# Willful Ignorance and Reference Dependence of Self-Image Concerns*

Vasilisa Petrishcheva†

October 25, 2022

**Abstract**

Self-image concerns play a crucial role in economic decision-making. I conduct a laboratory experiment and offer a theoretical framework to examine whether self-image concerns are reference-dependent and whether individuals react differently to gains and losses in self-image. I focus on intelligence, a self-image-relevant domain, and find that individuals update their beliefs about performance more strongly if they experience losses than gains in self-image. On average, individuals tend to avoid self-image-relevant feedback in case of gains and losses in self-image. However, the willingness to acquire self-image-relevant feedback increases if the difference between the posterior and the prior beliefs about performance increases.

**JEL Codes:** C91, D91

**Keywords:** Willful ignorance; Information avoidance; Self-image concerns; Motivated beliefs; Laboratory experiment

**AEA RCT Registry:** Petrishcheva, Vasilisa. 2021. "Willful Ignorance and Reference-Dependence of Self-Image Concerns." AEA RCT Registry. June 09.

*I am grateful to Hannah Schildberg-Hörisch, Paul Heidhues, Joël van der Weele, Mats Köster, Jens Prüfer, Gerhard Riener, Christopher Snyder, Lisa Bruttel, Maximilian Andres, Tobias Werner and the participants of the Economic Science Association Global Meeting 2021, the Berlin Behavioral Economics Workshop and the DICE Brown Bag Seminar for helpful comments and suggestions. I also thank Marlene Merker for her excellent research assistance. Financial support from DFG through the grant for GRK 1974 Competition Economics is gratefully acknowledged.

†University of Potsdam, August-Bebel-Straße 89, 14482 Potsdam, Germany; E-mail: petrishcheva@uni-potsdam.de.

# 1 Introduction

Information is a key element in most economic decisions. Individuals tend to seek certainty and avoid ambiguity. Yet, in many situations, people prefer to deliberately avoid information and remain willfully ignorant.[1] Examples of information avoidance range from everyday interactions to high-stake decisions with long-term effects. For instance, individuals may not want to learn that the holiday season made them put on some weight or that there is a better deal for a recent purchase (Sweeny et al., 2010). In a health context, people tend to avoid learning about their genetic risks for cancer or Huntington's disease (Oster et al., 2013; Ropka et al., 2006) and their HIV status even when offered monetary incentives to do so (Thornton, 2008). In a finance context, investors tend to monitor their portfolios and balances closely on "good days", e.g., on paycheck days or when the market goes up, and avoid logging into their accounts on "bad days" (Karlsson et al., 2009; Olafsson et al., 2018). In a workplace, managers often forego helpful feedback to avoid learning that their earlier decisions were incorrect because they want to maintain their professional self-image (Deshpande and Kohli, 1989; Schulz-Hardt et al., 2000; Zaltman, 1983). In the case of prosocial behavior, individuals often prefer to remain uninformed about the actual effectiveness of their altruistic actions or charitable donations and carry on a feeling of warm glow due to the fact of their deed but not to the impact on its recipient (Niehaus, 2014). Similarly, people tend to avoid learning about the potential harm their actions may yield to others (Dana et al., 2007; Grossman and Van Der Weele, 2017; Serra-Garcia and Szech, 2022).

This paper studies the implications of changes in self-image for the demand for feedback as well as the evolution of the self-image itself. I conduct a laboratory experiment to analyze individuals' willingness to avoid self-image-relevant feedback after having them work on more difficult or easier tasks in the first part of the experiment. The key novelty of the paper is two-fold. First, varying the complexity of the tasks allows for inducing exogenous shock to subjects' performance measurable on an individual level. Second, by complementing this approach with multiple elicitations of individuals' beliefs about their performance, I observe the impact of exogenous positive and negative shocks to self-image on an individual level as well. I investigate whether subjects who expect positive feedback are more likely to acquire information than those who expect negative feedback. I also test for reference-dependence of self-image concerns as well as for loss aversion in the self-image domain.

The experimental data provide strong evidence of information avoidance independently of whether the expected feedback is positive or negative. Individuals tend to have a stronger willingness to avoid feedback if they expect it to be negative. In line with expectations, subjects update their beliefs about their performance upward if they work on easier tasks in the first part of the experiment, which translates into an improvement in their self-image. Subjects who work on harder tasks first update their beliefs downward, indicating the dete-

[1] For a comprehensive multidisciplinary literature overview of information avoidance, see Golman et al., 2017.

rioration of their self-image. Moreover, subjects update their beliefs about their self-image only slightly after the easier task, while subjects who have done the harder task first update much stronger. At the end of the experiment, after individuals worked on both hard and easy tasks, their beliefs about their performance in the intelligence test go back to the pre-experiment levels. This result indicates that subjects did not find the overall complexity of the IQ test surprising.

I propose a stylized theoretical framework that offers a useful explanation of the experimental findings, in particular, the surprising patterns in belief updating. The framework captures the idea of disappointment aversion (Gul, 1991). It follows closely the setup of Gollier and Muermann, 2010 and models the trade-off between the ex-ante feelings and the risk of ex-post disappointment. In this framework, agents, who derive utility from self-image, first manage expectations and choose an optimal degree of optimism. Then, they decide whether they want to acquire self-image-relevant information. In the context of my experiment, subjects may choose to be optimistic about their performance and derive utility ex-ante at the cost of a possible disappointment at the end of the experiment. Alternatively, participants may stay pessimistic about their beliefs throughout the experiment and likely be positively surprised at the end.

My experimental setup uses intelligence as a self-image-relevant domain and lets subjects work on an IQ test.[2] To induce exogenous gains and losses in their self-image, I randomize whether subjects first complete a more difficult or easier part of a standard IQ test. This design feature allows inducing a sharp symmetric heterogeneous shock in subjects' performance in the first part that I observe on an individual level. After working on the easier part, subjects on average perceive their performance positively and thus expect good feedback. On the contrary, when initially working on the harder part perceived performance is on average worse, as is the feedback they expect. I employ a continuous willingness-to-pay (WTP) measure to elicit subjects' exact willingness to acquire feedback.

In order to be able to study perceived gains and losses in self-image, I elicit subjects' beliefs about their performance at three points over the course of the experiment. Prior belief elicitation takes place before subjects work on the IQ test. After inducing a gain or a loss in self-image by letting them work on easier or harder tasks, respectively, I elicit their beliefs again. The second belief elicitation allows observing whether they update their beliefs and whether they do so differently when they expect more positive and negative feedback. Furthermore, after the whole IQ test is complete and all subjects worked on the exactly same tasks, I elicit beliefs to analyze whether they recovered from the exogenous shock in self-image.

Multiple belief elicitations are an important feature of my design. While in standard economic theory the ultimate purpose of beliefs is to assist in the decision-making processes,

[2] Intelligence or IQ is a commonly used self-image-relevant domain. See, e.g., Fein and Spencer, 1997, Santos-Pinto and Sobel, 2005 and Castagnetti and Schmacker, 2020.

many recent studies have shown, both theoretically and experimentally, that individuals tend to hold motivated beliefs and argued that beliefs can be a choice variable (Bénabou and Tirole, 2002; Kőszegi, 2006). Experimental evidence shows that people dislike updating their beliefs negatively and react to noisy negative signals much weaker than to positive ones (Coutts, 2019; Eil and Rao, 2011; Golman et al., 2017; Zimmermann, 2020). In other words, a gain in self-image might be internalized stronger than a loss of the same magnitude. In contrast, individuals react stronger to losses than to gains in monetary and material domains (Kahneman, Knetsch, et al., 1990; Kahneman and Tversky, 1979) as well as with respect to health outcomes (Bleichrodt et al., 2001) and social image (Petrishcheva et al., 2022). It is important to observe not only actual differences in one's performance but also perceived ones. This paper focuses on disentangling these effects by looking at the willingness to acquire feedback of individuals who experience measurable perceived gains and losses in self-image. When analyzing subjects' willingness to acquire feedback, I take into consideration how they updated their beliefs.

