theoretical models, we interpret our manipulation of observability as an exogenous increase in the individual parameter governing the reputational payoff, that is, how much subjects care about reputation. This interpretation relies on the following observations: i) being exposed to another observer beyond the experimenter may increase the social pressure on the decision-maker, ii) the fact that observers can tie the identity of decision-makers with their reports may make reputational concerns stronger, and iii) other student participants may constitute a more relevant audience than the experimenter for the decision-makers, as this is the audience they would usually be exposed to in everyday life. An increase in the reputation parameter predicts a decrease in lying in models that do not allow downward lying (see the Proposition in Dufwenberg and Dufwenberg (2018) and Proposition 5 in Khalmetski and Sliwka (2019)). Our results confirm this prediction and go in the same direction as the full observability treatments in Abeler et al. (2019) and Gneezy et al. (2018), but we observe a smaller effect than the ones they report (see Section 4), suggesting that private information has indeed a crucial impact on lying behavior.[4]

[3] Some studies attempt to *reduce* observability vis-à-vis the experimenter by performing double-blind procedures and manipulating probabilities of getting caught; however, they do not report a significant impact on behavior (Mazar et al., 2008; Fischbacher and Föllmi-Heusi, 2013).

[4] In the model by Abeler et al. (2019), which allows for downward lying, an increase in the reputation parameter may have two counterbalancing effects. On the one hand, it makes high reports more costly, and hence induces people to report lower numbers; on the other hand, there will be fewer expected liars at the high reports and (potentially) more expected liars at the low reports, which would make high reports become more attractive. In relation to this framework, our data would indicate that the first effect dominates the second because people report on average a smaller outcome when they are observed.

The remainder of the paper is structured as follows. In Section 2, we describe the experimental design and procedures. In Section 3, we report the results of our study. In Section 4, we discuss the results and conclude.

2 Experimental Design and Procedure

Our experimental setup is closely based on the die-rolling paradigm introduced by Fischbacher and Föllmi-Heusi (2013). Subjects were asked to roll a six-sided die and report the outcome of the die roll on their computer. Depending on their report, they were able to earn any amount from 0 to 5 euros. The payoff is equal to the reported outcome minus 1 euro, i.e., for a report of 1 subjects earned 0 euros, for a report of 2 subjects earned 1 euro, etc. As in Fischbacher and Föllmi-Heusi (2013), subjects were told to roll the die minimally twice, which can facilitate lying; however, they were explicitly told to report the outcome of the first die roll (for experimental instructions, see Appendix A.1). Subjects were asked to roll the die in a non-transparent plastic cup. The cup ensured that only they could observe the outcome which was visible only from directly above the cup.[5]

We designed three different treatments: *Self-image*, *Social image*, and *Control* treatment. In the *Self-image* treatment, we exogenously manipulated self-awareness, a mediator of a person's focus on his or her private self. In order to increase self-awareness, we exposed the subjects to their own image. In particular, from the moment subjects sat in the cubicles, a camera installed on the top of the monitor was capturing the image of their face, and playing it in real time on their computer screen (see Figures A2 and A3 in Appendix). The camera was positioned in such a way that the subjects could not evade its visual field, but also, that it was obvious that the plastic cup on the table was outside of this visual field. Additionally, we used a software which automatically detected and zoomed on subjects' faces. The video was placed in the upper part of the screen, while the instructions and the decision screen were placed below it. The subjects were fully informed that the video was not being recorded, and that only they were able to see it. In order to give some meaning to the camera, subjects were also informed that they would

[5] The setup is similar to the die-under-cup paradigm by Shalvi et al. (2011b).

answer a few short questions on the camera technology and settings at the end of the experiment.

In the *Social image* treatment, we exogenously manipulated observability by exposing subjects to the observation of other participants, i.e., observers. Upon their arrival at the laboratory, subjects were randomly assigned one of two roles, decision-makers or observers. Decision-makers were facing identical procedures and decisions as in *Self-image*, but instead of viewing their face in the video, they saw the face of their observer in real time. Observers were also seated in private cubicles at the same time as the decision-makers, but had no decisions to make. Each observer was paired with one decision-maker. Each observer saw i) the video footage of their paired subject's face in real time and ii) the decision screen of their paired subject in real time. Hence, both the observer and the paired subject saw each other, and additionally, the observer saw the decision-maker's screen. This was common knowledge. The procedure made observers fully aware of the reported outcome, but not of the actual die-roll outcome. At the end of the decision-making part, observers left the laboratory before decision-makers. This was publicly announced at the beginning of the study. Alongside the questionnaires from *Self-image*, decision-makers were also asked if they had ever seen their observer before, and if so, what was their relationship with them. Only one subject indicated knowing the paired observer. Removing this subject from the data does not change our results.

We compare the decisions in *Self-image* and *Social image* to a *Control* treatment. To design a comparable *Control* treatment, we address two concerns. First, subjects in both of our treatments were exposed to a video. If such a distraction drains cognitive resources, subjects could be more inclined to act affectively, following their automatic response and potentially biasing the results.[6] Second, in both treatments subjects saw a person looking at them: the observer in *Social image* and themselves in *Self-image*. Several studies have shown that being exposed to simple social cues such as a pair of observing eyes can influence one's behavior (Haley and Fessler, 2005; Bateson et al., 2006; Rigdon et al., 2009).[7] To address these two issues, our *Control* treatment is

[6] Dishonest behavior has been considered a cognitively demanding process linked to brain areas responsible for cognitive control (Sip et al., 2008; Greene and Paxton, 2009). On the one hand, studies show that when cognitive control is low due to cognitive depletion, people's automatic response is to act more selfishly (Achtziger et al., 2015), and more dishonestly (Gino et al., 2011), which could bias our results upwards. On the other hand, other studies suggest that people's automatic response is to behave more prosocially (Rand et al., 2012; Schulz et al., 2014). If subjects care about how much money the experimenter is left with, or if prosociality as a positive trait is connected to honesty, this could bias our results downwards.

