

Petrishcheva, Vasilisa; Riener, Gerhard; Schildberg-Hörisch, Hannah

**Conference Paper**

## Loss aversion in social image concerns

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Petrishcheva, Vasilisa; Riener, Gerhard; Schildberg-Hörisch, Hannah (2020) : Loss aversion in social image concerns, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/224581>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Loss Aversion in Social Image Concerns\*

Vasilisa Petrishcheva,<sup>†</sup> Gerhard Riener<sup>‡</sup> and Hannah Schildberg-Hörisch<sup>§</sup>

## Abstract

We conduct a laboratory experiment to explore whether loss aversion applies to social image concerns. First, subjects are ranked publicly in a social image relevant domain, an IQ test, to establish own rank as a within-subject reference point. We then induce an exogenous change in within-subject rank and offer scope for lying about it to test whether subjects whose rank exogenously deteriorates (loss in social image) lie more than those who experience an equally-sized gain in rank (gain in social image). We find evidence for loss aversion in social image concerns, however, only for subjects who strongly care about their social image and have reputation to lose (high initial reference point).

**JEL Codes:** C91, D91

**Keywords:** loss aversion, social image concerns, lying behavior, laboratory experiment

**AEA RCT Registry:** <https://www.socialscienceregistry.org/trials/3422>

---

\*We are grateful to Loukas Balafoutas, Jonas Fooker, Lorenz Götte, Paul Heidhues, Florian Kerzenmacher, Daniel Müller, Hans-Theo Normann, Lionel Page, Wojtek Przepiorka, Sigrid Suetens, Agnieszka Tymula, Joël van der Weele, Jana Willrodt and Florian Zimmermann as well as participants of the Meeting on Experimental and Behavioral Social Sciences in Utrecht (IMEBESS 2019), 12th Maastricht Behavioral and Experimental Economics Symposium (M-BEES 2019), Workshop on Belief-Dependent Preferences in Copenhagen (PGT Workshop, 2019), Sydney Workshop on Experimental Economics and Theory (SWEET 2019) and GfW-Jahrestagung 2019 in Düsseldorf for helpful comments and suggestions. We also thank Robin Bitter and Vladyslav Zhurakhinsky for excellent research assistance. Financial support from DFG through the grant for GRK 1974 Competition Economics is gratefully acknowledged.

<sup>†</sup>University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany; E-mail: petrishcheva@dice.hhu.de.

<sup>‡</sup>University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany; E-mail: riener@dice.hhu.de.

<sup>§</sup>University of Düsseldorf, Düsseldorf Institute for Competition Economics (DICE), Universitätsstr. 1, 40225 Düsseldorf, Germany; E-mail: schildberg-hoerisch@dice.hhu.de; IZA, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany

# 1 Introduction

Humans care how they as persons are perceived by their fellow humans and go a great length to build up a positive image of themselves (e.g. Andreoni and Bernheim, 2009; Ariely et al., 2009; Bursztyn and Jensen, 2017; Ewers and Zimmermann, 2015; Soetevent, 2011). There is however always a risk that this carefully crafted image is at threat of being lost. As humans dislike losses more than they appreciate gains (Kahneman and Tversky (1979), Camerer (1998), Wakker (2010), Barberis (2013))—at least in the monetary domain—how do they react when they experience their social image declining?

Casual observations suggest that when social image is at stake people engage in lies and denial of their deeds to keep it. Does trying to shield oneself from a loss in social image generally lead to morally more deviant behavior than striving for a gain in social image? Or is it a particular behavior of those people who are more inclined to immoral behavior that can lead to tragic fall in the first place? Exogenous variation in the loss of social image is hard to imagine in the field and the extent of lying difficult to observe. Hence, we design a parsimonious laboratory experiment to test for the presence of loss aversion in social image concerns.

To fix ideas, we develop a simple model of loss aversion in social image concerns and derive testable hypothesis. In the experiment, we induce exogenous variation in reference points such that subjects either experience a potential loss or gain in their social image, while keeping average social image constant across treatments. We then offer subjects scope for improving their social image by lying about their true type. This allows us to test whether—on average—subjects lie more when they experience losses than when they experience gains in their social image.

We find evidence for loss aversion in social image concerns for subjects who strongly care about their social image—as measured by survey instruments—and those with high initial social image, i.e. those who have a potential to lose reputation.

Image concerns expand over various domains: People care about being perceived smart and skillful, prosocial, altruistic, and trustworthy, generous, or wealthy<sup>1</sup> and also forgo considerable benefits in order to avoid a loss of social image. In our experiment, we induce social image concerns by letting subjects perform an IQ test and report its results publicly. However, signaling skillfulness can be a two-sided sword as Austen-Smith and Fryer Jr (2005) show in a two-audience signaling model. For example, high ability students may under-invest in education because such investments lead to rejection by their peer group Bursztyn, Egorov, et al. (2019).<sup>2</sup> So it is important to establish that a test of intelligence is indeed suitable to induce social image that is worth striving for our university student sample. Ewers and Zimmermann (2015) document that in a student sample similar to the one used in this study, subjects misreport their private information on ability in a lab context in order to appear more skillful even when strong monetary incentives are given to tell the truth.

This study also contributes to the growing literature on lying costs extensively summarized in Abeler et al. (2019).<sup>3</sup> Abeler et al. analyze in detail the preferences for truth-telling and identify two main channels, namely, lying costs and image concerns for being perceived as an honest person. Our theoretical framework and experiment design build on their work. For example, our experiment design ensures that lying cannot be detected at the individual level such that image concerns for being seen as an honest person by others cannot play a role in the context of our experiment. Moreover, our design uses the extent of lying, i.e. the lying costs subjects are willing to incur, to quantify how much they suffer from losing or gain from improving their social image. Therefore, our finding that subjects with strong image concerns who have a reputation to lose are more likely to report dishonestly than others also speak to situations in which honest reporting of private information is key but not incentive-compatible. Our findings suggest that monitoring efforts should be targeted at individuals who have a high reputation and care strongly

---

<sup>1</sup>Bursztyn and Jensen (2017) present a detailed overview of recent literature on social image concerns covering a wide range of domains.

<sup>2</sup>Bursztyn, Egorov, et al. show that students are less likely to sign up for an SAT preparation course, as well as to take an SAT exam itself, if their choices are observable. They therefore forgo educational investment due to social stigma.

<sup>3</sup>Abeler et al. provide a web interface where they present recent experiments on lying in great detail.

about it. Moreover, one should try to make it harder to lie while keeping a good reputation, e.g., via transparency, naming-and-shaming or reputation systems (see also Abeler et al. (2019)).

Furthermore, we relate to the literature which links the concept of loss aversion (in the monetary domain) to lying behavior. Grolleau et al. (2016) and Schindler and Pfattheicher (2017) compare the extent of lying for individuals who face monetary losses and gains and find that participants misreport more to avoid a monetary loss than they do to increase their monetary gain.

As loss aversion is widely documented in terms of money (e.g. Booij and Van de Kuilen (2009) and Pennings and Smidts (2003)) as well as material goods (e.g. Kahneman, Knetsch, et al. (1990)) and health outcomes (Bleichrodt et al. (2001)) we expect that humans have the same inclination when it comes to social image utility. To the best of our knowledge, we are the first to consider whether misrepresentation of one's own ability is different in order to compensate for losses than to benefit from gains in the domain of social image. Importantly, our findings imply that loss aversion can also play a role in the non-material domain of social image.

Section 2 presents a theoretical framework combining social image concerns and loss aversion. Section 3 describes the experimental design and procedures, before we outline our hypotheses in Section 4. Results are presented in Section 5 and Section 6 concludes.

## 2 Theoretical Framework

Our model setup incorporates three key psychological features entering into individual utility. Agents prefer higher social image to lower social image. Agents experience loss aversion in the social image domain, i.e. they have a value function with standard properties with respect to their social image. Agents dislike lying, i.e. they experience costs of misreporting the true state of the world.

Consider a two-period game with  $t \in \{1, 2\}$  where in the first period an agent  $i$  establishes a reference point concerning her social type  $s_{i1}$ . In the second stage she learns about her true social type  $s_{i2}$ . Before her type is revealed to an observer, she may misrepresent her type through making a false claim. In each period  $t$  she derives  $u(s_{it})$  from her social image.  $u(\cdot)$  is assumed to be differentiable and weakly increasing in  $s_i$ .

The cost of misrepresentation of her type follows the model by Abeler et al. (2019) and Khalmetski and Sliwka (2017). We denote the true state of the world  $\omega \in [-\bar{\omega}, \bar{\omega}]$  which is independently identically distributed across individuals. The agent's report of the true state is  $r \in [-\bar{r}, \bar{r}]$  with  $\bar{\omega} = \bar{r} > 0$ . In period  $t = 2$ , her public final score consists of her actual performance  $s_{i2}$  plus her report of the true state  $r_i$ . Individual  $i$  dislikes misreporting the true state and experiences lying costs  $c(\omega_i, r_i)$ . Lying costs are zero if the state is reported truthfully, i.e.  $c(\omega_i, \omega_i) = 0$ , and positive otherwise. Lying costs depend on the size of misreporting and are symmetric around  $\omega_i$ , i.e.  $c(\omega_i, \omega_i + a) = c(\omega_i, \omega_i - a)$  for all  $a$ . In other words, an agent experiences the same lying costs when misreporting upwards and downwards to the same extent. Moreover, the agent has a lying cost sensitivity parameter  $\theta_i$ . Note that we do not account additionally for the social image concerns of being seen as a liar as we rule them out in our experimental design.

We define agent  $i$ 's total utility in stage  $t$  as an additive combination of her image concern and her lying cost:

$$\phi_{it} = u(s_{it} + r_i I_t) - \theta_i c(\omega_i, r_i) I_t, \quad (1)$$

where  $I_t$  is an index which equals to one in period  $t = 2$ , and equals zero otherwise. Additionally, we assume the following differentiable value function:

$$v(s_i) : v(\Delta s_i) < -v(-\Delta s_i)$$

where  $\Delta s_i = s_{i2} - s_{i1}$ . The value function satisfies the standard assumptions of prospect theory (Kahneman and Tversky, 1979). Negative deviations from the reference points have a larger absolute impact on

utility than similar sized positive deviations, i.e.  $v'(\Delta s_i) < v'(-\Delta s_i)$ . Additionally, the value function is concave for gains ( $v''(\Delta s_i) < 0$  for  $\Delta s_i > 0$ ) and convex for losses ( $v''(\Delta s_i) > 0$  for  $\Delta s_i < 0$ ).