This paper is organized as follows. Section 2 discusses the experimental design, implementation and technical details. In Section 3, I formulate the hypothesis. I present the results in Section 4. I discuss my results and propose a stylized theoretical framework in Section 5. Section 6 concludes.

## 2 Experimental design

My experimental setup includes three stages as shown in Figure 1. In Stage 1, I elicit subjects' prior beliefs about their performance in the upcoming IQ test. I treat prior beliefs as a within-subject reference point in intelligence, a self-image-relevant domain.[3] In Stage 2, I induce an exogenous shift in self-image. By introducing treatments *Loss* and *Gain*, I put subjects' self-image at either loss or gain by varying the task complexity. I then ask subjects whether they are willing to acquire feedback about their performance and elicit their willingness-to-pay to do so and their beliefs about their performance. The second belief elicitation is necessary to see whether the treatment variation worked, i.e., whether subjects indeed expect losses and gains when I assume they do. In Stage 3, I let subjects work on the remaining tasks of the IQ test and elicit their performance beliefs upon completion.

First, I analyze belief updating for those with gains and losses in self-image. I focus on the two main aspects: Is subjects' belief updating (a) going in the direction suggested by the treatment and (b) symmetric for gains and losses of self-image?

Additionally, this design allows for analyzing subjects' willingness to pay to acquire self-image-relevant feedback both unconditionally and conditionally on belief updating. Varying task complexity allows inducing an *objective* performance shift. Since subjects do not receive any signals about their performance except their *subjective* perception of it before they

[3] See, e.g., Fein and Spencer, 1997, Santos-Pinto and Sobel, 2005 and Castagnetti and Schmacker, 2020.

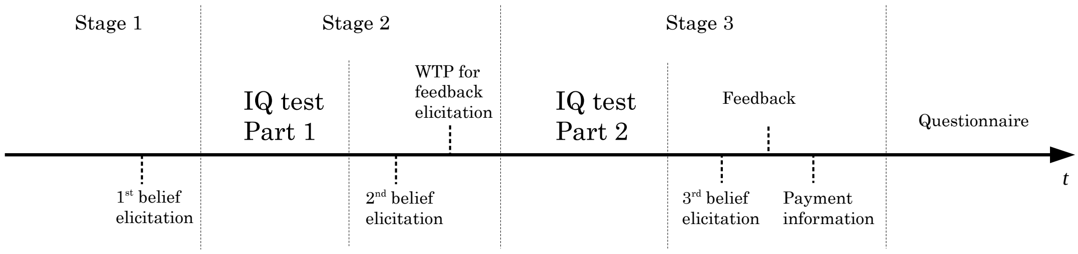



Figure 1: Timeline of the experiment

report the willingness to pay to acquire feedback, it is crucial to control for their beliefs when analyzing their WTP for feedback. I test whether subjects who care about their self-image avoid ego-relevant feedback. Then, I analyze whether those who experience a loss in self-image are more willing to acquire feedback than those who experience gain. I also test whether subjects with marginal self-image losses have a disproportionately higher willingness to acquire feedback than those with marginal gains in self-image.

**IQ test** In this experiment, subjects work on Raven's Progressive Matrices (RPMs; Raven, 1983), which are designed to measure fluid intelligence and often used in economic experiments to induce image concerns (e.g., Zimmermann, 2020, and Ewers and Zimmermann, 2015). In Figure 2, there are two examples of RPMs. They are picture puzzles with a missing piece. Among the available answers, subjects should choose the best logical fit to fill in the blank space. RPM tests commonly consist of five sets of matrices (A to E), with 12 matrices in each set. These sets progress in difficulty. Set A includes the easiest matrices; Set B is slightly harder, and so on. Set E contains the 12 hardest matrices. In Figure 2, the left matrix is one of the easier matrices from set B (B3), and the right one is among the most complicated tasks from set E (E10). Based on the reference data,[4] I expect student subjects to solve all the matrices in set A correctly. Hence, I do not use the 12 easiest matrices in this experiment but the 48 matrices from sets B to E.

I split 48 matrices into two parts: *Easy* and *Hard*. Matrices from sets B and C belong to the *Easy* part. Matrices from sets D and E form the *Hard* part. Both parts are progressive, i.e., they start with easy tasks and get more complicated over time. Matrices in parts *Easy* and *Hard* do not repeat or overlap. Subjects get one point if they solve a matrix correctly and get zero points otherwise. Subjects have a time limit of 30 seconds per matrix, which ensures that their performance is comparable within the experiment and to the reference sample, where the same time limit was imposed.

[4]The reference sample includes 413 observations (students) from a previous experiment that took place at the same lab in 2014 who worked on the full set (all 60) of the same RPM matrices.

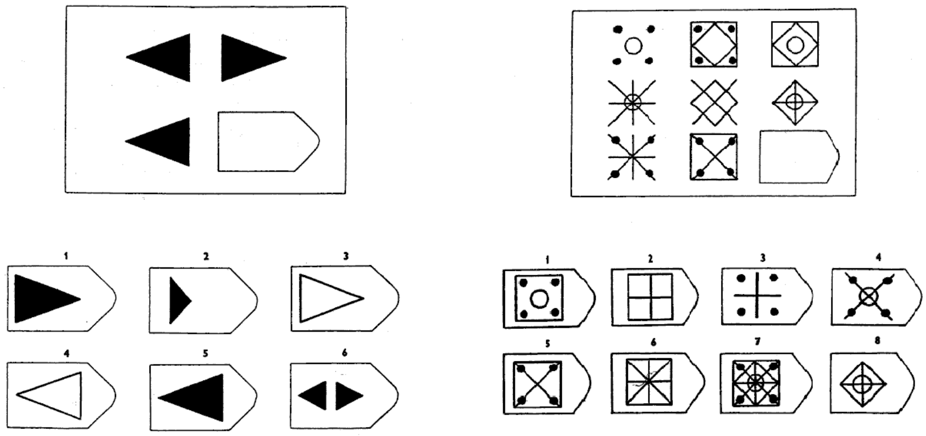



Figure 2: Examples of Raven's progressive matrices

**Stage 1** After reading general instructions and answering control questions,[5] subjects proceed to the first belief elicitation. I elicit their prior beliefs about their overall performance, i.e., the number of correctly solved matrices in both parts.

**Belief elicitation procedure** In the belief elicitation screen, subjects get the following information:[6]

- A summary of the performance of the reference sample. I tell subjects that in 2014, 413 individuals worked on the same picture puzzles in the DICE Lab. Additionally, I give them a short description of the data, namely: (a) no previous participant solved all 48 matrices, (b) the average participant solved 39 matrices, and (c) all previous participants solved at least 20 matrices or more.
- A figure with the performance of the reference sample. I show a histogram with scores displayed on the horizontal axis and the frequency (i.e., percent of the participants) on the vertical axis.
- A disclaimer saying "Carefully and honestly answering the question is in your best interest". Following Danz et al. (2020), I do not explain the exact monetary incentive structure in advance to reduce errors in belief elicitation. Instead, I tell them that the precise payment rule details are available by request at the end of the experiment.
- A slider with values between 0 and 48 and no default value where subjects should indicate how many matrices they think they will solve correctly.
- A phrase "I think I will solve X out of 48 picture puzzles correctly. It means that I think

[5]Original and translated versions of instructions and control questions (in German and English, respectively) are available in Appendix C.

[6]See a complete belief elicitation screen in Appendix C.

I will perform better than Y% of previous participants", which completes automatically when they choose or adjust the slider.

I incentivize the decision using the binarized scoring rule (Danz et al., 2020; Hossain and Okui, 2013). According to the binarized scoring rule, an individual may earn a fixed payment. The probability of receiving it increases the closer is her guess to the true outcome. In the context of my experimental design, participants can earn one euro in each belief elicitation task. Throughout the experiment, I used experimental currency units (ECU). The exchange rate was 1 euro = 20 ECU.[7] If their belief is correct, i.e., their perceived number of correctly solved matrices corresponds to their actual number of correctly solved matrices, they get a bonus of 20 ECU with a probability of one. Importantly, with the binarized scoring rule, subjects still have a small probability to get paid for the belief elicitation task, even if their guess and their actual performance differ a lot. Hence, their payoffs are not (directly) indicative of their performance.