[7] Note, however, that several studies fail to find an effect of social cues in different settings (Fehr and Schneider,

identical to the *Self-image* and *Social image* treatments, except that instead of seeing their own face or the observer's face subjects saw a mute video footage of a famous German news presenter (see Figure A4 in Appendix). As the context of the video was immediately recognizable, the subjects were perfectly aware that the video was prerecorded. Moreover, the news presenter is a non-controversial public person working for a mainstream public service, and as such does not trigger any tendentious associations.[8]

The lying task lasted less than 10 minutes, and it was run right after another task in line with Fischbacher and Föllmi-Heusi (2013). In the preceding task, subjects were exposed to the identical treatment manipulations as in the subsequent die-rolling experiment.[9] This means that the same manipulations were present from the moment subjects entered the cubicles until reaching the short questionnaire at the end of the session. Hence, the cameras were not abruptly turned on when reaching the lying task, and their function was clear from the beginning of the session. A total of 685 subjects participated in the study (59.7% female), out of which 531 subjects participated as decision-makers and 154 as observers. In *Social image*, observers earned 8 euros each for their participation in the entire session. Decision-makers were not informed about the payment of observers to avoid influences on their behavior driven by social comparison such as, e.g., inequity considerations (Fehr and Schmidt, 1999). Subjects were primarily students of the University of Bonn and were recruited with hroot (Bock et al., 2014). Sessions were conducted at the BonnEconLab in February 2016 and July 2017. Each session was dedicated to one treatment, and treatments were balanced within and across days of conduction. The experiment was programmed using Z-Tree (Fischbacher, 2007), and it followed ethical guidelines for study procedures from the BonnEconLab, University of Bonn.

In addition to our main experiment, we report results from a short survey which was run on a separate set of 100 subjects. The survey aims to complement our *Social image* treatment findings by exploring whether people i) perceive that higher reports are more likely to be a lie

2010; Lamba and Mace, 2010), and question the validity of previous evidence (Carbon and Hesslinger, 2011).

[8]Comparing *Self* or *Social image* treatment with the *Control* treatment ensures that any difference can be interpreted as the effect of increased self-awareness or observability, respectively, as any potential effects of social cues or cognitive depletion would be present in all three treatments. Notice, furthermore, that Falk (2018) uses an identical control treatment in an investigation of self-image concerns for moral behavior and compares it with a control treatment without any video, finding no difference between the two.

[9]In the preceding task, subjects played a dictator game (Bašić et al., 2020).

and ii) attribute negative traits to subjects who report high numbers. These questions are highly relevant in understanding potential reputation effects which can come from reports. Specifically, given that we maintain the outcome of the die roll private in our paradigm, we investigate whether reports alone can send signals about decision-makers which are perceived as informative by others. Upon accepting to participate in the survey, subjects were shortly informed about the die-rolling paradigm. Then, they were asked to evaluate the probability of a subject being a liar conditional on each of the six reports. Moreover, for hypothetical reports they were confronted with 6 different statements. For each of the statements they had to indicate how much they agreed on a scale from 0 to 7. The statements were: "I find this person trustworthy", "I would accept this person as a flatmate in my shared flat", "I would lend money to this person", "I would employ this person", "I would buy a car from this person", and "I would vote for this person". The hypothetical reports differed within-subject across the six statements, and the order of the hypothetical reports was randomized between-subject. The survey was conducted with students in front of the University of Bonn library and canteen. Each participant earned 5 euros for participating.

3 Results

We divide our results section in two subsections. In the first, we report the comparison between the *Control* and the *Self-image* treatments. In the second, we contrast the results of the *Control* and the *Social image* treatments, and complement this comparison with the results of our survey experiment.

3.1 The Influence of Self-image

In this subsection, we focus on the difference between *Control* ($n = 188$) and *Self-image* ($n = 189$). First, we check if people overreport the outcome of die-rolling. To do that, we contrast the reports of the die rolls with the uniform distribution by applying the Kolmogorov-Smirnov test for discrete data (henceforth KS d; Jann et al., 2008). In Figure 1, we report the frequency of each reported outcome (left panel) and the average reported outcome (right panel) in the two

treatments. We find that the distribution of reports is significantly different from the uniform distribution in both treatments (two-sided KS d test, $p < 0.001$ for both treatments). Next, we compare average reports between the two treatments. Figure 1 (right panel) shows that people report on average 4.62 in *Self-image* and 4.70 in *Control* with no significant differences across the two (two-sided t-test, $p = 0.630$). The result remains insignificant if we perform the analysis using an OLS regression and including control variables (see Table 1 in the following subsection). Next, we compare the distributions between the two treatments (Figure 1, left panel). We do not find any significant difference in this case either (Fisher's exact test, $p = 0.310$; Epps-Singleton two-sample test, $p = 0.270$).