Then agent  $i$ 's utility in part 1 is

$$\phi_{i1} = u(s_{i1})$$

and in part 2 it is

$$\phi_{i2} = u(s_{i2} + r_i) + v(s_{i2} + r_i - s_{i1}) - \theta_i c(\omega_i, r_i),$$

which she maximizes with respect to her report  $r_i$ . We treat  $s_{it}$  for  $t \in \{1, 2\}$  as exogenous. Another interpretation of  $s$  is that it is an endogenously chosen variable and has already been optimized with respect to effort provision, social preferences or any other type of costs.

The agent's first order condition for optimal reporting is then:

$$\frac{d\phi_{i2}}{dr_i} = \underbrace{\frac{du(s_{i2} + r_i)}{dr_i}}_{\geq 0} + \underbrace{\frac{dv(s_{i2} + r_i - s_{i1})}{dr_i}}_{> 0} - \theta_i \frac{dc(\omega_i, r_i)}{dr_i} \stackrel{!}{=} 0. \quad (2)$$

Note that  $\frac{du(s_{i2} + r_i)}{dr_i} \geq 0$  because if the score increases, the agent obtains a non-negative marginal utility;  $\frac{dv(s_{i2} + r_i - s_{i1})}{dr_i} > 0$  is independent of whether an individual is in the loss or gain domain (or shifts from the loss to the gain domain);  $\frac{dc(\omega_i, r_i)}{dr_i} \geq 0$  if  $\omega_i \leq r_i$  and  $\frac{dc(\omega_i, r_i)}{dr_i} < 0$  if  $\omega_i > r_i$ .

The first result is straight forward:

**Proposition 1.** *Individuals never underreport the true state.*

*Proof.* Given the true state  $\omega_i$  an agent always strictly prefers to report  $r_i = \omega_i$  to any  $\tilde{r}_i$  such that  $\tilde{r}_i < \omega_i$  because underreporting leads to lower utility due to three factors. On the one hand, an individual obtains weakly lower utility derived from the social image:  $u(s_{i2} + \tilde{r}_i) \leq u(s_{i2} + r_i)$ . On the other hand, the level of value function is lower at  $\tilde{r}_i$  than at  $r_i$  for any value of  $\Delta s_i$ , i.e.  $v(\Delta s_i + \tilde{r}_i) < v(\Delta s_i + r_i)$ , because  $\frac{dv(s_{i2} + r_i - s_{i1})}{dr_i} > 0$ . Finally, reporting  $r_i = \omega_i$  yields zero lying costs while reporting  $\tilde{r}_i$  is misreporting the true state and is thus costly, i.e.  $c(\omega_i, \tilde{r}_i) > c(\omega_i, \omega_i)$ .

Therefore, given state of the world  $\omega_i$ , an agent never reports  $r_i < \omega_i$  which leads to Proposition 1. Additionally, if  $\omega_i = \bar{\omega}$  and an agent does not lie down, it directly implies that the agent will only report truthfully (i.e.  $r_i = \omega_i$ ) which leads to Lemma 1.  $\square$

**Lemma 1.** *Individuals always report truthfully if  $\omega_i = \bar{\omega}$ .*

Next, we study the conditions under which agents will engage in full and partial lying. We assume lying costs to be non-differentiable at zero. Equation 2 states that  $\frac{d\phi_{i2}}{dr_i} \Big|_{r_i = \omega_i} > 0$  if an agent reports truthfully which directly implies that there exists a small enough upward deviation  $r_i - \omega_i$  which maximizes this agent's utility. If we assume the lying cost function to be differentiable at zero, all individuals, even those with extremely high lying sensitivity, would engage in overreporting.<sup>4</sup> We thus believe it is more plausible to assume lying costs which are non-differentiable at zero with  $\frac{dc(\omega_i, r_i)}{dr_i} \Big|_{r_i = \omega_i}^+ > 0$  and  $\frac{dc(\omega_i, r_i)}{dr_i} \Big|_{r_i = \omega_i}^- < 0$ .

**Proposition 2.** *Individuals report truthfully if  $\theta_i \geq \theta_i^{true}$  or  $\omega_i = \bar{\omega}$ .*

*Proof.* If the lying cost function is non-differentiable at zero (for example, constant lying costs), individuals with high sensitivity to lying will report truthfully. This happens if:

$$\theta_i \geq \theta_i^{true} = \left( \frac{dc(\omega_i, r_i)}{dr_i} \Big|_{r_i = \omega_i}^+ \right)^{-1} \left( \frac{dv(s_{i2} + r_i - s_{i1})}{dr_i} \Big|_{r_i = \omega_i} + \frac{du(s_{i2} + r_i)}{dr_i} \Big|_{r_i = \omega_i} \right).$$

$\square$

<sup>4</sup>See detailed proof in Appendix.

A lower lying sensitivity increases misreporting but not necessarily fully, i.e. up to  $r_i = \bar{\omega}$ . There are several factors which influence the choice between partial and full lying:

- Lying sensitivity  $\theta_i$ : if the agent is very insensitive to lying, she might engage in lying fully. However, if her sensitivity parameter is relatively high (but not that high to report truthfully), she chooses to lie partially.
- True state  $\omega_i$ : if the true state is relatively bad, i.e.  $\omega_i$  is small enough, the difference between the actual state  $\omega_i$  and the best state  $\bar{\omega}$  is rather large. A large difference offers a lot of scope for lying but also means that lying costs may potentially get very high. Therefore, partial lying is more likely in the bad true states. If, on the contrary, the true state is very good, the lying costs to reach  $\bar{\omega}$  are quite small, so lying to the full extent is more likely.
- Curvature of the lying cost function, i.e.  $\frac{d^2 c(\omega_i, r_i)}{dr_i^2}$ : if marginal costs of lying increase steeply, the agent will engage in partial lying. If, on the contrary,  $\frac{dc(\omega_i, r_i)}{dr_i}$  is increasing in  $r_i$  rather slowly, individuals will choose to lie all the way up to  $\bar{\omega}$ .
- Marginal utility from social image: if the agent cares a lot about social image, and every additional score point brings her a lot of utility, she is more likely to engage in misreporting all the way up to  $\bar{\omega}$ .
- Marginal value of increasing the gain, reducing the loss or shifting from loss to gain in the social image domain

**Proposition 3.** *Individuals lie fully if  $\theta_i \leq \theta_i^{part}$  and  $\omega_i < \bar{\omega}$ .*

*Proof.* The agent chooses to lie fully if reporting  $r_i = \bar{r} = \bar{\omega}$  yields marginal costs the same or lower than marginal benefits of lying:

$$\underbrace{\frac{dv(s_{i2} + \bar{\omega} - s_{i1})}{dr_i}}_{MB \text{ of lying}} + \frac{du(s_{i2} + \bar{\omega})}{dr_i} \geq \underbrace{\theta_i \frac{dc(\omega_i, \bar{\omega})}{dr_i}}_{MC \text{ of lying}}.$$

By rearranging with respect to  $\theta_i$  we get

$$\theta_i \leq \theta_i^{part} = \left( \frac{dc(\omega_i, r_i)}{dr_i} \Big|_{r_i = \bar{\omega}} \right)^{-1} \left( \frac{dv(s_{i2} + r_i - s_{i1})}{dr_i} \Big|_{r_i = \bar{\omega}} + \frac{du(s_{i2} + r_i)}{dr_i} \Big|_{r_i = \bar{\omega}} \right).$$

Therefore, if the agent is quite insensitive to lying, i.e.  $\theta_i \leq \theta_i^{part}$ , she will lie all the way up to  $r_i = \bar{\omega}$ .  $\square$

**Lemma 2.** *Individuals lie partially if  $\theta_i \in (\theta_i^{part}, \theta_i^{true})$  and  $\omega_i < \bar{\omega}$ .*

If the agent's lying sensitivity is high enough not to engage in full lying but still not high enough to be willing to report truthfully, she will engage in partial lying. She will report a state which is between the true state and the best possible state, i.e.  $r_i \in (\omega_i, \bar{\omega})$ .

Next, we want to analyze the behavior with respect to gains and losses in social image concerns.

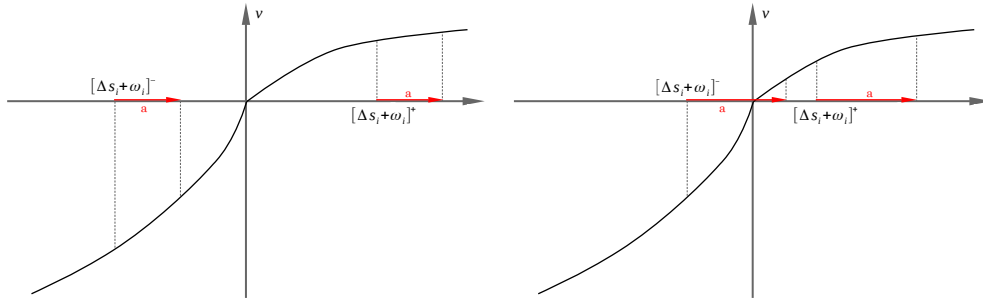
**Proposition 4.** *There is more incentive to lie if an agent experiences a loss in social image than a gain in social image of the same size.*

*Proof.* The next step is to show that individuals lie more if they experience a loss in social image, i.e. Condition 2 is stricter if  $\Delta s_i > 0$ . Rearranging the condition yields:

$$\frac{dv(s_{i2} + r_i - s_{i1})}{dr_i} = \theta_i \frac{dc(\omega_i, r_i)}{dr_i} - \frac{du(s_{i2} + r_i)}{dr_i}.$$