Subjects' prior beliefs about their performance in the IQ test serve as a within-subject reference point in intelligence. The procedure of belief elicitations is always the same. I always ask subjects about their beliefs about their overall performance. Payoffs of multiple belief elicitations are independent.

**Stage 2** Subjects work on Part 1 of the test. In treatment *Gain*, Part 1 is *Easy*, such that subjects, on average, solve more matrices than they expected and hence can expect positive feedback about their performance. In treatment *Loss*, on the contrary, subjects work on *Hard* tasks, so they, on average, perform worse than expected. After participants completed 24 tasks in Part 1, I elicit their beliefs following the same procedure as described above.

After the second belief elicitation, I ask subjects about their willingness to pay to get feedback using the Becker-DeGroot-Marschak mechanism (Becker et al., 1964; BDM; see the screen in Appendix C). On a scale from -100 to 100 ECU (-5 to 5 euro), they report how much they would like to pay for feedback. Subjects are aware that WTP of -100 ECU guarantees that they will not receive information about their performance. WTP of 100 ECU means that they will certainly get feedback, and WTP of zero yields a 50% chance of receiving feedback about the number of matrices they solved correctly. I draw a random price for feedback from a uniform distribution with a support on the interval $[-100; 100]$. If the random price for feedback is smaller than or equal to the participants' WTP, they pay the price and receive feedback. If the random price for feedback is higher than their WTP, they do not pay the price and do not receive the feedback.

**Stage 3** Subjects work on the remaining 24 RPM tasks. It means that subjects from treatment *Gain* work now on the *Hard* part, while those from treatment *Loss* work on the *Easy* part. After Stage 3, all subjects have worked on the same 48 picture puzzles described

[7] In the instructions, I refer to ECU as thalers (Taler) which is a commonly used ECU in the DICE Lab.

above. Once subjects complete the task, I elicit beliefs about their performance again before they receive (or not) their feedback. I display their feedback in the same format as belief elicitation, i.e., it says "You solved X out of 48 picture puzzles correctly. This means that you performed better than Y% of previous participants".

**Questionnaire** After the main experiment is complete, subjects fill out a questionnaire. It contains the main sociodemographic characteristics such as age, gender, field of study, occupation, current GPA (or last degree GPA), high school GPA as well as average monthly budget and spending. Additionally, I ask them about their experience in the lab and collect independent measures of loss aversion in the monetary domain (Fehr and Goette, 2007; Gächter et al., 2007), risk aversion, and time preferences (Falk et al., 2016). Furthermore, subjects report the intensity of their social image concerns by answering the question "How important is the opinion that others hold about you to you?" following Ewers and Zimmermann (2015). I measure their overconfidence by letting them work on real-effort slider tasks and eliciting their beliefs about their performance (similar to S. Chen and Schildberg-Hörisch, 2019). Additionally, I elicit the intensity of self-image concerns following the approach of Aquino and Reed II (2002) and Grossman and Van Der Weele (2017). Subjects get a list of six statements about the importance of being kind, generous, and fair to their sense of self. They can choose whether they agree or disagree with those statements on a six-point Likert scale (from 0 indicating "strongly disagree" to 5 indicating "strongly agree"). Following Grossman and Van Der Weele (2017), I sum the points from evaluating all six statements to generate a measure of self-image concerns. The exact wording of each question is in Appendices C.7 and C.8. The independent measure of loss aversion in monetary domains is a set of incentivized lotteries. There are six lotteries and subjects can decide whether they accept or reject participation in each of them. One of the lotteries is paid out randomly at the end of the experiment. Each lottery yields a 50% chance of winning 12 ECU and a 50% chance of losing 4, 6, 8, 10, 12, or 14 ECU. Subjects do not earn any additional money if they rejected a lottery.

The independent measure of overconfidence is incentivized as well. There are 11 slider tasks, and subjects should position each slider in the middle (between 49 and 51 on a 0-100 scale). For each correctly solved slider task, subjects received 2 ECU. Furthermore, subjects could receive an additional 10 ECU if their guess about how many sliders they solved correctly was sufficiently accurate according to the binarized scoring rule (Hossain and Okui, 2013).

**Payment structure** Total earnings are only paid out upon completion of the experiment to prevent subjects from potentially dropping out. Subjects received a show-up fee of 3.70 euro as well as a 5 euro endowment at the beginning of the experiment, which might be used to pay for the feedback about their performance. The 5 euro endowment assures that, to ensure (not) getting feedback, the stakes are relatively high. However, subjects cannot make an absolute loss after their decision is realized. Additionally, subjects face three rounds of belief elicitations (before the experiment, after Part 1, and after Part 2) which pay 1 euro each with a probability that depends on the correctness of their belief. On top of that, loss aversion and overconfidence measures were monetarily incentivized.

**Technical details and procedure** This experiment was conducted online with subjects from the DICE Lab, University of Düsseldorf, in June 2021. For each session, all subjects took part in a web-conference call where they could ask clarifying questions or receive technical support if needed.[8] The experiment is preregistered on AEA RCT Registry[9], received an IRB approval[10] and was programmed using oTree (D. L. Chen et al., 2016). Subjects were recruited via Orsee (Greiner, 2015). Original instructions (in German) and the translated version of the instructions (in English) are in Appendix C. Subjects earned 13.3 euro on average for the experiment, which lasted approximately 45 minutes.[11] No subjects dropped out of the experiment. During the experiment, participants could not communicate with or see each other.

I conducted six online sessions with 20-24 participants each. In total, 132 subjects participated in the experiment: 67 of them were assigned to treatment *Gain* and the remaining 65 to treatment *Loss*. As reported in Table A1, the sample is well-balanced with respect to individual characteristics between treatments, such that no exclusion criteria apply.

# 3 Hypotheses

In this section, I formulate four pre-registered hypotheses regarding belief updating and information avoidance in my experiment.[12] First, I formulate a hypothesis about how subjects update their beliefs about their performance in the IQ test when I introduce positive and negative shocks to their performance.

**Hypothesis 1.** (Asymmetric belief updating)
*Individuals who care about their self-image may update their beliefs stronger if they experience a gain in a self-image-relevant domain compared to a loss in a self-image-relevant domain of the same size.*

In line with motivated beliefs literature (Coutts, 2019; Eil and Rao, 2011; Golman et al., 2017; Zimmermann, 2020), I hypothesize that the absolute difference between prior beliefs

[8] Li et al. (2020) find that using web-conference calls in online experiments leads to outcomes comparable to those the laboratory experiments for various economic games.

[9] Petrishcheva, Vasilisa. 2021. "Willful Ignorance and Reference-Dependence of Self-Image Concerns." AEA RCT Registry. June 09.

[10] IRB Approval No. 49nWIXIa

[11] Subjects earned at least 9.7 and at most 17.7 euro in this experiment. In addition to a show-up fee of 3.7 euro and an endowment of 5 euro, subjects' earnings depended on numerous decisions, namely, belief elicitations, willingness-to-pay for feedback, loss aversion lotteries, and performance in the overconfidence tasks. Subjects were not able to make an absolute loss in this experiment.

[12] The content of all hypotheses is the same as pre-registered. The order in which I present my hypothesis in the paper differs from the one in the pre-registration to mimic the order in which I present results in Section 4.

and the first posterior beliefs will be larger for subjects in *Gain* than in *Loss*. It implies that subjects who on average experience gains in their self-image update their beliefs stronger than those who experience losses of the same size in their self-image. In presence of a rather strong but very noisy signal about their performance (their perception of their performance), I expect subjects who observe a negative signal to be more hesitant to update their beliefs about their IQ compared to those who observed a positive signal. Next, I hypothesize that the share of subjects with a negative willingness to pay to acquire feedback relevant to their self-image in the IQ domain will be non-negligible.