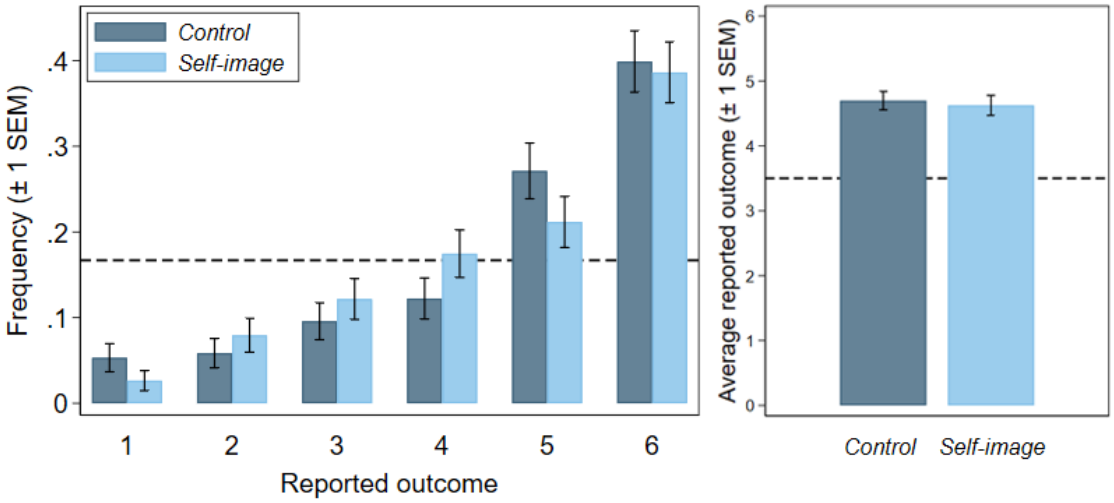

Figure 1: Frequencies of each reported outcome (left panel) and the average reported outcome (right panel) in *Self-image* and *Control*. Error bars indicate standard error of the means. The dashed line represents the expected frequency of each outcome (left panel) and the expected average outcome (right panel).

Finally, we turn to the analysis of report frequencies for each possible outcome. We observe that in both *Control* and *Self-image* people overreport the outcome 5 (two-sided binomial test against the expected true value of 0.167, $p < 0.001$ for *Control*, $p = 0.097$ for *Self-image*) and the outcome 6 (two-sided binomial test against the expected true value of 0.167, $p < 0.001$ for both treatments). If we compare the two distributions, we observe that the frequencies of all outcomes

are very similar across the two treatments. Only the frequencies for outcomes 4 and 5 exhibit a noticeable difference across treatments; however, this difference is not statistically significant (two-sided binomial test of proportions, $p = 0.154$ for outcome 4, $p = 0.183$ for outcome 5).

Overall, higher self-awareness and emphasized self-image concerns have no significant effect on the reported outcomes.

3.2 The Influence of Social Image

In this subsection, we analyze the difference between *Control* ($n = 188$) and *Social image* ($n = 154$). In Figure 2 we report the frequency of each reported outcome (left panel) and the average reported outcome (right panel) in the two treatments. First of all and similar to *Control* and *Self-image*, subjects in *Social image* significantly overreport their die-roll outcome compared to the uniform distribution (two-sided KS d test, $p < 0.001$). However, as shown in Figure 2 (right panel), people on average report less in *Social image* (4.34) than they do in *Control* (4.70). This difference is statistically significant (two-sided t-test, $p = 0.038$). When comparing observed distributions (Figure 2, left panel), we do not find a significant difference (Fisher's exact test, $p = 0.351$; Epps-Singleton two-sample test, $p = 0.344$).

Next, we focus on the reports for each possible outcome. Similar to *Control*, we observe that subjects in *Social image* overreport the outcome 5 and the outcome 6 (two-sided binomial test against the expected true value of 0.167, $p = 0.025$ for outcome 5, $p < 0.001$ for outcome 6). When comparing all 6 possible outcomes across the two treatments, we observe that the percentage of all reported outcomes is closer to the expected true value in *Social image* than in *Control*, which can explain the observed shift of the average reported outcome. However, these changes in report frequencies are rather small, and most outcomes in *Social image* are not significantly different from *Control* (two-sided binomial test of two proportions, $p > 0.196$ for outcomes 1, 3, 4, 5 and 6, $p = 0.082$ for the outcome 2).

Next, in Table 1, we report an OLS regression analysis to confirm the robustness of our findings in *Self-image* and *Social image* to further controls. We report 3 regression models where we use the reported outcome as dependent variable and two treatment dummies (with the *Control* treatment as omitted category). Model (1) reports these estimates without controls and confirms