Note that the right hand side depends on  $r_i$  but not on  $\Delta s_i$ . The incentive to lie is higher if the left hand side gets higher. We compare cases denoted  $(\Delta s_i + \omega_i)^+$  and  $(\Delta s_i + \omega_i)^-$  in which  $(\Delta s_i + \omega_i)^+ =$

Figure 1: Value function



$-(\Delta s_i + \omega_i)^-$ . Those cases are driven by changes in  $s_{i1}$  or  $\omega_i$ , i.e. holding  $s_{i2}$  constant, and they both imply zero lying costs and symmetry. We assume that an individual makes a lying decision after observing true state  $\omega_i$ . It follows from Proposition 1 that individuals will not lie downwards and therefore we only consider the case of  $r_i \geq \omega_i$ . We know that for  $r_i = \omega_i$  the following holds:

$$v'((\Delta s_i + \omega_i)^+) < v'((\Delta s_i + \omega_i)^-) \quad (3)$$

(see Figure 1) as well as that the value function is convex for losses, i.e. for  $a > 0$  it is true that

$$v'((\Delta s_i + \omega_i)^-) < v'((\Delta s_i + \omega_i + a)^-),$$

and concave for gains, such that

$$v'((\Delta s_i + \omega_i)^+) > v'((\Delta s_i + \omega_i + a)^+).$$

Then condition 3 also holds for  $r_i = \omega_i + a$ :

$$v'((\Delta s_i + r_i)^+) < v'((\Delta s_i + r_i)^-) \quad (4)$$

and therefore reporting  $r_i = \omega_i + a > \omega_i$  is more attractive if an individual is in the loss domain than the gain domain. Note that if  $a$  is sufficiently large,  $v((\Delta s_i + \omega_i)^-) < 0$  but  $v((\Delta s_i + r_i)^-) > 0$  which means that the agent has been in the loss domain before reporting but has entered the gain domain by overreporting. Condition 4 still holds in this case.  $\square$

### 3 Experimental design

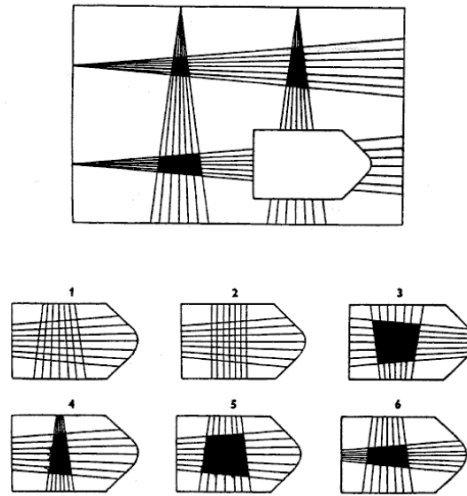
**General setup** We attempt to create social image concerns through reporting of a subject's ranking in a standardized test of fluid intelligence, i.e. logical reasoning, and abstract thinking—Raven's matrix test (Raven, 1983). As intelligence is an image providing trait for university students—our sample—reporting of results shall create social image utility. In order to strengthen this link we explicitly mention in the instructions that the matrices are designed to measure fluid intelligence, and that it is an important part of an individual's overall IQ, and that such or related tasks are often employed in recruitment processes.<sup>5</sup> Figure 2 provides an example of a Raven's Progressive Matrix (RPM). Subjects have to choose that box below the picture puzzle which is the best logical fit to the empty box within the picture. Progressive means that the matrices are increasing in difficulty.

Our experiment consists of two stages. Stage 1 is designed to establish a personal reference point for social image utility—the reported rank—against which the subjects can fall short of or improve their image in Stage 2. The second part induces an exogenous change of the rank by design. Subjects are then informed about their true rank and offered scope to manipulate the reporting of their rank to their

<sup>5</sup>Moreover, our approach is similar to Falk and Szech (2019), Ewers and Zimmermann (2015) and Burks et al. (2013) who also use the performance in IQ or knowledge tests to induce image concerns.

peers. We test whether (i) subjects whose rank deteriorates—what we associate with a loss in social image—misreport their rank more than those who experience an increase in their rank and (ii) whether lying depends on the reference point from the first stage—namely the question of how high is a potential fall?

Figure 2: Example of a Raven’s progressive matrix



**Procedures** The original Raven’s Progressive Matrices test consists of 60 matrices that are divided into 5 equally sized sets (A to E) which increase in difficulty. In our design, we do not use the 12 matrices of the easiest set A since we expect our student subjects to solve them all correctly. We split the remaining 48 matrices in two parts consisting of 24 matrices each. One part is easier (*Easy*) and the other is harder (*Hard*). While both parts contain tasks from sets B to E, we calibrated the two sets such that *Hard* has a higher likelihood to contain matrices that have been solved by fewer subjects in a reference sample.<sup>6</sup> Subjects of the reference group solved exactly the same overall 48 matrices as our subjects.<sup>7</sup> In both parts, the difficulty of the tasks is gradually increasing over time. Matrices in part *Easy* and *Hard* do not repeat or overlap.

At the beginning of each session, two subjects per session are randomly assigned the role of peer observers. We randomly draw one observer from all male subjects and the other from all female subjects. We did so in order to avoid possible gender-specific observer effects. After the observers have been determined, they stand up in front of the other subjects and announce “I am one of the two observers”.

**Treatments** The remaining subjects are randomly assigned to one of two treatments that vary the sequencing of the parts over the two stages of the experiment. We label the treatments as *HardEasy* and *EasyHard* where subjects work on the *Hard* part in Stage 1 and the *Easy* part in Stage 2 and the *Easy* part in Stage 1 and the *Hard* part in Stage 2, respectively. This implies that at the end of the experiment, subjects in both treatment groups have worked on the exactly same 48 matrices. All subjects (including the observers) read the same instructions. Figure 3 illustrates the timeline of the experiment.

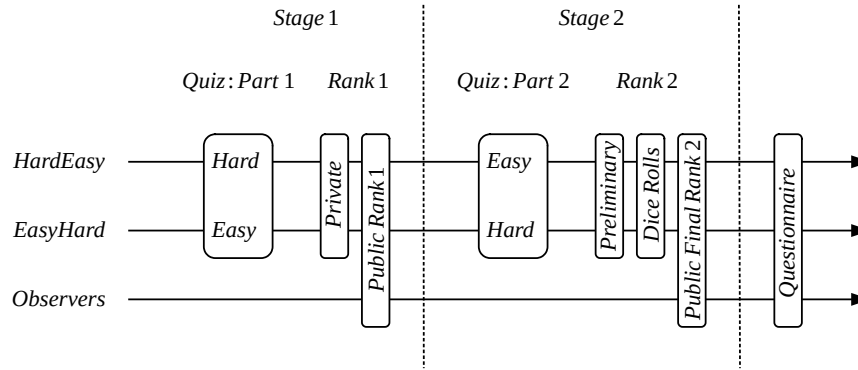
**Reporting in Stage 1** After completing the treatment specific tasks, subjects received private feedback on their relative performance (i.e. Rank 1) on their screen telling them that “ $X$  % of the participants of the reference group have a higher rank than you in part 1”. A higher rank (lower  $X$ ) implies better relative performance. To determine the rank, we compare the share of correctly solved matrices among the first 24 matrices to the distribution of the share of correctly solved matrices among all 48 matrices of

<sup>6</sup>The reference sample includes 413 observations (students) from a previous experiment which took place at the same lab in 2014.

<sup>7</sup>Part *Easy* consists of the following matrices: B1, B5, B6, B7, B8, B9, B10, B11, B12, C1, C2, C3, C7, C8, C9, C10, C12, D2, D3, D5, D7, E2, E6 and E11. Part *Hard* contains the following matrices: B2, B3, B4, C4, C5, C6, C11, D1, D4, D6, D8, D9, D10, D11, D12, E1, E3, E4, E5, E7, E8, E9, E10 and E12.



Figure 3: Timeline



the predetermined reference sample. Our calibration of the matrix distribution between *Easy* and *Hard* ensures that subjects in treatment *EasyHard* will on average rank better than subjects in treatment *HardEasy* in Stage 1 since both groups are compared to the same reference sample but the first 24 matrices are easier for subjects in treatment *EasyHard* than in *HardEasy*.

Subjects report their rank in the first stage to the observers. This establishes the individual Rank 1 as a personal reference point for social image concerns. We give both subjects and observers detailed instructions on the reporting procedure to control the reporting process and keep it as similar as possible across sessions and observers. We instruct subjects to fill in the report sheet named “Rank 1” and “Rank 2” in Stages 1 and 2, respectively, and present these sheets to observers who verify the report. We allow no further verbal communication between subjects and observers, i.e. the entire reporting procedure happens in silence. Report sheets contain two pieces of information: a 4-digit individual code and a rank. After each Stage, observers see a table on their computer screen in which each individual code corresponds to a rank, and thus can compare the report sheet to the true information from the table. If the reported rank matches the true rank, observers stamp the report sheet to verify it.<sup>8</sup>

**Reporting in Stage 2** Subjects work on the remaining 24 matrices. For subjects in treatment *EasyHard*, Part 2 is more complicated than Part 1. Thus, on average, we expect them to be ranked worse than in Part 1, such that their rank deteriorates. For subjects in treatment *HardEasy* the average rank shall improve compared to Rank 1. We now construct a *Preliminary* Rank 2, using the all 48 matrices comparing to the reference group. Consequently, we do not expect on average over the two treatments that the Preliminary Rank 2 differs. After completing the task in this stage, both Rank 1 and the Preliminary Rank 2 are displayed privately to each subject, so that subjects can compare their ranking in the two stages. While average Preliminary Rank 2 does not differ systematically across treatments, subjects’ average reference point (Rank 1) will be better in treatment *EasyHard* than *HardEasy*.