**Hypothesis 2.** (Willful ignorance)

*Individuals who care about their self-image may avoid feedback relevant to their self-image.*

This experimental design induces changes in subjects' performance in an IQ test, a self-image-relevant domain. Acquiring or avoiding feedback may influence subjects' utility derived from their self-image. Hence, following the literature on information avoidance and willful ignorance (e.g., Golman et al., 2017; Kőszegi, 2006), I expect subjects may avoid information relevant to their self-image.

Next, I formulate the following hypothesis for *perceived* gains and losses of the same size:

**Hypothesis 3.** (Reference-dependence)

*On average, individuals who care about their self-image and expect a loss in their self-image are more willing to acquire self-image-relevant information than those who expect a gain in their self-image of the same size.*

I expect that individuals with a perceived loss will be more willing to acquire feedback about their performance than those with a perceived gain in self-image. The key novelty of this paper is analyzing the reference dependence of self-image concerns. More specifically, I test whether subjects who expect a loss in self-image have a higher willingness to pay to acquire feedback than those who expect a gain in self-image. If an individual expects a loss in self-image, positive feedback may serve as a tool to avoid this loss. Moreover, this paper focuses on individuals who experience marginal gains and losses. According to prospect theory (Kahneman and Tversky, 1979), there is a kink in the value function for changes in self-image which results in a kink in incentives to acquire self-image-relevant feedback. I hence formulate Hypothesis 4:

**Hypothesis 4.** (Loss aversion)

*Individuals who care about their self-image and expect a marginal loss in their self-image are more willing to acquire information than those who expect a marginal gain in their self-image.*

# 4 Results

This section is organized as follows. First, I discuss results related to subjects' performance in the IQ test in Parts 1 and 2 in Section 4.1. Next, I analyze subjects' beliefs about their intelligence in Section 4.2. In Section 4.3, I discuss their willingness-to-pay to receive self-image-relevant feedback.

## 4.1 IQ

Despite not being monetarily incentivized, subjects exerted substantial effort in solving the Raven's Progressive Matrices. On average, they solved 36.4 matrices correctly. Out of 48 matrices that subjects worked on, they gave at least 24 and at most 44 correct answers.[13]

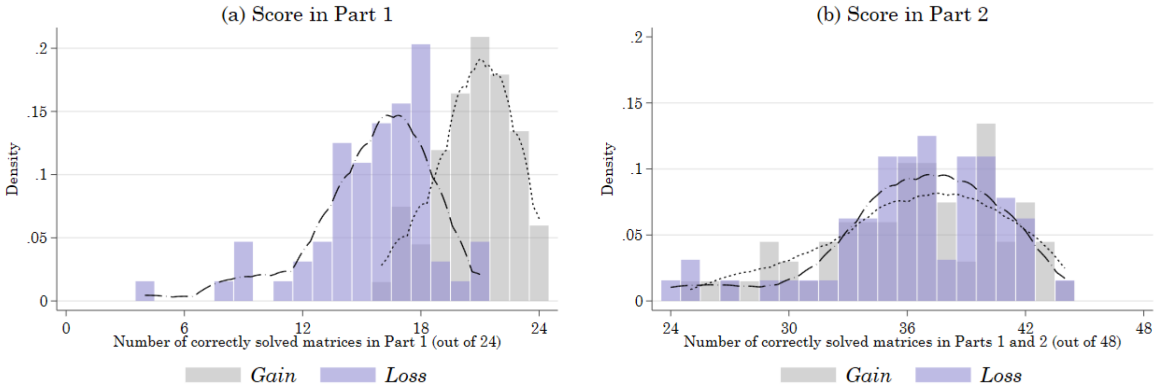



Figure 3: Score distributions by treatment

*Note*: Figures (a) and (b) display score distributions by treatment for Parts 1 and 2, respectively. In Figure (a), the horizontal axis shows the total number of correctly solved matrices after subjects completed Part 1. In Figure (b), the horizontal axis displays values between 24 and 48 because no subject solved less than 24 matrices correctly. In Figure (b), the horizontal axis shows the total number of correctly solved matrices after subjects completed Part 2. The vertical axis shows density. I show the histograms of score distributions and the kernel density estimates for treatments *Gain* and *Loss* in each figure. I estimate density using Epanechnikov kernels with an optimal bandwidth.

As intended by the experimental design, there are no significant differences in the distributions of the overall performance of subjects in treatments *Gain* and *Loss* (p=0.937).[14] I display the distributions of the score in Part 2 (overall performance) by treatment in Figure 3(b). The average number of correct answers is 36.3 and 36.4 in *Gain* and *Loss*, respectively. Working on part *Easy* first and on part *Hard* second (treatment *Gain*) leads to similar overall scores as working on part *Hard* first and on part *Easy* second (treatment *Loss*). Hence, there is no evidence for order effects in my experiment.

After subjects worked on Part 1 of the IQ test (the first 24 tasks), I document a substantial difference in performance between treatments *Gain* and *Loss*. The average number of correctly solved matrices is 20.7 in treatment *Gain* and 15.6 in treatment *Loss*. The difference in performance between treatments is highly statistically significant. In Figure 3(a), I illustrate the distributions of performance in Part 1 by treatment.

My experiment introduces a shock to subjects' self-image by affecting their score in

[13]Only one participant did not solve any matrices correctly by letting the 30-second timers run out. I exclude this subject from further analysis.

[14]In my analyses, I report two-sided Mann-Whitney U test results unless specified otherwise. I refer to results as (highly/weakly) statistically significant if the respective p-values are below 0.05 (0.01/0.1).

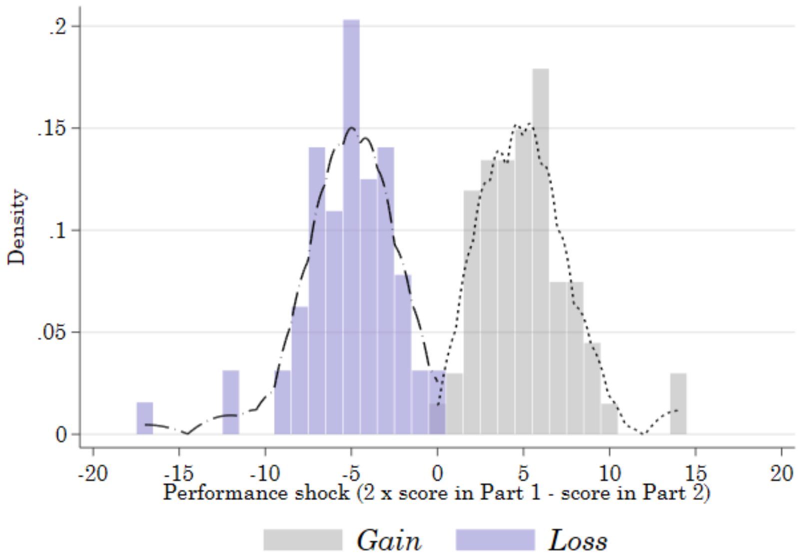



Figure 4: Performance shock (by treatment)

*Note*: This Figure displays histograms of the performance shock (2 × score in Part 1 - score in Part 2) by treatment. The dashed line corresponds to the density estimates with Epanechnikov kernels and an optimal bandwidth.

Part 1. I define a shock by comparing subjects' total number of correctly solved matrices (score in Part 2) and an extrapolated number of correctly solved matrices, i.e., the number of matrices they would have correctly solved if they carried on the same performance (2 × score in Part 1). The distributions of the performance shock are shown in Figure 4. The difference in shock distributions is highly statistically significant ($p<0.001$). Moreover, in absolute terms, performance shocks in *Gain* and *Loss* do not differ significantly ($p=0.904$) which indicates their symmetry for treatments *Gain* and *Loss*. Additionally, the performance shock I introduce in Part 1 aligns with the treatment assignment. As shown in Figure 4, there are no overlapping values of shock for treatments *Gain* and *Loss* except for zeros which account for two observations in treatment *Loss* and only one observation in treatment *Gain*. Hence, the score in Part 1 can act as a precise continuous individual-level measure of treatment that I will rely on in my analyses.