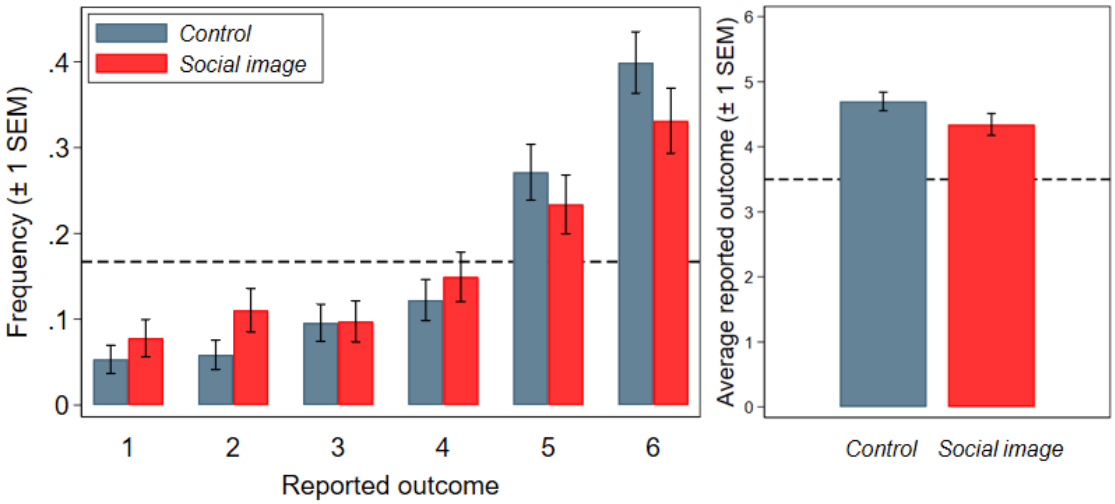

Figure 2: Frequencies of each reported outcome (left panel) and the average reported outcome (right panel) in *Social image* and *Control*. Error bars indicate standard error of the means. The dashed line represents the expected frequency of each outcome (left panel) and the expected average outcome (right panel).

the results from non-parametric tests. Model (2) indicates that these findings are robust when controlling for gender. Model (3) shows the same results, also controlling for age and personality characteristics (*Social image* remains significant at a 10% level; $p = 0.059$). In addition, we also show that females report less than males in the overall sample, replicating a standard finding in die-roll experiments (see Abeler et al., 2019). Finally, we observe that reports increase with extraversion and decrease with agreeableness.

Next, we report the results of the survey conducted on uninvolved subjects ($n = 100$) to complement our social image results. In Figure 3, we report the average expected probability of the decision-maker being dishonest conditional on the report (left panel) and the average agreement with the statements conditional on the report (right panel). The average belief about the probability that subject is a liar increases with the report (OLS regression with standard errors clustered at the individual level, $p < 0.001$). In particular, we observe a monotonic increase of beliefs ranging from 6% to 55%. With respect to the agreement with the statements, we observe that on average, survey participants perceive a person reporting higher numbers as

Table 1: OLS regressions estimates of treatment effects

Variables	Dependent variable: reported outcome		
	(1)	(2)	(3)
Self-image	-0.072	-0.069	-0.061
	(0.150)	(0.148)	(0.148)
Social image	-0.353**	-0.346**	-0.317*
	(0.171)	(0.169)	(0.167)
Female (=1)		-0.516***	-0.375**
		(0.130)	(0.150)
Age			0.001
			(0.014)
Big5: Extraversion			0.220***
			(0.075)
Big5: Agreeableness			-0.258***
			(0.072)
Big5: Neuroticism			0.038
			(0.081)
Big5: Conscientiousness			-0.048
			(0.071)
Big5: Openness			0.026
			(0.069)
Constant	4.697***	5.001***	4.889***
	(0.108)	(0.128)	(0.359)
Observations	531	531	529
R-squared	0.009	0.037	0.068

The table presents OLS regressions using reported outcome as the dependent variable. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

less trustworthy, and they declare to be less willing to consider such a person as a flatmate, to lend them money, employ them, buy a car from them, or vote for them (Spearman's rho, $p = 0.004$ for "Flatmate" comparison, $p < 0.001$ for all other comparisons).

To sum up, we find that increased observability decreases the average reported outcome. Furthermore, we observe that reporting higher numbers signals a higher likelihood of being a liar, and a higher likelihood of being untrustworthy and having undesirable traits in many other domains. This further supports the reputation channel as the explanation behind the results of our social image manipulation.