**Misreporting possibility** After learning about their ranks, subjects are asked to throw a dice twice and report the rolled numbers. Each subject has a dice in the cabin, such that no one, including the experimenters, can observe the actually rolled numbers. The first reported number is added to the number of correctly solved matrices in the reference group. The second reported number is added to a subject’s own number of correctly solved matrices.<sup>9</sup>

<sup>8</sup>We organized our laboratory setup in a way that subjects cannot see observers’ computer screens while reporting their rank. Additionally, to assure anonymity, we use 4-digit individual codes instead of cubicle numbers which, in the unlikely case of a subject seeing the table on the observer’s screen, makes it uninformative. Report sheets contain rank information in the same form as one’s private feedback, i.e. each subject fills in the blank space in the sentence: “\_ % of the participants of the reference group have a higher rank than me in part 1”. Examples of filled in and verified report sheets (in German) as well as their translations to English are shown in the Appendix (see Figures 9-10 and 11-12 for Ranks 1 and 2, respectively).

<sup>9</sup>This task is a modified version of the dice roll task by Fischbacher and Föllmi-Heusi (2013). In Fischbacher and Föllmi-Heusi, 2013, subjects roll a dice once, report some supposedly rolled number (not necessarily the truly rolled number), and are paid according to the reported number (i.e., higher numbers give a higher payoff except for 6, which pays zero). In the dice roll task, lying cannot be detected at the individual level. However, the underlying distribution of true dice roll

Including two dice rolls instead of only one allows us controlling for several aspects. First, the subject's final rank can either be better or worse than the Stage 1 rank. Adding a smaller number to the reference sample's score than to the own score will improve a subject's Final Rank 2 compared to Preliminary Rank 2, and vice versa. Second, if subjects have a preference for telling the truth, two dice rolls help to satisfy various preferences for truth-telling: Subjects can, for example, tell the truth about the dice rolls to the experimenters by reporting the actual numbers they have rolled. Alternatively, subjects can tell observers the truth about their Preliminary Rank 2 by reporting the same number for both dice rolls such that the Final Rank 2 is exactly the same as their Preliminary Rank 2. This option would not be possible with only one dice roll.

Additionally, if we based the ranking system on comparing subjects only within the current experiment (for example, ranking them from best to worst scores), there would be an incentive to add a higher number to the own score if subjects expect others to add a high number to their score. So, subjects' lying behavior would depend on their beliefs on others' lying. In order to avoid this and to be able to interpret lying as a reflection of image concerns independent of individual beliefs, it is important to construct a ranking system which compares subjects to a predetermined reference group one by one.

Observers only know about the existence of a "further task" on top of the second part of the matrices quiz in Stage 2 and that the score in this task will feed into a subject's Final Rank 2. Observers are not informed about the exact nature of the dice roll task and this is common knowledge to all subjects. Consequently, subjects do not risk losing social image because of possible reputation cost of being seen as a liar. The remaining subjects receive the instructions regarding the dice roll task on their computer screen after they have worked on part 2 of the quiz.

Once the reported dice rolls have been added and Final Rank 2 calculated, subjects go to observers again and report their Final Rank 2. After Stage 2, observers' information tables include, for each subject, the individual code, Final Rank 2, Rank 1 and the difference between Final Rank 2 and Rank 1. This is common knowledge for all subjects. Reporting procedures are the same as in Stage 1.

**Procedural details and implementation** Our experimental design and hypotheses are preregistered on AEA RCT Registry.<sup>10</sup> We conducted our experiment using zTree (Fischbacher, 2007). After two pilot sessions as a prerequisite for power calculations, we run 19 main experimental sessions in the DICE Lab, University of Düsseldorf between November 2018 and November 2019.

383 subjects participated, 38 as observers. 177 subjects (51%) were randomly assigned to treatment *HardEasy* and the remaining 168 (49%) to treatment *EasyHard*. We randomized within each session in order to balance the two treatments with respect to possible confounding factors such as day of the week, time of the day or weather. Our sample mainly consists of a student population and was recruited using ORSEE (Greiner2004). 142 subjects were male (67 in treatment *HardEasy* and 75 in *EasyHard*), 203 were female (110 in treatment *HardEasy* and 93 in *EasyHard*). Age varied between 18 and 63 years with a median age of 23 years and 95% of subjects being younger than 33 years. No particular exclusion criteria applied.

All participants received a show-up fee of 12 Euro, but no additional performance-contingent payment for correctly solving the matrices. This enables us to test whether solving matrices is indeed an image-relevant task: Subjects' behavior indicates image concerns if they exert effort to solve the matrices correctly, although this does not increase their monetary payoff.

Subjects have 30 seconds to work on each matrix. The time limit ensures that performance is comparable across subjects: both within our experiment and with respect to the reference sample we use, in

---

outcomes is known such that it can be observed whether subjects lie as a group at large. We build on the original dice roll task but adjust it for our purposes in two respects. First, instead of using monetary payoffs, we reward subjects with additional points which add up to the number of correctly solved matrices. Thus, lying enables subjects to improve their rank. Second, our subjects are told to throw the dice twice.

<sup>10</sup>Petrishcheva, Vasilisa, Gerhard Riener and Hannah Schildberg-Hörisch. 2019. "Loss Aversion in Social Image Concerns." AEA RCT Registry. April 09. <https://doi.org/10.1257/rct.3422-5.0>.

which subjects also had 30 seconds to work on each matrix. On average, it took subjects 11.5 seconds to answer a matrix. 2.7% of answers were provided in the last 5 seconds and in only 0.7% of cases subjects ran out of time, which suggests that the time limit was not generally binding. For each correctly solved matrix, subjects get one point. Wrong answers or no answer within the 30 seconds time limit do not give any points.

**Post-experimental questionnaire** The questionnaire provides information on socio-economic and demographic characteristics (age, gender, high school GPA, last math grade at school, student status and field of study, previous participation in experiments). It also assessed subjects’ general willingness to take risks, based on a question from the German Socio-Economic Panel (GSOEP) questionnaire as well as the importance of social image, using the question “How important is the opinion that others hold about you to you?” from the Multidimensional Personality Questionnaire (MPQ, Tellegen and Waller (2008), also used by e.g., Ewers and Zimmermann (2015) and Burks et al. (2013)). Additionally, following Gächter et al. (2007) and Fehr and Goette (2007), we measure loss aversion in the monetary domain using a set of incentivized lotteries which subjects can choose to accept or decline. Subjects received on average €12.65, which includes €12 flat payment plus one lottery outcome. In total, the experiment lasted about 90 minutes (including payment). Appendix A.4 provides the exact wording of all items.

## 4 Hypotheses

First, we want to test that our RPM-based task is indeed image-relevant for our subjects. Since their payment is unrelated to performance, exerting effort on solving the matrices will provide evidence for the relevance of either social and/or self-image concerns in our experimental design. This leads us to Hypothesis 1a.

**Hypothesis 1.** (Image relevance of task)

- (a) *Subjects will exert substantial effort on solving the matrices.*
- (b) *In both treatments, subjects will on average overreport their score.*

If subjects report higher dice rolls for themselves than for the reference group to be able to report a better Final Rank 2 to the observers, this establishes the relevance of social image concerns for our subjects as a whole.

**Hypothesis 2.** (Unconditional loss aversion in social image concerns) *Subjects in treatment EasyHard report significantly higher dice roll differences than subjects in treatment HardEasy.*

We hypothesize that subjects in treatment *EasyHard* (who on average experience a loss in social image since their rank deteriorates from Part 1 to Part 2) lie more than subjects in treatment *HardEasy* (who on average experience a gain in social image since their rank improves from Part 1 to Part 2). We compare the average difference in dice roll reports (average reported number to be added to own performance minus average reported number to be added to the reference group’s performance) from treatments *HardEasy* and *EasyHard*. If this difference is significantly higher in treatment *EasyHard* than in treatment *HardEasy*, this provides evidence for loss aversion in social image concerns unconditional on reference point because it implies that subjects who risk losing social image are ready to lie more than those with social image gains.

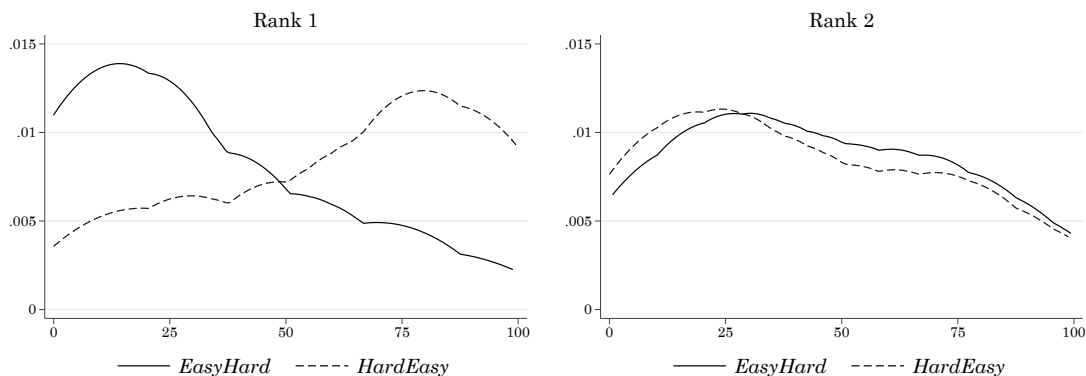
**Hypothesis 3.** (Loss aversion in social image concerns conditional on reference point) *Subjects in treatment EasyHard with a good Rank 1 report significantly higher dice roll differences than subjects in treatment HardEasy.*

We also test for loss aversion in social image concerns conditional on reference point: We hypothesize that subjects who performed well in Part 1 resulting in a good Rank 1 (reference point), i.e. those who have reputation to lose, are more strongly loss averse than those who have less reputation to lose.

## 5 Results

As intended by our design, subjects in treatment *HardEasy* performed worse in Part 1 than subjects in treatment *EasyHard*. On average, subjects in treatment *HardEasy* solved 2.6 matrices less in Part 1 than in treatment *EasyHard*. This then also reflects in the average Rank 1 in treatment *HardEasy* that was 61.48%, while it was 29.74% in treatment *EasyHard*. Figure 4 displays the kernel densities of Rank 1 (left) and 2 (right) by treatment. The difference of Rank 1 distributions between treatments is highly significant (Mann-Whitney  $U$  test, MWU,  $p < 0.0001$ ).<sup>11</sup> Thus, the exogenous manipulation of Rank 1, the reference point for social image concerns, worked as expected.