## 4.2 Beliefs about IQ

There are three belief elicitations in this experiment. I denote them Beliefs 1, 2, and 3, respectively. Belief 1 corresponds to the participants' prior beliefs about their performance which I elicit before they start working on the IQ test. Belief 2 is a subject's first posterior belief which I elicit in the middle of the IQ test, namely after they worked on the first 24 out of 48 matrices and after the treatment variation took place. Belief 3 is a second posterior belief. Its elicitation takes place after subjects worked on all 48 matrices. I present summary statistics of subjects' beliefs in Table 1 and distributions of beliefs in Figure 4.2.

I measure Belief 1 before the treatment variation affects the course of the experiment,

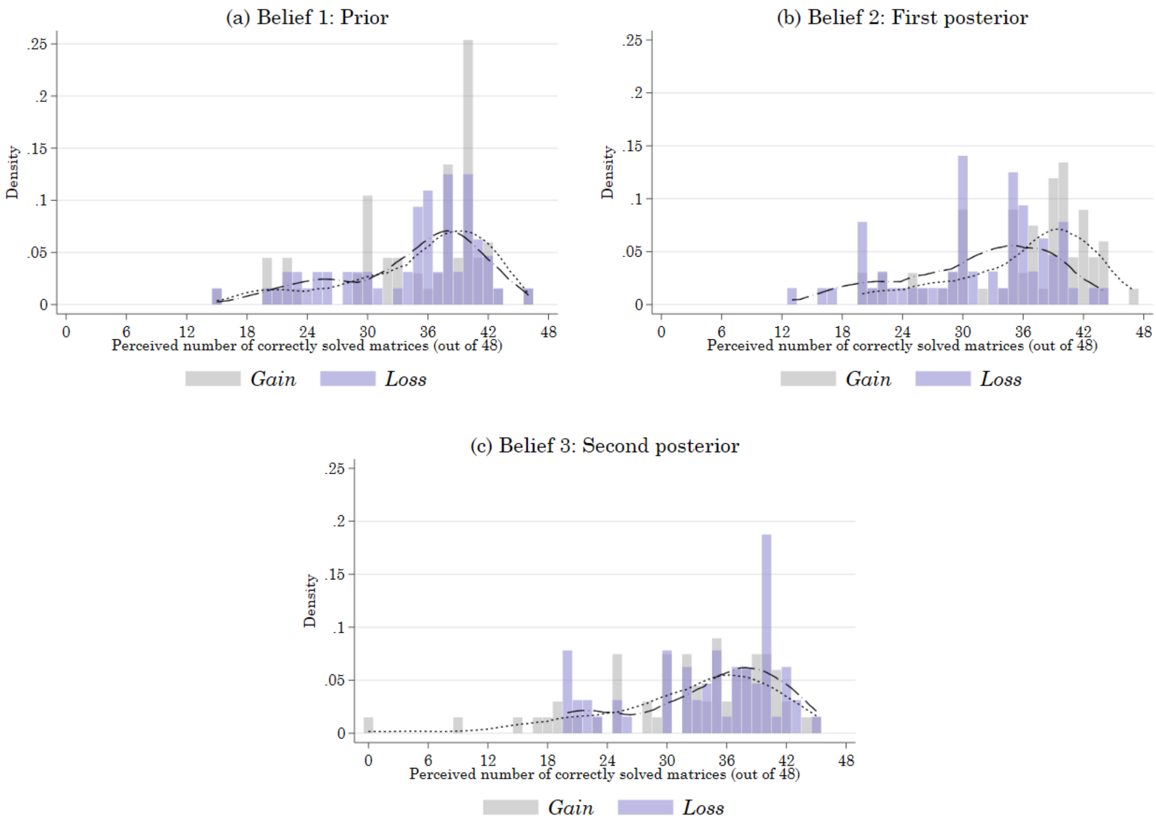



Figure 5: Belief distributions by treatment

*Note*: Figures (a)-(c) display distributions of Beliefs 1, 2, and 3 by treatment, respectively. The horizontal axis shows the total number of matrices that subjects expect to solve correctly (out of 48). The vertical axis shows density. I show the histograms of score distributions and the kernel density estimates for treatments *Gain* and *Loss* in each figure. I estimate density using Epanechnikov kernels with an optimal bandwidth.

Table 1: Summary statistics: Beliefs (by treatment)

| | *Loss* | *Gain* | Difference |
|---|---|---|---|
| Belief 1 | 34.30 | 35.03 | p=0.325 |
| Belief 2 | 31.58 | 36.07 | p<0.001 |
| Belief 3 | 34.05 | 32.27 | p=0.224 |
| N | 64 | 67 | 131 |

*Note*: I show mean values of Beliefs 1, 2 and 3 for treatments *Loss* and *Gain*. Beliefs 1, 2 and 3 indicate subjects' guesses about their number of correctly solved matrices (0 to 48). I compare distributions of Beliefs 1, 2 and 3 between treatments and report two-sided MWU test p-values.

hence creating a belief baseline for my analysis. Unsurprisingly, participants of treatments *Loss* and *Gain* do not differ in their prior beliefs about performance in the IQ test (p=0.325). On average, subjects believe they will solve 34.3 and 35.0 Raven's Progressive Matrices correctly in treatments *Loss* and *Gain*, respectively.

My treatment manipulation is designed to affect Belief 2. I shift participants' beliefs in the positive direction in treatment *Gain*, such that their Belief 2 is more optimistic than their Belief 1. In treatment *Loss*, on the contrary, participants update their beliefs negatively, i.e., Belief 2 is less optimistic than Belief 1. I find a highly significant difference in Belief 2 between subjects from *Gain* and *Loss* ($p<0.001$).

I define the belief difference as the difference between the first posterior beliefs about IQ and the prior beliefs about IQ: (Belief 2 - Belief 1). Hence, positive belief difference implies updating beliefs positively, i.e., subjects thinking they will solve more matrices than they initially assumed. Negative belief difference, on the contrary, refers to updating beliefs negatively. Subjects expect to solve fewer matrices correctly than they thought before.

Subjects' average belief difference is -2.72 in treatment *Loss* and 1.04 in treatment *Gain* ($p<0.001$). Thus, subjects (a) update beliefs about their performance in the IQ test according to their treatment assignment and (b) update their beliefs stronger if they experience a loss in the self-image domain. The latter result is statistically significant as well and provides evidence for asymmetric belief updating ($p=0.018$). Moreover, I find that these results hold on an individual level. Subjects in both treatments update their beliefs weaker than the performance shock they experience. Subjects' belief difference is 4.07 matrices lower in treatment Gain and 2.47 higher in treatment Loss than the performance shock they experience.

According to previous findings, individuals hold motivated beliefs, dislike updating their beliefs negatively, and react to noisy negative signals much weaker than to positive ones (Coutts, 2019; Eil and Rao, 2011; Golman et al., 2017; Zimmermann, 2020). In contrast to those findings, subjects in my experiment update stronger in absolute terms when facing a negative shock to their self-image than a positive one. This result is in line with the subjects' inclination to avoid possible disappointment at the end of the experiment. I discuss the mechanism which could lead to these patterns in belief updating in Section 5.

**Self-image or task complexity?** In my experiment, the aim is to shift the task complexity to induce a self-image shock. The crucial assumption is that subjects attribute this shock to their "wrong" prior belief about their performance, not to the difference in complexity between Part 1 and Part 2. This assumption is reasonable due to two factors: (a) it is common knowledge that the IQ test is of a representative difficulty, and (b) subjects observe the percentile of the reference sample each number of correct answers is associated with in all belief elicitation. Then, if subjects attribute their experience in Part 1 exclusively to the level of complexity being different from the one they expected, they would not update their beliefs about the overall performance at all. Alternatively, if subjects are unsure about the complexity of Part 1, it necessarily implies that they are uncertain about the overall complexity of the IQ test. However, it is not the case due to (a) and (b).