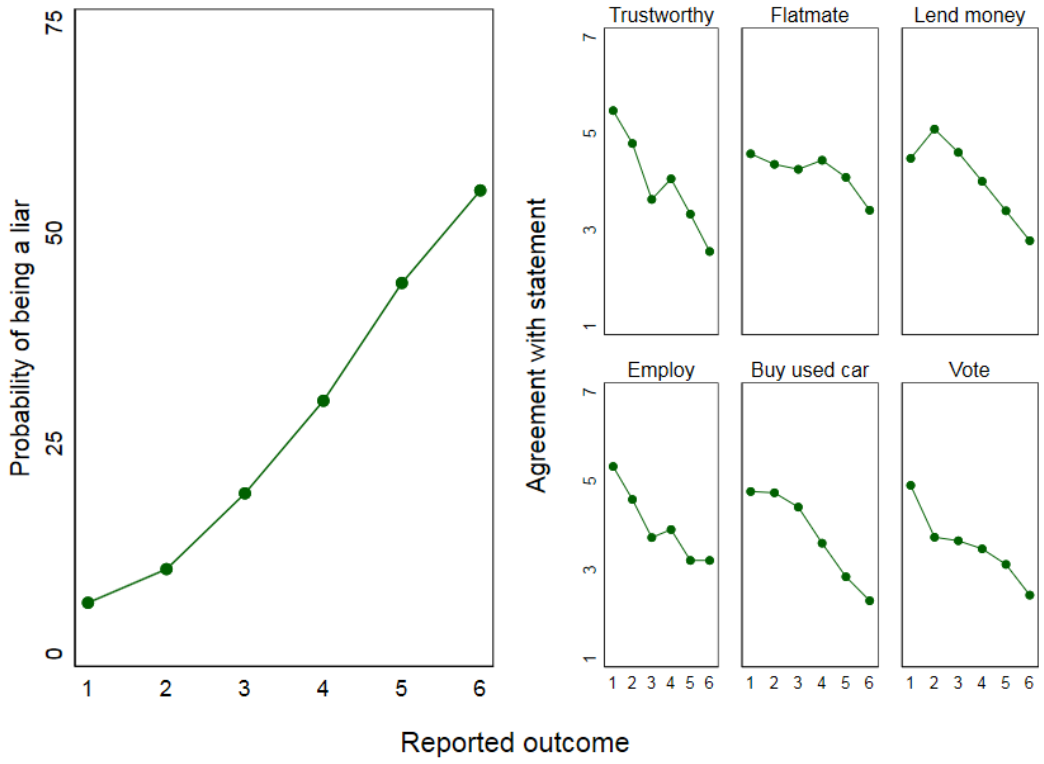

Figure 3: Survey results

The average expected probability that the subject is lying conditioning on the reported outcome (left panel), and the average agreement with the statement conditioning on the reported outcome (right panel). The statements are "I find this person trustworthy", "I would accept this person as a flatmate in my shared flat", "I would lend money to this person", "I would employ this person", "I would buy a car from this person", and "I would vote for this person".

4 Discussion and Conclusion

In this paper, we have investigated the influence of self and social image concerns on lying behavior. We have exogenously manipulated self-awareness and observability by exposing subjects to their own image or to the observation of another participant in real time, respectively. We have shown that the increase of self-awareness has no effect on the average report.

There are several reasons why this may be the case. One possibility is that the importance of self-image concerns might be low in certain domains. Falk (2018) uses a similar self-image

manipulation as ours, and shows that subjects care about their self-image when confronted with the choice of administering an electric shock to another individual for money. Moreover, using an identical manipulation, Bašić et al. (2020) find that increasing self-awareness significantly increases generosity. This suggests that honesty might not be a (salient) inner standard, and the act of lying might have a negligible impact on one's self image. Alternatively, self-image concerns could be an important determinant of lying and cheating behaviors; but in the die-rolling task, where the only "victim" of immoral behavior is the experimenter, and the negative externality is arguably weaker than when administering an electric shock or being selfish to another participant, the strength of self-image concerns could be insufficient to generate a significant shift in behavior. Hence, it would be important to study self-image concerns in settings where the externality of lies is more pronounced, as for example, in sender-receiver games (Crawford and Sobel, 1982; Gneezy, 2005).[10]

The lack of self-image effect might suggest that there could be other psychological mechanisms at the origin of intrinsic lying costs in die-rolling experiments.[11] While several motives have already been proposed, e.g., social norms and guilt aversion, it was recently shown that these motives cannot reconcile all the findings from previous die-rolling experiments (Abeler et al., 2019). One potential explanation is that honesty could have components of heuristical behavior, that is, subjects display automatic honest behavior that has been transmitted by parents and/or other role models (see Bénabou et al., 2020), and do not question this behavior even if characteristics of the decision environment change. Alternatively, in line with Dufwenberg and Dufwenberg (2018), subjects could have no intrinsic lying costs, but only reputational concerns vis-à-vis an external audience. In order to reconcile previous findings, and in line with Dufwenberg and Dufwenberg (2018), these reputational concern should be related not only to being a liar or not, but also to the size of lies.

With regard to social image concerns, we have shown that the increase in subjects' observ-

[10] Additionally, it is possible that for a certain proportion of subjects, their inner standard is self-interest and not honesty, which could in turn explain why the average report does not change in our self-image manipulation. This explanation, however, is not consistent with standard conceptualization of self-image concerns (e.g., Bénabou and Tirole, 2006). Moreover, with such polarization of inner standards, we would also expect a polarization of reports in the *Self-image* treatment, which is not supported by our data.

[11] Note that the existence of intrinsic lying costs was identified in the die-rolling paradigm (Abeler et al., 2019; Gneezy et al., 2018; Khalmetski and Sliwka, 2019).

ability significantly decreases the average report, and that reporting high paying numbers ties subjects with a stigma of likely being a liar and having undesirable traits. The results clearly indicate that reputation constitutes an important component of lying costs. In contrast to previous literature, however, we have reported a manipulation that keeps information about the true outcome of the die roll private, and hence maintains the private-information property of standard lying situations. The direction of our effect confirms theories that predict a decrease in lying due to increased reputational concerns (Dufwenberg and Dufwenberg, 2018; Khalmetski and Sliwka, 2019). As we have noted, we find a smaller effect compared to the effects found when the outcome of a randomizing device is public. In particular, we observe a Cohen's d effect size of 0.226, while Abeler et al. (2019) observe an effect size of 0.761, and Gneezy et al. (2018) of 0.284. This suggests that private information indeed has an important impact on lying behavior.

References

Abeler, J., Becker, A., and Falk, A. (2014). Representative evidence on lying costs. *Journal of Public Economics*, 113:96–104.

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

Achtziger, A., Alós-Ferrer, C., and Wagner, A. K. (2015). Money, depletion, and prosociality in the dictator game. *Journal of Neuroscience, Psychology, and Economics*, 8(1):1–14.

Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, 2(3):412–414.

Bašić, Z., Falk, A., and Quercia, S. (2020). Self-image, social image and prosocial behavior. *Mimeo*.

Beaman, A. L., Klentz, B., Diener, E., and Svanum, S. (1979). Self-awareness and transgression in children: Two field studies. *Journal of Personality and Social Psychology*, 37(10):1835–1846.

Bénabou, R., Falk, A., and Tirole, J. (2020). Narratives, imperatives, and moral persuasion. Working paper available at https://www.briq-institute.org/wc/files/people/armin-falk/working-papers/narratives-imperatives-and-moral-reasoning.pdf.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *The American Economic Review*, 96(5):1652–1678.