Figure 4: Distributions of Rank 1 and Preliminary Rank 2 by treatment



Note: Densities are estimated using Epanechnikov Kernels with a bandwidth of 15.

The total number of solved matrices after completion of Part 2 is similar across the treatments. On average, subjects in treatment *HardEasy* and *EasyHard* solve 39.2 and 38.5 matrices, respectively, which results only in a small, average difference in Preliminary Rank 2 of 3.7 percentage points between treatments. The difference in distributions of Preliminary Rank 2 between treatments is not significant (MWU test,  $p = 0.2027$ ). This ensures that possible differences in average lying across treatments do not reflect differences in Preliminary Rank 2 but only in the reference point for social image concerns, Rank 1.

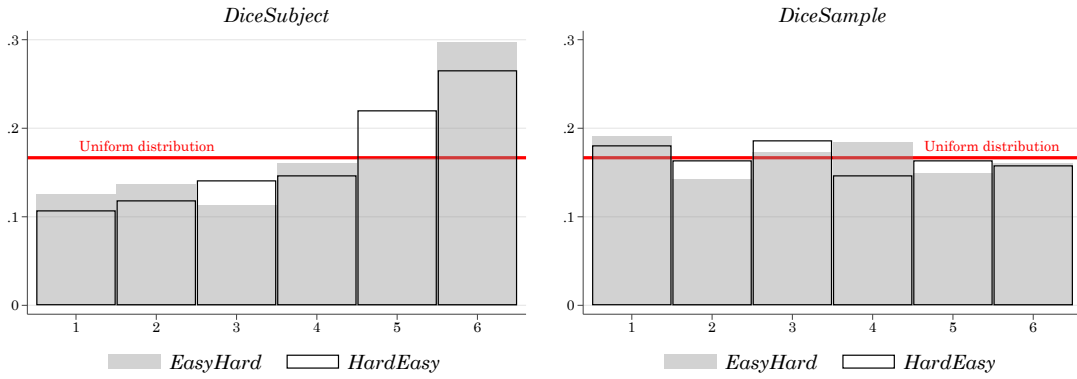
Moreover, the numbers above underline that participants exerted substantial effort on the quizzes. They solved an average of 38.8 out of all 48 matrices correctly and no subject solved less than 20 matrices, and more than 90% of subjects gave 34 or more correct answers. Since correct answers are not incentivized monetarily, substantial effort provision suggests that solving the Raven’s matrices is an image relevant task.

How do subjects report their dice rolls? Remember that they reported two values: The variable *DiceSubject* which is added to their own score and the variable *DiceSample* which is added to the scores of all subjects in the reference sample. In the absence of lying, dice roll reports for each of the variables should follow a discrete uniform distribution between 1 and 6 with an average of 3.5. Figure 5 displays histograms of *DiceSubject* (left) and *DiceSample* (right) as well as the uniform distribution (red line). The Figure shows that *DiceSubject* has an average of 4.03 and we can reject the null hypothesis for the point prediction ( $t$ -test,  $H_0: \text{DiceSubject} = 3.5$ ,  $p < 0.0001$ ). The distribution of *DiceSubject* is also significantly different from the discrete uniform distribution (Pearson’s  $\chi^2$ -test,  $p < 0.0001$ ) and left-skewed. In contrast, *DiceSample* has an average of 3.43 which is not significantly different from 3.5 ( $t$ -test,  $p = 0.4614$ ). Moreover, the distribution of *DiceSample* does not differ significantly from the discrete uniform distribution (Pearson’s  $\chi^2$ -test,  $p < 0.881$ ). Interestingly, while subjects do lie to add to their own score, they truthfully report the dice roll for the reference sample on average.

Subtracting *DiceSample* from *DiceSubject* results in the dice roll difference, *DiceDiff*, which indicates whether subjects improve or worsen their Final Rank 2 by lying. The higher *DiceDiff*, the more

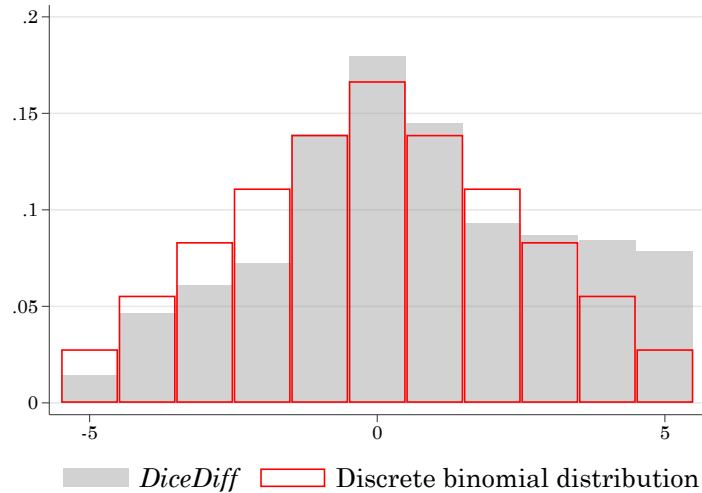
<sup>11</sup>Throughout the paper, we report two-sided tests and refer to results as (weakly/highly) significant if the two-tailed test’s  $p$ -value is smaller than 0.05 (0.10/0.01).

Figure 5: Distributions of *DiceSubject* and *DiceSample* by treatment



Final Rank 2 improves. In principle, *DiceDiff* can vary between -5 and 5, and, in the absence of lying, follows a discrete binomial distribution with a mean of zero. Pooling the data from both treatments, our subjects report an average dice roll difference of 0.59 which is significantly different from zero (*t*-test,  $p < 0.0001$ ). As illustrated in Figure 6, the values of 4 and 5 are significantly overreported (binomial probability tests, two-sided  $p = 0.0253$  and  $p < 0.0001$  for the values of 4 and 5, respectively).<sup>12</sup> In other words, subjects lie significantly and to their favor. Overreporting high values of *DiceDiff* provides further evidence that subjects perceive our matrices task as image-relevant and additionally shows that *social image concerns* matter.

Figure 6: Reported dice difference



**Result 1.** *On average, public reporting of own performance in the Raven’s matrices induces social image concerns. Subjects do not only exert substantial effort on solving the matrices, but are also willing to incur lying costs to positively misreport their score on average.*

Having established that the basic features of our design worked as intended, we now investigate whether subjects show loss aversion with regard to social image. Obviously, loss aversion in social image can only be observed for those subjects who indeed care about their social image and do so sufficiently strongly to incur lying costs. While we have shown above that many of our subjects do, it is also well documented that people are heterogeneous in the degree of their social image concerns (see, (bursztyn2017social) and (friedrichsen2018cares)) and lying costs (AbelerNosenzoRaymond:2019). This is also true in our sample, as Figure 13 in Appendix A.5 shows. Therefore, we present three sets of

<sup>12</sup>Thus, subjects engage both in a full lying (maximal overreporting) as well as partial lying (less than maximal overreporting).

Table 1: Regression analysis: dice roll difference (two-limit Tobit)

	Whole sample		Strong social image concerns		Weak social image concerns	
	(1)	(2)	(3)	(4)	(5)	(6)
EasyHard	-0.050 (0.388) [0.908]	0.392 (0.366) [0.298]	0.356 (0.529) [0.528]	1.522** (0.684) [0.026]	-0.440 (0.514) [0.431]	-0.828 (0.607) [0.189]
EasyHard × Rank 1		-0.008 (0.009) [0.387]		-0.024* (0.011) [0.050]		0.009 (0.014) [0.513]
Rank 1		0.006 (0.005) [0.362]		0.016** (0.007) [0.016]		-0.004 (0.009) [0.710]
Constant	0.695*** (0.242)	0.319 (0.313)	0.425 (0.353)	-0.528 (0.470)	0.912*** (0.278)	1.156* (0.655)
Number of obs.	345	345	173	173	172	172

Note: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  based on score-bootstrap  $p$ -values. In columns (3) to (6), we split our sample based on the median of the importance of social image concerns that subjects reported on a 11-point Likert-scale. Individuals who reported 6 or higher are categorized as “Strong Social Image” and as “Weak Social Image” otherwise. Rank 1 ranges from 0 to 100, with lower values indicating better rank. All columns display two-limit tobit estimates with score bootstrap clustering at the session level. Standard errors clustered at the session level in parentheses. Score-bootstrap  $p$ -values in brackets.

results: on top of presenting evidence for our sample as a whole, we make use of our direct measure of social image concerns at the individual level. Based on the question “How important is the opinion of others to you?” (11-point Likert scale, see Ewers and Zimmermann (2015) and Burks et al. (2013)), we perform a median split and additionally present separate results for subjects with strong and weak social image concerns. Subjects who answer 6 or above are classified as individuals with strong social image concerns, the remaining subjects as having weak social image concerns.<sup>13</sup>

Table 1 displays our main results that are all based on two-limit tobit models to account for the censored nature of the dependent variable *DiceDiff*. Columns (1), (3), and (5) test for unconditional loss aversion in social image concerns for the various samples; columns (2), (4), and (6) for loss aversion in social image concerns conditional on the reference point, Rank 1. Standard errors are clustered at the session level in order to account for possible within-session correlations generated by observers. We use score bootstrap clustering with null imposed, Rademacher weights and 999 replications in order to account for the small number of clusters ( $N = 19$  sessions).

Regarding Hypothesis 2, we do not find evidence for unconditional loss aversion in social image concerns: while the constant terms indicate that subjects overreport their own score, subjects in treatments *HardEasy* and *EasyHard* do not differ significantly in their average overreporting (see columns (1), (3), and (5)). Similarly, MWU tests confirm that subjects’ overall lying behavior does not differ across treatments. They yield  $p = 0.9108$  for *DiceDiff*,  $p = 0.8970$  for *DiceSubject* and  $p = 0.9232$  for *DiceSample*, respectively.