**Under-confidence about IQ** I compare subjects' prior beliefs about their IQ and their actual performance in the IQ test. Contrary to the consensus in economic and psychological literature,[15] I detect significant under-confidence using the Wilcoxon matched-pairs signed-ranks test (p=0.044). On average, participants believe they will solve 1.7 matrices less than they do. The degree of under-confidence in the IQ domain does not vary between treatments (p=0.721). Furthermore, subjects remain under-confident after they have completed the task, i.e., all 48 matrices. While the average performance results in 36.4 correct answers, the average Belief 3 is only 33.1 correct answers, and the difference is highly statistically significant (Wilcoxon matched-pairs signed-ranks test, $p<0.001$). The degree of under-confidence does not differ significantly between treatments (p=0.258). Crucially, this under-confidence is intelligence-specific. The survey measure of confidence, based on 11 real-effort slider tasks, shows that subjects are significantly overconfident (Wilcoxon matched-pairs signed-ranks test, $p<0.001$) and expect to solve 1.52 tasks more correctly than they do.

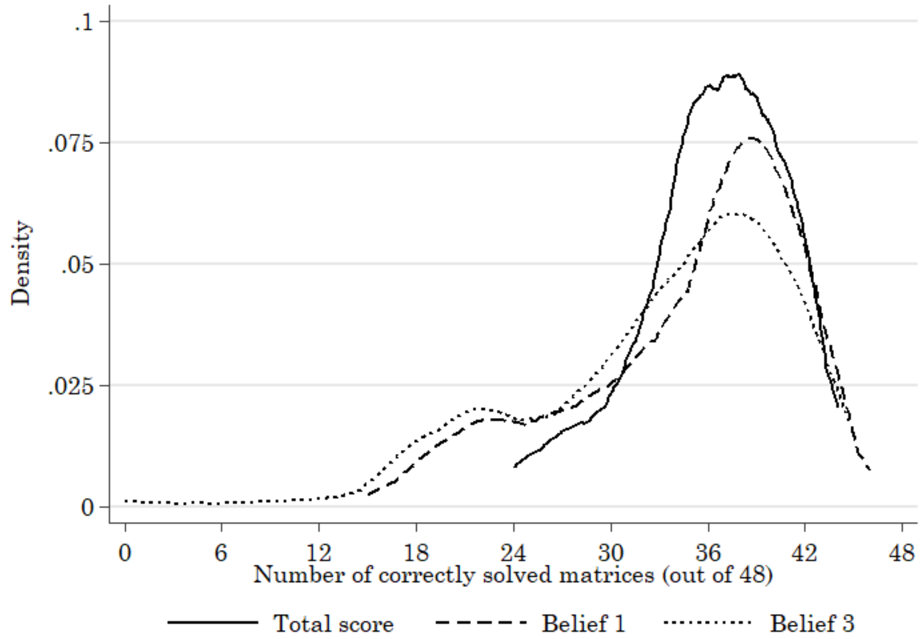



Figure 6: Performance in the IQ test and beliefs about it

*Note*: This Figure displays distributions of total scores (actual performance in the IQ test) as well as prior and second posterior beliefs about performance in the IQ test. Horizontal axis shows the total number of correctly solved matrices. Vertical axis shows the kernel density estimates using Epanechnikov kernels with an optimal bandwidth.

In Figure 6, I display kernel density estimates for subjects' total performance along with their prior and second posterior beliefs about it. Prior beliefs are unaffected by treatment assignment by design. Second posterior beliefs (Belief 3) are elicited at the end of the experiment, i.e., after subjects observed and worked on all matrices.[16] In Figure 6, I observe that belief distributions are more left-skewed than the distribution of total scores.

In belief elicitation instructions, I gave subjects an overview of the performance of the

[15]See, e.g., Burks et al. (2013) and Heck et al. (2018).

[16]I do not compare their total performance and Belief 2 because Belief 2 is directly affected by treatment, which leads to positive or negative belief shocks in treatments *Gain* and *Loss*, respectively.

reference sample, where, among other information,[17] I included the following statements: (a) no previous participant solved all 48 matrices, (b) the average participant solved 39 matrices, and (c) all previous participants solved at least 20 matrices or more. Subjects could see these statements in all belief elicitations. Interestingly, 1.5% and 5.3% of subjects still reported their perceived number of correct answers to be less than 20 in Beliefs 1 and 3, respectively. Moreover, 63.4% and 68.7% of them thought they would perform worse than an average participant of the reference sample (i.e., solve less than 39 matrices) in Beliefs 1 and 3, respectively.

In Result 1, I summarize the findings of belief updating.

**Result 1.** ***(Belief updating)***

*(a) Subjects have on average a negative belief difference in treatment Loss and a positive belief difference in treatment Gain.*

*(b) Subjects in treatment Loss update their beliefs stronger in absolute terms than subjects in treatment Gain.*

*(c) Subjects are on average under-confident in their beliefs about their performance. The degree of under-confidence does not differ between treatments.*

I find that subjects update their beliefs asymmetrically indicating that subjects hold motivated beliefs in the intelligence domain. This effect is, however, in the opposite direction as postulated in Hypothesis 1. Subjects in treatment *Loss* update their beliefs stronger than subjects in treatment *Gain*. While this result contrasts previous findings in the literature, I offer a simple possible explanation that is also in line with under-confidence and reference dependence in the intelligence domain in Section 5.

### 4.3 Willful ignorance

In this subsection, I analyze subjects' willingness to pay to acquire feedback. The WTP measure varies between -100 ECU and 100 ECU, where -100 means that an individual certainly wants no feedback, 100 implies that an individual definitely wants feedback, and 0 corresponds to a 50% chance of getting feedback.

On average, subjects reported a negative willingness to pay of -9.5 ECU. I show the distribution of the WTP in Figure 7. 42.0% of subjects reported a positive willingness-to-pay for feedback, implying they were ready to forego monetary benefits to increase their chances of acquiring feedback. However, only 2.3% of all participants had a WTP of 100. In total, 28.2% of subjects reported a negative willingness to pay to receive feedback. Moreover, 10.7% of all participants had a willingness-to-pay of -100 which guarantees no feedback about their performance in the IQ test.

**Result 2.** ***(Willful ignorance)***

[17] See detailed screenshots in Appendix C.

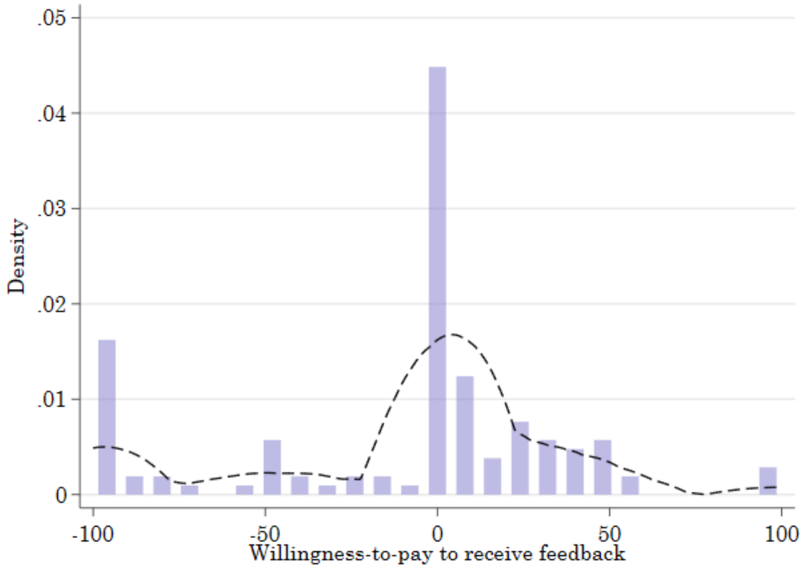



Figure 7: Willingness-to-pay for feedback

*Note*: This Figure displays a histogram of the willingness-to-pay to receive feedback. The dashed line corresponds to the density estimates with Epanechnikov kernels and an optimal bandwidth.