Bénabou, R. and Tirole, J. (2011). Laws and norms. *NBER Working Paper*, (No. 17579).

Bock, O., Baetge, I., and Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120.

Bodner, R. and Prelec, D. (2003). Self-signaling and diagnostic utility in everyday decision making. *The Psychology of Economic Decisions*, 1:105–26.

Carbon, C.-C. and Hesslinger, V. M. (2011). Bateson et al.'s (2006) cues-of-being-watched paradigm revisited. *Swiss Journal of Psychology*, 70:203–210.

Carver, C. S. (1975). Physical aggression as a function of objective self-awareness and attitudes toward punishment. *Journal of Experimental Social Psychology*, 11(6):510–519.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Dufwenberg, M. and Dufwenberg, M. A. (2018). Lies in disguise – A theoretical analysis of cheating. *Journal of Economic Theory*, 175:248–264.

Duval, R. A. and Wicklund, S. (1972). *A theory of objective self awareness*. Oxford: Academic Press.

Duval, S. and Wicklund, R. A. (1973). Effects of objective self-awareness on attribution of causality. *Journal of Experimental Social Psychology*, 9(1):17–31.

Falk, A. (2018). Facing yourself – A note on self-image. Working paper available at https://www.briq-institute.org/wc/files/people/armin-falk/working-papers/facing-yourself-a-note-on-self-image.pdf.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 144(3):817–868.

Fehr, E. and 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. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

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

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

Gino, F., Schweitzer, M. E., Mead, N. L., and Ariely, D. (2011). Unable to resist temptation: How self-control depletion promotes unethical behavior. *Organizational Behavior and Human Decision Processes*, 115(2):191–203.

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

Gneezy, U., Kajackaite, A., and Sobel, J. (2018). Lying aversion and the size of the lie. *American Economic Review*, 108(2):419–53.

Greene, J. D. and Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences*, 106(30):12506–12511.

Haley, K. J. and Fessler, D. M. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human behavior*, 26(3):245–256.

Jann, B. et al. (2008). Multinomial goodness-of-fit: Large-sample tests with survey design correction and exact tests for small samples. *Stata Journal*, 8(2):147.

Khalmetski, K. and Sliwka, D. (2019). Disguising lies – Image concerns and partial lying in cheating games. *American Economic Journal: Microeconomics*, 11(4):79–110.

Lamba, S. and Mace, R. (2010). People recognise when they are really anonymous in an economic game. *Evolution and Human Behavior*, 31(4):271–278.

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

Rand, D. G., Greene, J. D., and Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489(7416):427–430.

Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology*, 30(3):358–367.

Schulz, J. F., Fischbacher, U., Thöni, C., and Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology*, 41:77–87.

Shalvi, S., Dana, J., Handgraaf, M. J., and 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. J., and De Dreu, C. K. (2011b). Ethical manoeuvring: Why people avoid both major and minor lies. *British Journal of Management*, 22(s1):S16–S27.

Sip, K. E., Roepstorff, A., McGregor, W., and Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, 12(2):48–53.

A Appendix

A.1 Experimental instructions

The following section contains experimental instructions translated from German.

A.1.1 Die-roller: *Control*, *Self-image*, and *Social image* treatments

The general instructions and the instructions for the first task were printed and left in front of the subjects' computer screens before they entered the lab. At the beginning of the experiment, the instructions were read out loud by one of the experimenters.

Welcome to this study!

You are participating in an economic study. Depending on your answers, you can earn a certain amount of money. The money will be paid out at the end of the study in cash. It is therefore very important that you read the instructions carefully, and that you understand them.

Only for the Control treatment. As you can see, there is a video playing on your computer screen. This video will also be played during the study.

Only for the Self-image treatment. As you can see, there is a camera installed above the computer screen. The image that the camera is capturing is shown on your computer screen in real time. **Please note: No video streams are saved, and only you and no other person can see your camera video.** At the end of the study we will ask you several short questions about the camera technology and camera settings.

Only for the Social image treatment. As you can see, there is a camera installed above the computer screen. You can see another participant of the study. Simultaneously, the participant can also see you. **It is strictly forbidden to communicate in any way with this other participant, e.g, through waving, signs, facial expressions, or similar.** This other participant has the role of observer. Your observer has received his own instructions, in which his task is clearly explained. He has only one task, and that is to observe you and your decisions.

Your observer sees your computer screen in real time. That means that all the movements that you do with your mouse, and all the decisions that you take during this study, will be seen by your observer. Please note that there can be short delays in the transmission of the camera video. The transmission of your screen and mouse movements occurs with no delay. **Please note: No video streams are saved, only your observer and no other person can see your camera video.** If you disagree with this, you can finish your participation on the study now. At the end of the study, we will ask you several short questions about the camera technology and camera settings. After the end of the study, your observer will leave the laboratory before you.

All statements made in these instructions are true. This holds generally for all studies conducted at the Bonn Laboratory for Experimental Economic Research, and also for this study.

During the study, communication between participants is forbidden. If you have questions, then please direct them only to us. Please raise your hand and a member of the experimental team will come to answer privately. Violating this rule leads to exclusion from the study.

At this point, the experimenter read the instructions of the first task, which was followed by the task itself. After the task, we presented the subjects with the instructions of the die-rolling task on their computer screens.

Die roll, introductory screen. The first part of the study is now finished. The second part of the study is not connected to the first. For the following task you will require a cup and a die. Please wait until we bring it to you.