**Result 2.** *We do not find evidence in favor of loss aversion in social image concerns irrespective of the reference point for social image.*

We continue by investigating Hypothesis 3 that postulates the existence of loss aversion in social image concerns conditional on having a high reference point in columns (2), (4), and (6) of Table 1. Intuitively,

<sup>13</sup>A MWU test shows that social image concerns do not differ significantly across treatments,  $p = 0.1514$ .

subjects who perform well in Part 1 and therefore have a reputation to lose in Part 2 may be more averse to losing social image than subjects who have no reputation to lose since they are ranked less favorable in Part 1. This is indeed what we find for subjects with strong social image concerns, but—perhaps unsurprisingly—not for those with weak social image concerns.

In particular, estimation results in column (4) imply that for high initial performance in Part 1 (high reference point which means low Rank 1) subjects in treatment *EasyHard* overreport substantially more than subjects in treatment *HardEasy* who on average have lower reference points. For the best possible Rank 1 of zero, the difference in overreporting to one’s own advantage is substantial: it amounts to 1.5 units on the dice roll difference scale from -5 to 5. Moreover, for each 10 percentage point increase in Rank 1 (decrease in reference point), e.g. moving from 10 to 20 in Rank 1, subjects in treatment *EasyHard* lie 0.08 units less ( $= (-0.016 + 0.024) \times 10$ ). Thus, in treatment *EasyHard*, subjects with higher reference points in social image do display loss aversion in social image, while those who have less reputation to lose do so much less. In other words, those who report to care about their reputation also experience loss aversion in this domain, and their loss aversion is stronger if their reference point for reputation is higher. Interestingly, we observe the exactly opposite pattern for subjects with strong social image concerns in treatment *HardEasy*: they overreport more strongly in Part 2, the worse their initial reputation in Part 1, i.e. the higher Rank 1. This explains why we do not observe a significant treatment difference in column (3), which tests for loss aversion in social image unconditional on the reference point.

Unsurprisingly, subjects who care less about their social image do not display loss aversion in social image concerns, independent of whether we do not or whether we do condition on their reference point for social image (see columns (5) and (6) of Table 1, respectively). This results in non-significant and substantially smaller coefficients for the sample as a whole (see columns (1) and (2) of Table 1, respectively).

**Result 3.** *We observe loss aversion in social image concerns only for those subjects who strongly care about their social image and have reputation to lose, i.e. who have a high initial reference point for social image.*

## Robustness checks

In our main regression analysis we split the sample according to the median of the importance of social image concerns that subjects reported on a 11-point Likert-scale. Individuals who reported 6 or higher are categorized as “Strong Social Image” and as “Weak Social Image” otherwise. In the following, we analyze how robust our findings are to alterations of this threshold.

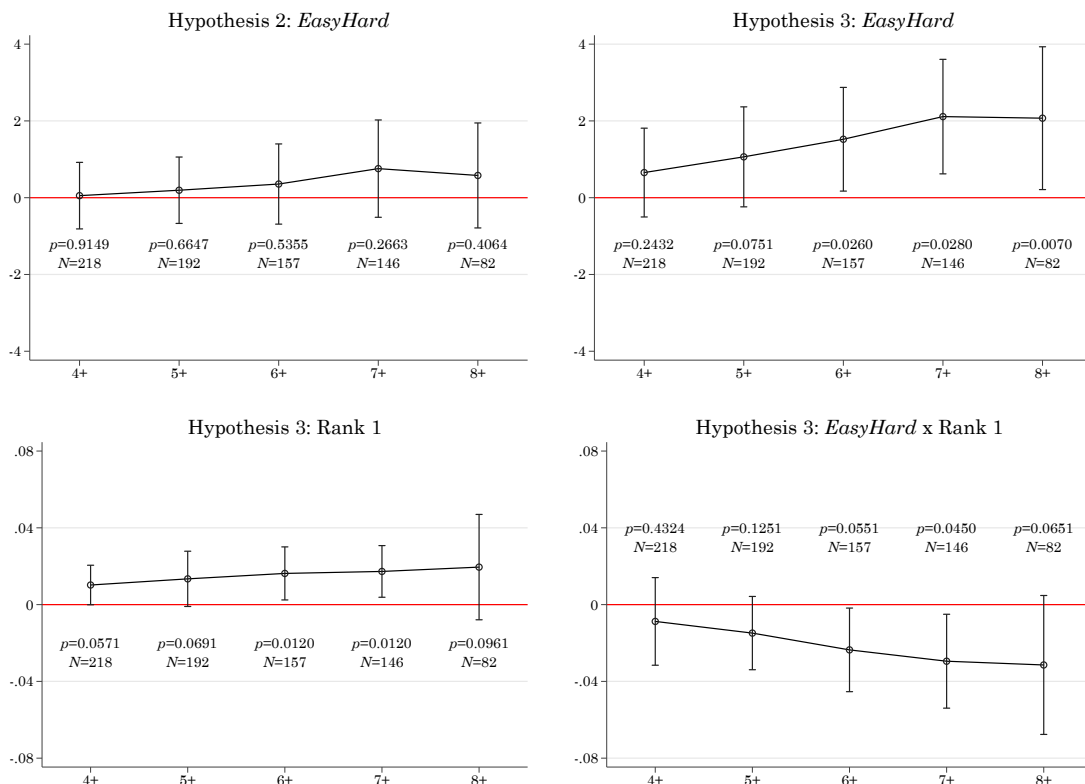
Figure 7 shows the coefficients for the treatment dummy *EasyHard*, Rank 1 and their interaction for an increasingly restricted sample. We indicate the sample as “4+” if it includes subjects who reported the importance of social image to be 4 or higher, “5+” for those who reported it to be 5 or higher, etc. Then, “8+” sample includes only subjects who are extremely concerned with their social image and reported its importance to be 8, 9 or 10.<sup>14</sup> For each of those subsamples, we estimate the same two-limit tobit models with standard errors clustered at the session level as reported in Table 1. We label a robustness check “Hypothesis 2” if its underlying regression corresponds to columns (1), (3), and (5) in Table 1 testing for unconditional loss aversion in social image concerns, and “Hypothesis 3” if it follows the specification of columns (2), (4) and (6), and tests for loss aversion in social image concerns conditional on one’s reference point represented by Rank 1.

Our findings are robust for a variety of subsamples. In particular, we do not find evidence for unconditional loss aversion irrespective of the threshold we apply for social image concerns (see upper left panel of Figure 7). However, conditional loss aversion is strongly supported by our robustness checks: subjects who have a high reference point (low Rank 1) lie significantly more on average in treatment

<sup>14</sup>Note that we do not consider subsamples of those who reported the importance of social image to be 9 or higher, or 10 due to a small number of observations: There are only 30 subjects in our data who reported the importance of social image to be 9 or 10.

*EasyHard* than in *HardEasy*, and the effect gets more and more pronounced for stronger image concerns (see upper right panel of Figure 7). Expanding our sample by adding subjects who report the importance of social image to be 7 and 6 as opposed to 8 or higher gradually reduces the average treatment effect (which, however, stays high and significant). After including those with image concerns of 5 and 4, the effect remains positive but becomes smaller and insignificant as we have previously documented for a sample as a whole. The estimated coefficient of Rank 1 remains rather stable for the various thresholds (see lower left panel of Figure 7). Finally, the estimated coefficient of the interaction term becomes more and more negative the stronger the social image concerns are (see lower right panel of Figure 7), gradually offsetting the larger level effect of *EasyHard* that is displayed in the upper right panel of of Figure 7.

Figure 7: Robustness check: individuals with strong image concerns, by threshold



*Note:* The vertical axis displays the magnitude of each respective coefficient along with a 95% confidence interval. Number of observations and score bootstrap  $p$ -values correspond to each estimated regression. The horizontal axis indicates the subsample used for the estimation. "4+" ("5+"/"6+"/"7+"/"8+") includes subjects who reported the importance of social image to be 4(5/6/7/8) or higher. Each reported coefficient and its 95% confidence interval is estimated using two-limit tobit models with standard errors clustered at the session level. We label a robustness check "Hypothesis 2" if its underlying regression corresponds to columns (1), (3), and (5) in Table 1 and "Hypothesis 3" if it follows the specification of columns (2), (4) and (6).

## 6 Conclusion

Does loss aversion apply to social image concerns? The answer provided by our experimental results is yes and no. In sum, we do not find evidence for loss aversion in social image concerns irrespective of the reference point in social image, but for loss aversion in social image conditional on a high initial reference point. In particular, we observe loss aversion in social image concerns for those individuals who strongly care about their reputation and have a reputation to lose, i.e. have a high initial reference point for social image. On average, however, individuals do not suffer more strongly from losses than they benefit from equally sized gains in social image. Given that individuals differ in the importance they assign to social



image concerns, it is conceivable that only individuals who sufficiently care about social image can suffer from loss aversion in the domain of social image.

More generally, our findings suggest that loss aversion can also play a role in the non-material domain. While loss aversion is a well-established phenomenon for money and material goods (Kahneman, Knetsch, et al., 1991), our findings take a first step in a new line of research investigating the relevance of loss aversion to non-material sources of utility such as various drivers of reputation or self-image.

Since our experimental paradigm quantifies utility changes due to changes in social image by the amount of lying that individuals are willing to engage in, our findings also speak to the manifold situations in which honest reporting of private information is of great importance but not necessarily incentive-compatible. Our results reveal that two groups of individuals are more likely to report more dishonestly than others: individuals with strong social image concerns who have a high reputation to lose and individuals with strong social image concerns and particularly low reputation. Thus, monitoring efforts should be targeted at those groups. Moreover, one should try to make it harder to lie while keeping a good reputation, e.g., via transparency, naming-and-shaming or reputation systems (see also Abeler et al., 2019).