*(a) On average, subjects report a negative willingness to pay for self-image-relevant feedback.*

*(b) 28.2% of subjects report a negative willingness-to-pay to acquire feedback.*

*(c) 10.7% of subjects report a willingness-to-pay of -100 ECU that guarantees no feedback about their performance in the IQ test.*

In line with Hypothesis 2, a non-negligible share of participants has a negative WTP for feedback relevant to their self-image. Moreover, approximately one in ten participants choose to avoid feedback with certainty.

## 4.4 Reference dependence of self-image concerns

Participants were on average willing to pay -7.7 ECU in treatment *Gain* and -11.3 ECU in treatment *Loss* but the difference is not statistically significant (p=0.814). I report summary statistics of the WTP by treatment in Table 2. The shares of subjects who reported a positive, zero, and negative WTP is similar in treatments *Gain* and *Loss*. My sample size can detect a difference in willingness-to-pay of 9.3 ECU with 80% power and a significance level of 5%. With a scale from -100 to 100 ECU, the minimal detectable size of 9.3 ECU accounts for only 4.65% of the maximal shift and thus represents a minimal economically significant effect size.

To analyze reference dependence of self-image concerns, I account for how subjects update their beliefs when analyzing their WTP. I design treatments *Gain* and *Loss* to shift subjects' first posterior beliefs about their performance in the IQ test (Belief 2) by influencing their performance in Part 1. Hence, subjects endogenously update their beliefs taking into account their exogenous prior beliefs and an exogenous shock to their score in Part 1. In Table 3, I

Table 2: Summary statistics: Willingness-to-pay for feedback (by treatment)

| WTP | *Loss* | *Gain* | Difference |
|---|---|---|---|
| Negative | 0.297 | 0.269 | p=0.846 |
| Zero | 0.281 | 0.313 | p=0.707 |
| Positive | 0.422 | 0.418 | p=1.000 |
| N | 64 | 67 | 131 |

*Note*: This table shows shares of subjects whose reported WTP to receive feedback is negative, zero and positive for treatments *Loss* and *Gain*. I compare these shares between treatments and report two-sided Fisher's exact test p-values.

conduct 2SLS regressions to analyze the impact of beliefs on willingness to pay for feedback. I estimate the following regression:

$$\text{WTP}_i = \alpha + \beta(\text{Belief difference})_i + \gamma(\text{Belief 1})_i + \varepsilon_i,$$

which is equivalent to

$$\text{WTP}_i = \alpha + (\gamma - \beta)(\text{Belief 1})_i + \beta(\text{Belief 2})_i + \varepsilon_i.$$

I estimate the effect of belief difference on the willingness to pay for feedback on an individual level ($i$). The prior belief about the performance in the IQ test is exogenous. It is a proxy for the subjects' ability, and I elicit it before the treatment variation happens. The endogeneity concern arises with respect to Belief 2. Since I find evidence for motivated beliefs in my experimental data, I expect subjects to make decisions to update from Belief 1 to Belief 2 endogenously. Arguably, there might be unobservable individual effects that influence subjects' belief differences through Belief 2 and could be correlated with the error term. One likely candidate is the degree of optimism about the performance in Part 1 that can be potentially associated with belief updating and willingness to pay for feedback. Therefore, I instrument belief difference with the score in Part 1. Additionally, I include Belief 1 in both stages to account for differences in WTP for subjects with different levels of perceived ability.[18]

**Relevance** In the first stage, Belief 2 forms under the influence of two main criteria: previous beliefs about IQ and an exogenous shock introduced by the treatment. As I discussed in Section 4.1, the treatment shock affects subjects not only by treatment but also individually. Hence, using the score in Part 1 as an instrument provides me with greater precision on an individual level.

Belief difference is strongly correlated with the score in Part 1. The correlation coefficient

[18] My results are robust to excluding Belief 1 as presented in Table A2. Furthermore, I report simple OLS estimates with robust standard errors in Table A3. Both alternative specifications indicate robustness of the effect in Table 3.

is 0.41 and is highly statistically significant (p<0.001). This correlation emerges through the correlation between the score in Part 1 and Belief 2 (corr=0.47, p<0.001) but not between the score in Part 1 and Belief 1 (corr=0.13, p=0.134).

**Exogeneity** For the IV approach to be valid, the instrumental variable should be exogenous. In the discussed setup, the score in Part 1 should influence willingness to pay for feedback only through Belief 2.

The score in Part 1 is unobservable to subjects. They observe the complexity of the tasks in Part 1 and receive no additional signal about their performance. Hence, the only available information they have about the score in Part 1 *is* their perceived performance in Part 1. Since subjects' perceived performance in Part 1 is fully reflected in Belief 2, there is no other channel through which score in Part 1 influences WTP for feedback except Belief 2. I rely on the assumption of the maximal effort provision in the IQ test that, as discussed in Section 4.1, is consistent with the observed performance.

Table 3: Instrumental variable approach: Willingness-to-pay for feedback

(a) First stage

| Dependent variable: Belief 2 | |
|---|---|
| Score in Part 1 | 0.782*** |
| | (0.124) |
| Belief 1 | -0.384*** |
| | (0.078) |
| Constant | -1.734 |
| | (3.252) |
| N | 131 |
| F statistic | 39.72 |

(b) Second stage

| Dependent variable: WTP | |
|---|---|
| Belief 2 | 2.966** |
| | (1.375) |
| Belief 1 | 0.773 |
| | (0.676) |
| Constant | -33.942 |
| | (22.952) |
| N | 131 |

*Note*: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors are in parentheses. In these tables, I present the instrumental variable regression estimated via 2SLS. I report results of the first stage in Table (a) and results of the second stage in Table (b).

**Interpretation** On average, both the score in Part 1 and Belief 1 have a strong and highly significant impact on Belief 2. The F statistic of 39.72 indicates that the score in Part 1 is a strong instrument. According to the first-stage results presented in Table 3(a), one additional correctly solved matrix increases belief difference by approximately 0.8 matrices. The results are highly statistically significant.

The second stage shows the impact of the prior beliefs and the posterior beliefs on the willingness to pay to acquire self-image-relevant feedback. For a fixed prior, one standard deviation increase in posterior beliefs leads to a statistically significant increase in willingness to pay for feedback by 18.1 ECU. Prior beliefs about subjects' ability do not affect willingness-to-pay for self-image-relevant feedback significantly.

**Result 3.** ***(Reference dependence)***

*(a) A standard deviation increase in belief difference leads to a statistically significant increase in willingness to pay for feedback by on average 18.1 ECU.*

*(b) Prior beliefs about subjects' own ability do not affect willingness-to-pay for self-image-relevant feedback significantly.*

*(c) The difference in willingness-to-pay between participants of treatments Gain and Loss is not statistically significant (p=0.814).*

In line with Hypothesis 3, I find that participants' belief difference influences willingness-to-pay for feedback. However, contrary to Hypothesis 3, a higher belief difference leads to a higher willingness to pay. It indicates that participants who expect a gain in self-image are, on average, more willing to acquire information than those who, on average, expect a loss in their self-image. Indeed, subjects who expect bad news are more likely to avoid information than those who expect good news.

Importantly, this finding is belief-driven. A fixed performance shock introduced by the treatment assignment has no significant impact on subjects' average willingness to receive feedback. However, treatments affect subjects' beliefs about their performance in the IQ test asymmetrically. As discussed in Section 4.2, subjects update beliefs weaker when they expect a gain in self-image than when they expect a loss in self-image. Belief differences then lead to differences in willingness to pay on an individual level resulting in the higher WTP to receive feedback the larger the belief difference becomes.