At this point, the subjects were each given a plastic non-transparent cup and a die.

Die roll, instructions screen. Please do not use the die nor the cup before you are asked to do so. When you are asked to roll the die, roll it twice. More specifically, take it with your hand and roll it in the cup. Your task is to report which number you have rolled with the first

try. Depending on the reported number, you will receive a certain amount of money. How much money you earn for a given number is presented in the table underneath *(see Figure A1)*.

Rolled number	1	2	3	4	5	6
Payoff	0 euros	1 euro	2 euros	3 euros	4 euros	5 euros

Figure A1: Payment table

The second roll is to assure yourself that the die is not loaded. You may also roll the die more than two times; however, only the first roll counts. The money that you earn in this task will be added to the money you have earned so far and paid out at the end of the study in cash.

One more time: When you are asked to roll the die, roll it twice. Report the number that you rolled with the first try. The amount of money you will earn depending on your report is presented in the table.

Please do not start yet. If you have understood everything, press "next". If you have questions, raise your hand.

Die roll, decision screen. Roll the die twice now. Report which number you have rolled with the first try.

A.1.2 Observer

The instructions were printed and left in front of the subjects' computer screens before they entered the lab.

READ IMMEDIATELY

Welcome to this study!

Important: It is strictly forbidden to communicate with the participant you see in the video in any way, e.g., through waving, signs, facial expressions, or similar.

In this study, you are participating in a role of observer. Your only have one task, and that is to observe another participant of the study. For this task, you will receive a payment of 8 euros in cash at the end of the study.

On your screen, you can see a video of another participant and his decision screen in real time. That means that you will observe the decisions that this participant takes during the study. At the same time, this participant can also see you through the camera that is installed on your computer screen. **Please note: No video streams are saved. Except for the participant you see, no other person can see the video from your camera.**

During the study, communication between participants is forbidden. If you have questions, direct them to us. Raise your hand, and a member of the experimental team will come to answer privately. If you violate this rule, you will be excluded from the study.

The instructions that will be read out loud at the beginning of the study are for the subjects who will take decisions. That means, they are intended for the participant who you are observing. Listen carefully, so that you can understand what is the task of this participant.

A.2 Additional figures

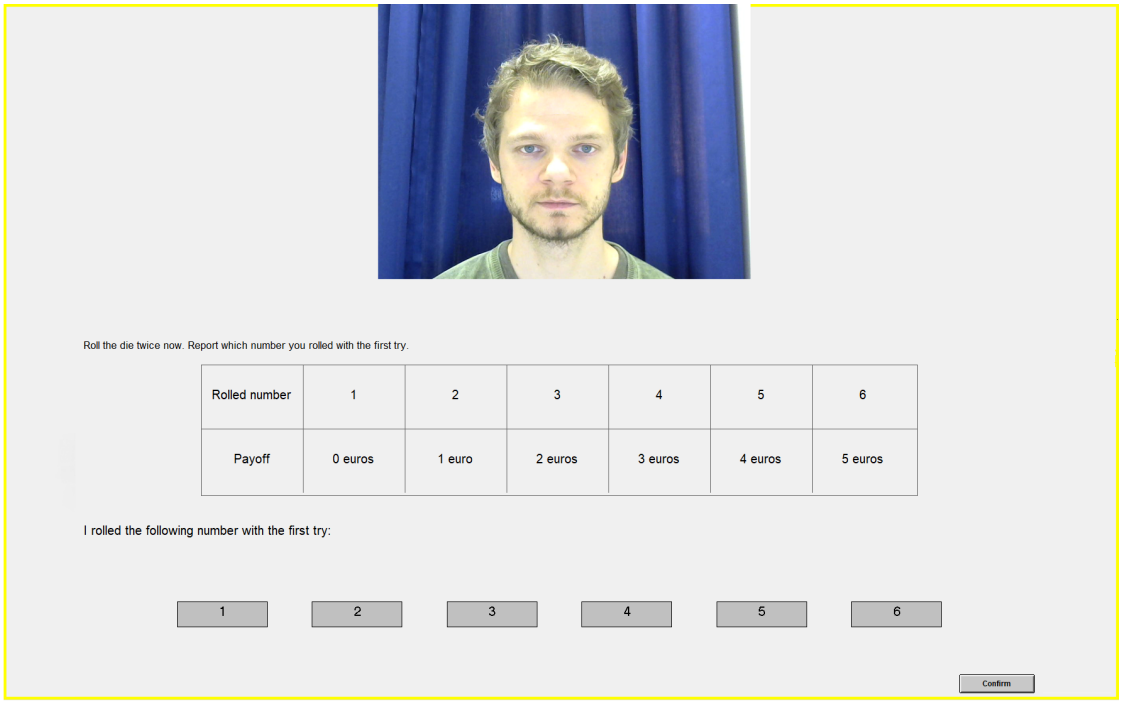

Figure A2: Decision screen in *Self-image* and *Social image* (translated from German)