## References

- Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2019). “Preferences for Truth-Telling”. In: *Econometrica* 87.4, pp. 1115–1153. DOI: 10.3982/ECTA14673. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14673>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA14673>.
- Andreoni, James and B Douglas Bernheim (2009). “Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects”. In: *Econometrica* 77.5, pp. 1607–1636.
- Ariely, Dan, Anat Bracha, and Stephan Meier (2009). “Doing good or doing well? Image motivation and monetary incentives in behaving prosocially”. In: *American Economic Review* 99.1, pp. 544–55.
- Austen-Smith, David and Roland G Fryer Jr (2005). “An economic analysis of “acting white””. In: *The Quarterly Journal of Economics* 120.2, pp. 551–583.
- Barberis, Nicholas C (2013). “Thirty years of prospect theory in economics: A review and assessment”. In: *Journal of Economic Perspectives* 27.1, pp. 173–96.
- Bleichrodt, Han, Jose Luis Pinto, and Peter P Wakker (2001). “Making descriptive use of prospect theory to improve the prescriptive use of expected utility”. In: *Management science* 47.11, pp. 1498–1514.
- Booij, Adam S and Gijs Van de Kuilen (2009). “A parameter-free analysis of the utility of money for the general population under prospect theory”. In: *Journal of Economic psychology* 30.4, pp. 651–666.
- Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini (2013). “Overconfidence and social signalling”. In: *Review of Economic Studies* 80.3, pp. 949–983.
- Bursztyn, Leonardo, Georgy Egorov, and Robert Jensen (2019). “Cool to be smart or smart to be cool? Understanding peer pressure in education”. In: *The Review of Economic Studies* 86.4, pp. 1487–1526.
- Bursztyn, Leonardo and Robert Jensen (2017). “Social image and economic behavior in the field: Identifying, understanding, and shaping social pressure”. In: *Annual Review of Economics* 9, pp. 131–153.
- Camerer, Colin F (1998). “Prospect theory in the wild: Evidence from the field”. In:
- Ewers, Mara and Florian Zimmermann (2015). “Image and misreporting”. In: *Journal of the European Economic Association* 13.2, pp. 363–380.
- Falk, Armin and Nora Szech (2019). *Competing Image Concerns: Pleasures of Skills and Moral Values*. Tech. rep. Working paper.
- Fehr, Ernst and Lorenz Goette (2007). “Do workers work more if wages are high? Evidence from a randomized field experiment”. In: *American Economic Review* 97.1, pp. 298–317.
- Fischbacher, Urs (2007). “z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental economics* 10.2, pp. 171–178.
- Fischbacher, Urs and Franziska Föllmi-Heusi (2013). “Lies in disguise—an experimental study on cheating”. In: *Journal of the European Economic Association* 11.3, pp. 525–547.
- Gächter, Simon, Eric J Johnson, and Andreas Herrmann (2007). “Individual-level loss aversion in riskless and risky choices”. In:
- Grolleau, Gilles, Martin G Kocher, and Angela Sutan (2016). “Cheating and loss aversion: do people cheat more to avoid a loss?” In: *Management Science* 62.12, pp. 3428–3438.
- Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1990). “Experimental tests of the endowment effect and the Coase theorem”. In: *Journal of political Economy* 98.6, pp. 1325–1348.
- (1991). “Anomalies: The endowment effect, loss aversion, and status quo bias”. In: *Journal of Economic perspectives* 5.1, pp. 193–206.
- Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk”. In: *Econometrica* 47.2, pp. 263–292.
- Khalmetski, Kiryl and Dirk Sliwka (2017). “Disguising Lies-Image Concerns and Partial Lying in Cheating Games”. In:
- Pennings, Joost ME and Ale Smidts (2003). “The shape of utility functions and organizational behavior”. In: *Management Science* 49.9, pp. 1251–1263.

- Raven, John C (1983). “Manual for Raven’s progressive matrices and vocabulary scales”. In: *Standard Progressive Matrices*.
- Schindler, Simon and Stefan Pfattheicher (2017). “The frame of the game: Loss-framing increases dishonest behavior”. In: *Journal of Experimental Social Psychology* 69, pp. 172–177.
- Soetevent, Adriaan R (2011). “Payment choice, image motivation and contributions to charity: evidence from a field experiment”. In: *American Economic Journal: Economic Policy* 3.1, pp. 180–205.
- Tellegen, Auke and Niels G Waller (2008). “Exploring personality through test construction: Development of the Multidimensional Personality Questionnaire”. In: *The SAGE handbook of personality theory and assessment* 2, pp. 261–292.
- Wakker, Peter P (2010). *Prospect theory: For risk and ambiguity*. Cambridge university press.

# A Appendix

## A.1 Proofs

**Proposition 5.** *Under the assumption of differentiable lying costs, individuals never report truthfully unless  $\omega_i = \bar{\omega}$ .*

*Proof.* Players face a trade-off: on the one hand, they can report  $r_i = \omega_i$  which results in zero lying costs, however, also does not bring additional benefit to the social image component and value function. On the other hand, they can report  $r_i > \omega_i$  which is going to improve her social image and the value function but requires lying costs. Note that if lying cost function is convex with a minimum at  $r_i = \omega_i$ , that implies

$$\left. \frac{dc(\omega_i, r_i)}{dr_i} \right|_{r_i = \omega_i} = 0$$

and therefore no one reports truthfully if  $\omega_i < \bar{\omega}$  because condition 2 should be set to zero in order to maximize one's utility which is only possible under  $r_i > \omega_i$ .  $\square$

## A.2 Instructions of the Experiment

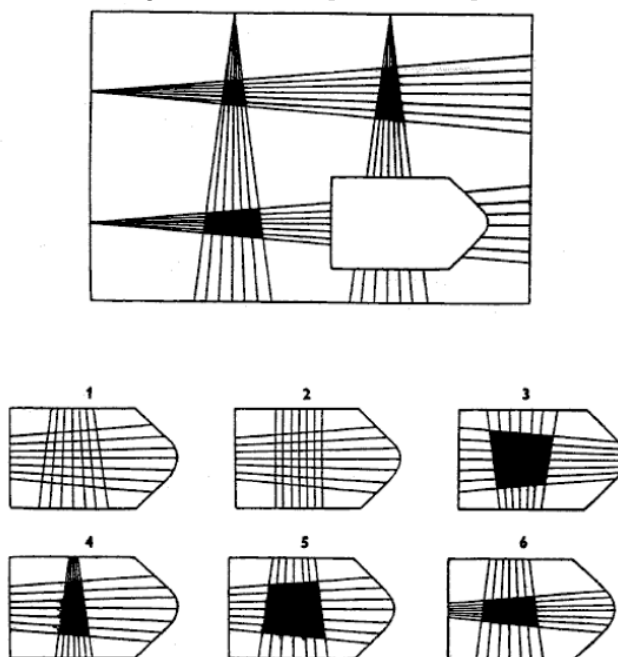
### General Instructions

We warmly welcome you to this economic experiment. Please read the following instructions carefully! If you have any questions, please raise your hand from the cubicle - we will then come to your seat. It is not allowed to talk to other participants of the experiment, use mobile phones or start other programs on the computer during the experiment. Non-compliance with these rules will result in exclusion from the experiment and all payments. You will receive a fixed payment of €12 for participating in this experiment, which will be paid in cash at the end of the experiment. On the following pages we describe the exact procedure of the experiment.

### Part 1 of the Experiment

Parts 1 and 2 consist each of 24 tasks, which are often used to measure so-called fluid intelligence of a person. The fluid intelligence is an important part of the general intelligence of humans. These or similar tasks are also often used by companies in the context of recruitment procedures. Each task corresponds to a picture puzzle. Here you can see an example:

Figure 8: Picture puzzle example



Each picture puzzle shows in its upper part a pattern in a box, in which a “piece of the puzzle” in the lower right corner is left out. Your task is to select one of the puzzle pieces listed below the box, which will logically fill the blank lower right corner of the pattern in the box. Please enter the number of the puzzle piece that you think fits best on the screen. The number of a puzzle piece is stated above each puzzle piece. There is always exactly one piece that fits best. You have 30 seconds to complete each picture puzzle. For each correctly completed picture puzzle you receive one point. As commonly done with intelligence tests, correct answers are not paid extra. You will receive 0 points for each wrongly answered picture puzzle or if you do not enter the best fitting piece of the puzzle within 30 seconds. After you have completed all 24 picture puzzles in Part 1, you will first receive a private feedback on your rank on the computer screen, indicating how well you performed in solving the picture puzzles. The feedback has the following form: “X % of the participants of the reference group have a higher rank than you in Part 1”. The reference group consists of 413 participants of a previous laboratory experiment conducted in 2014

here at the DICE Lab of the University of Düsseldorf, who have worked on the same picture puzzles as you do in the course of this experiment. So the feedback “9% of the participants of the reference group have a higher rank than you in Part 1” means that 9% performed better than you (i.e. solved a higher percentage of the total 48 picture puzzles from Parts 1 and 2 correctly than you) and 90% performed worse (i.e. solved a lower percentage of picture puzzles correctly than you). So you belong to the 10% of the best at answering the picture puzzles designed to measure individual fluid intelligence. The feedback “83% of the participants of the reference group have a higher rank than you in Part 1” means that 83% performed better and 16% worse than you. So you are among the 17% of the worst in answering the picture puzzles. Before Part 2 of the experiment starts, you have to inform two so-called “Observers” about your performance in the experiment. Please use the report sheet available in your cubicle. Your cubicle number is already entered. Please enter legibly the number, which you received as feedback on the computer screen, in the sentence “\_\_ % of the participants of the reference group have a higher rank than me in Part 1” in the report sheet “Rank 1”. Please enter your personal code, which is also displayed on the screen, in the free field next to it: “My personal code is \_\_”. Observers sit in the cubicles number 1 and 2 in the laboratory (directly in front of the entrance door). Please go there with the completed report sheet and show it silently to Observers as soon as your cubicle number is called by the experimenter. This ensures that each participant informs Observers individually without other participants knowing her/his rank. A two-column table will be displayed on the Observers’ computer screens, assigning each personal code the corresponding rank in Part 1. Each Observer will silently compare your report sheet with the information in the table and stamp it. Afterwards, please return to your cubicle in silence. Part 2 of the experiment will begin as soon as all participants have informed Observers of their rank.

### **The Different Participants in the Experiment**

At the beginning of the experiment, each participant randomly drew a chip with a number indicating his cubicle number. The cubicle numbers have the following additional meaning:

The participants who have randomly drawn cubicle numbers 1 and 2 have the role of “Observers” described above. Since the chips with even numbers were reserved for female participants and the chips with odd numbers for male participants, there is always one male Observer and one female Observer. These will introduce themselves to you shortly before the actual experiment begins by standing up and saying “I am one of the two Observers”. Observers — just like all other participants — will receive this printed explanation of the rules of the experiment, which you are reading, for information about the experiment. All other participants in the experiment with cabin numbers 3 or higher solve the picture puzzles described above. Each participant is randomly assigned to one of two groups: Group A or Group B. Throughout the whole experiment, all participants of both groups will solve exactly the same 48 picture puzzles, 24 in Part 1 and 24 in Part 2. The further task in part 2 of the experiment is also exactly the same for both groups. Only the order in which the picture puzzles are processed differs between group A and B. The group membership has no further meaning. In Parts 1 and 2 you belong to the same group.

### **Part 2 of the Experiment**

Part 2 of the experiment is very similar to Part 1. First you work on 24 more picture puzzles following the same rules (30 seconds time per puzzle, 1 point for correct answers, 0 points otherwise, etc.). After you have completed remaining 24 picture puzzles in Part 2, you will receive a private feedback on your preliminary rank in Part 2 on the computer screen, indicating how well you have done in the 48 picture puzzles in Parts 1 and 2. The feedback again has the following form: “X% of the participants of the reference group have a higher rank than you”. The reference group is again the 413 participants of a previous lab experiment here in the DICE Lab of the HHU from 2014, who have solved the same 48 picture puzzles as you. In addition, the rank you had in Part 1 of the experiment is displayed as a

reminder. The only difference to Part 1 is that you have one more task, which is also used to calculate your final rank in Part 2. You will then receive a private feedback on your final rank in Part 2, which is calculated based on the 48 picture puzzles in Parts 1 and 2 and your score in the further task in Part 2. Details of the further task and how exactly it is included in the calculation of the final rank in Part 2 will be explained on the computer screen during the course of the experiment. For calculation of your final rank the same reference group is used again as for your rank in Part 1 and the preliminary rank in Part 2. The detailed explanations of the further task in Part 2 are given only to the participants, but not to the two Observers. Just like at the end of Part 1, you still have to inform the two Observers about your performance, i.e. your final rank, in Part 2. Please use the report sheet which is available in your cubicle. In addition, under “Rank 2”, please enter legibly in the sentence “\_\_ % of the participants of the reference group have a higher rank than me”, which you have received as feedback on your final rank on the computer screen. Please go to two Observers with the completed report sheet and show it to them in silence as soon as your cubicle number is called up by an experimenter. This again ensures that each participant informs Observers individually without the other participants knowing her/his rank. A table with four columns is now displayed to Observers on your computer screen, which assigns to each personal code the corresponding rank in Part 1, the final rank and the difference in rank between the rank in Part 1 and the final rank. The observers will, also in silence, compare your report sheet with the information in the table and stamp it. Afterwards, please return to your cabin in silence.

### **End and Payment of the Experiment**

After Part 2 of today’s experiment, there will be some more screens with questions before we proceed to the payment of €12. We will call you individually by cubicle number for payment. If you have any questions now, please raise your hand out of the cubicle. Experiment supervisor will then come to your seat to answer your questions. Do not ask questions out loud!

### **A.3 Additional Instructions on the Computer Screen: Dice Roll Task**

There is a dice in your cabin. Please roll a dice twice in your cubicle. Please enter numbers between 1 and 6 which you rolled on the first and second roll of the dice on the computer screen. The first number you rolled is added to the number of correctly solved picture puzzles of each participant of the reference group. The second dice roll is added to the number of correctly solved picture puzzles. Your total output is equal to the number of correctly solved picture puzzles in Parts 1 and 2 of the 48 picture puzzles plus the number of points you entered on the second die roll. Your total score is used to calculate your final rank. Your overall performance is compared with the overall performance of the peer group. The total score of a competitor in the comparison group is equal to the number of correctly solved picture puzzles out of the 48 picture puzzles plus the number of points you entered on the first dice roll. Your final rank will be shown to Observers and you will report it to the Observers at the end.



## A.4 Questionnaire

Now please fill out the following questionnaire before we proceed to the payment. Please enter the following personal data:

If you want to enter decimal numbers, please use a dot (.) instead of a comma (,).

Age

Gender (male/female)

Final grade point average at high school (Abiturnote) (1.0-6.0)

Last math grade (1.0-6.0)

Last German grade (1.0-6.0)

Field of study/job

How much money do you have available each month (after deducting fixed costs such as rent, insurance, etc.)?

How much money do you spend each month (after deducting fixed costs such as rent, insurance, etc.)?

In how many economic science experiments have you (approximately) already participated?

On a scale of 0 to 10, how would you rate your willingness to take risks? 0 means not willing to take risks at all and 10 means completely willing to take risks.

How important is the opinion that others hold about you to you? Please answer on a scale of 0 to 10, where 0 is not important at all and 10 is extremely important.

Have you ever solved similar tasks as the picture puzzles before? (Yes/No)

If so, how long ago approximately? Please indicate the approximate number of months.

Below, please answer a few more questions about lotteries in which you can earn or lose money in addition to the €12 if you decide to accept the lotteries. Listed below are 6 different lotteries. For each of the 6 lotteries you can choose whether to accept or decline the lottery. If you choose to decline a lottery, your payout will not change. If you accept a lottery, you will realize either an additional gain or an additional loss based on the €12. At the end of the experiment, one of the 6 lotteries is randomly selected. So you should make each lottery decision as if it was your only decision. The selected lottery is then drawn to determine whether the additional gain or loss will be realized.

Lottery 1: With 50% probability you lose €2 and with 50% probability you win €6. (accept / reject)

Lottery 2: With 50% probability you lose €3 and with 50% probability you win €6. (accept / reject)

Lottery 3: With 50% probability you lose €4 and with 50% probability you win €6. (accept / reject)

Lottery 4: With 50% probability you lose €5 and with 50% probability you win €6. (accept / reject)

Lottery 5: With 50% probability you lose €6 and with 50% probability you win €6. (accept / reject)

Lottery 6: With 50% probability you lose €7 and with 50% probability you win €6. (accept / reject)

## A.5 Additional Figures

Figure 9: Rank 1 report sheet

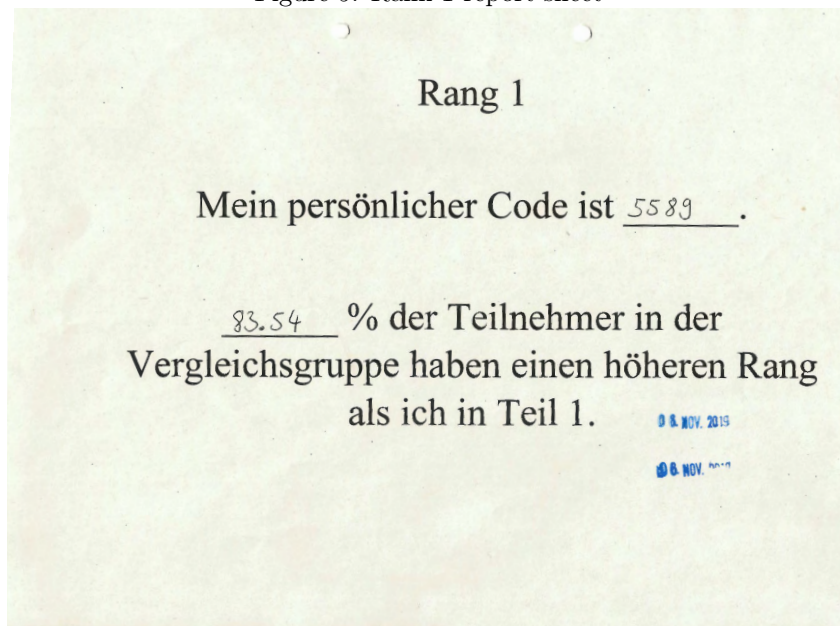


Figure 10: Rank 1 report sheet (translated to English)

**Rank 1**

My individual code is \_\_\_\_\_.

\_\_\_\_\_ % of the participants of the  
reference group have a higher rank than me  
in part 1.

Figure 11: Rank 2 report sheet

Rang 2

Mein persönlicher Code ist 5589.

54.00 % der Teilnehmer in der  
Vergleichsgruppe haben einen höheren Rang  
als ich.

13. NOV. 2019  
10. NOV. 2019

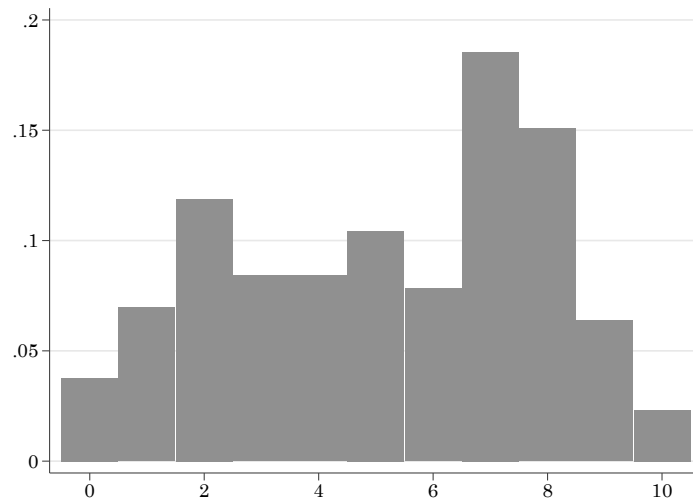
Figure 12: Rank 2 report sheet (translated to English)

Rank 2

My individual code is \_\_\_\_\_.

\_\_\_\_\_ % of the participants of the  
reference group have a higher rank than me.

Figure 13: Self-reported importance of social image



*Note:* Importance of social image concerns is measured on a 11-point Likert scale based on the question “How important is the opinion that others hold about you to you?” from the Multidimensional Personality Questionnaire (MPQ, Tellegen and Waller (2008), also used by e.g., Ewers and Zimmermann (2015) and Burks et al. (2013)).