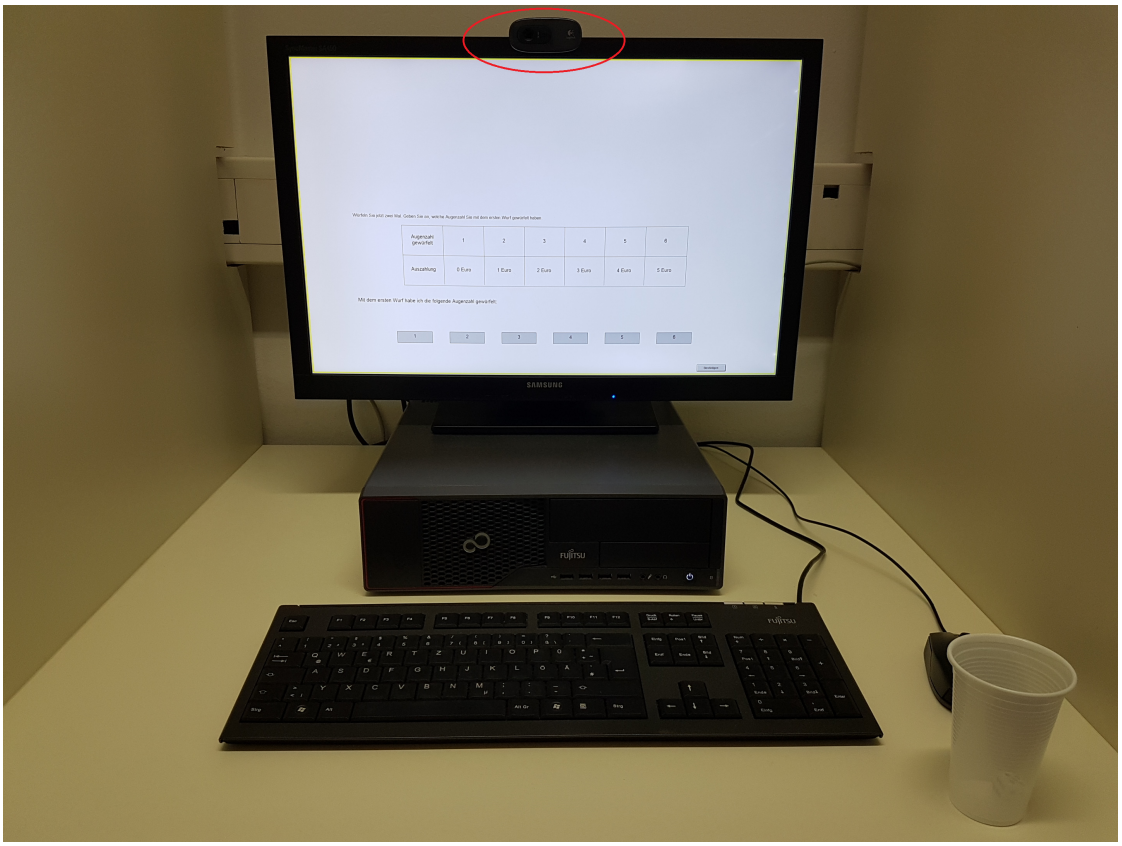

Figure A3: The cubicle in *Self-image* and *Social image* with a camera attached to the computer screen. The video is turned off for demonstration purposes.

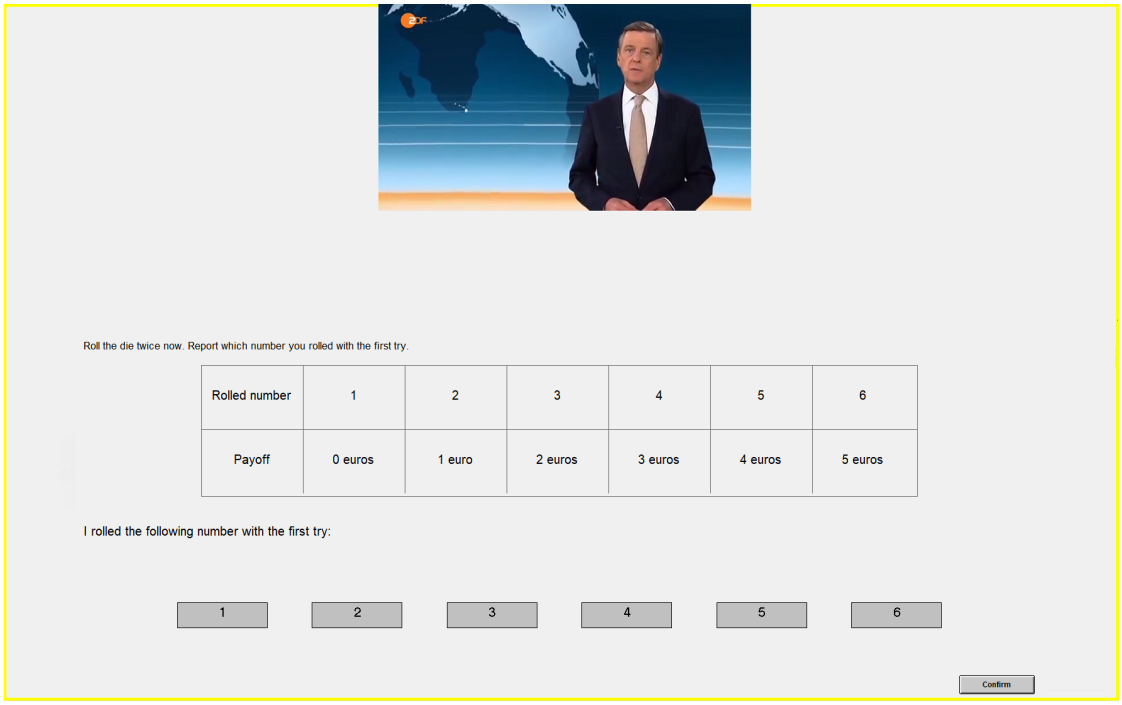

Figure A4: Decision screen in *Control* (translated from German)