## 4.5 Loss aversion in self-image concerns

To analyze whether loss aversion applies to self-image concerns, I focus on subjects with marginal perceived gains and losses. I apply a regression discontinuity design (RDD) to estimate local average treatment effects. Specifically, I use kink RDD. I aim at capturing the effect of small belief differences on willingness to pay to acquire feedback. Loss aversion implies a kink in incentives to receive feedback. Hence, I adjust my design to capture a kink, not a discontinuity. I present the results in Table 4.

Table 4 provides several specifications. Column (1) presents a kink RDD without additional control variables. Since belief difference is highly but not perfectly correlated with the treatment assignment, I include treatment as a control variable and present covariate-adjusted estimates in column (2). I further control for observable individual characteristics in column (3). These characteristics include age, gender, occupation, the field of study, monthly budget and spending, experience in laboratory experiments, number of correctly answered control questions, proxies for ability,[19] measures of risk aversion, time preferences, overconfidence, and intensity of social and self-image concerns.[20]

I find no evidence for loss aversion in self-image concerns. Subjects around the cut-off, i.e., those with belief differences close to zero, do not differ significantly in their willingness to

[19]Proxies for ability include current GPA, high school GPA, and IQ test results.

[20]See Section 2 and Table A1 for detailed explanation and summary statistics of all individual characteristics.

Table 4: Regression discontinuity design: Willingness-to-pay for feedback

| | (1) | (2) | (3) |
|---|---|---|---|
| RDD estimates | -0.949 | -1.894 | -4.381 |
| | (12.299) | (12.462) | (10.650) |
| Covariates | none | treatment | treatment and individual characteristics |
| N | 86 | 86 | 86 |

*Note*: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors are in parentheses. In this table, I present kink RDD local linear estimates with Epanechnikov kernels and a bandwidth of 5. In column (1), there are no additional control variables. In column (2), I control for treatment assignment. In column (3), I control for treatment assignment and individual characteristics. Individual characteristics include age, gender, occupation, field of study, monthly budget and spending, experience in laboratory experiments, number of correctly answered control questions, current GPA, high school GPA and IQ test results, measures of risk aversion, time preferences, overconfidence and intensity of social and self-image concerns. The reported number of observations indicates how many observations were actually used given a particular bandwidth selection criterion. Estimations are based on all 131 observations.

pay to acquire self-image-relevant feedback. This finding is robust for specifications presented in Table 4 and specifications with shorter and longer bandwidths reported in Table A4.

**Result 4.** ***(Loss aversion)***

*I find no significant difference in the effect of belief difference on willingness-to-pay for subjects with marginal gains and losses in self-image concerns.*

In this experiment, I find no evidence that supports Hypothesis 4. I observe no significant difference in the effect of belief difference on willingness-to-pay for feedback between subjects with small positive and negative belief differences.

# 5 Theoretical framework

In the following, I discuss a theoretical framework that offers an ex-post rationalization of my findings. The belief formation I observe in my experiment is in contrast to the pre-registered hypotheses. This mechanism is in line with the idea of disappointment aversion (Gul, 1991) and follows closely the setup of Gollier and Muermann (2010). Gollier and Muermann (2010) propose that the decision-maker faces a trade-off between the ex-ante feelings and the risk of ex-post disappointment and chooses an optimal degree of optimism. To capture the endogenous choice of beliefs, I consider a dual-self model following the setups of, e.g., Bénabou and Tirole (2002), Eil and Rao (2011), Fudenberg and Levine (2006), and Greiff (2019). In the context of my experiment, subjects may decide to be optimistic about their performance and derive utility ex-ante at the cost of a possible disappointment at the end of the experiment. Alternatively, participants may stay pessimistic about their beliefs throughout the experiment and likely be positively surprised at the end.

When reporting their Belief 2, subjects are aware that they only worked on Part 1 of the test, and there are 24 more matrices to solve. Arguably, subjects want to avoid a loss of self-image at the end of the experiment. Then, updating their beliefs weaker if participants are in gain can be optimal to avoid any possible disappointment at the end of the experiment. In other words, there might exist reference dependence not only within actions (willful ignorance) but also within the reported beliefs of the participants.

I observe that subjects who reported the WTP of -100 which guarantees that they do not receive feedback, were initially significantly more overconfident in the intelligence domain than others (p=0.024) and update their beliefs negatively (Belief 2 worse than Belief 1) in both treatments. The performance of these subjects in Parts 1 and 2 is not significantly different from other participants in the respective treatments (p=0.309 and p=0.287 in Parts 1 and 2, respectively). However, after they decide that they certainly do not want to receive feedback, their beliefs recover (Belief 3 better than Belief 2) in both treatments. Those findings might indicate that, at first, these subjects try to avoid disappointment in their performance by lowering the expectations, i.e., by adjusting their beliefs downwards. Yet, after they learn that they can avoid feedback altogether, they recover their beliefs accordingly.

I consider the following stylized framework to examine the mechanism of belief updating and incentives to acquire information that I observe in my experimental data.

**Setup** There are two stages denoted $t \in \{0, 1\}$. In $t = 0$, agents hold a prior belief about their type based on a self-image-relevant characteristic. In my experiment, this self-image-relevant characteristic is the number of correctly solved matrices in the IQ test. In $t = 1$, they face an exogenous shock to this characteristic, update their beliefs in response to the shock and choose whether to acquire or avoid information that affects their self-image.

I consider dual-self agents who derive reference-dependent utility from self-image. The concept of dual selves distinguishes the "rule chooser" and the "rule user", or a rational and an emotional self, for each agent (Bénabou and Tirole, 2002; Eil and Rao, 2011; Fudenberg and Levine, 2006; Greiff, 2019). The dual-self agent consists of two decision-makers: the rational self (R) and the emotional self (E). In $t = 1$, the emotional self shapes motivated beliefs, and the rational self takes beliefs as given and takes action which, in my setup, means it decides whether to acquire information.

An agent holds a prior belief about her type $n_0 \in [0, N]$ in period $t = 1$. She derives utility $\phi_0 = n_0$. In $t = 1$, agents experience an exogenous self-image shock $S \in \{s, -s\}$ with $s \in (0, 0.5)$ which can influence their perceived type either positively or negatively. Agents perceive this shock with a degree of sensitivity $\alpha \in (0, \bar{\alpha})$. Importantly, $\alpha > 0$ assures that they cannot fully neglect the shock. The agent forms two sets of beliefs in $t = 1$: I denote them $n_{1m}$ and $n_{1u}$ for motivated and unmotivated beliefs, respectively. The emotional self observes $n_{1u}$ and chooses an optimal $\alpha = \alpha^*$. The rational self observes $n_{1m}(\alpha^*)$.

The agent's motivated posterior beliefs are $n_{1m} = [1 + \alpha S]n_0$. Essentially, the agent's beliefs are influenced by a shock $S$, and the degree of optimism $\alpha$ determines the agent's sensitivity to this shock. Agents overestimate the shock when $\alpha > 1$ and underestimate the shock if $\alpha < 1$. I will refer to belief updating as "conservative" if $\alpha < 1$. For $\alpha = 1$, agents perceive the shock neutrally.

**Belief updating** In my experiment, subjects do not know about the possibility of acquiring or avoiding feedback before they arrive at the respective decision screen. Hence, when reporting Belief 2, I assume that their status quo is that they *will* receive feedback about their performance in the IQ test. It is plausible to assume that, without any additional indication, individuals who work on an IQ test would expect to receive results upon completing the test. Notably, if this assumption is reversed, the predictions of the model change drastically leading to outcomes summarized in hypotheses in Section 3.

The emotional self E endogenously chooses an optimal degree of sensitivity to the shock $\alpha^*$, while the rational self R takes it as given. E knows that R holds motivated beliefs $n_{1m}$ and that the agent will receive information about her ability and will have to update to $n_{1u}$ (unmotivated beliefs). Hence, her posterior beliefs will become $n_{1u}^{Gain} = (1+s)n$ if she is exposed to a positive shock, and $n_{1u}^{Loss} = (1-s)n$ if she is exposed to a negative one. Furthermore, E knows that R will experience losses whenever $n_{1m} < n_0$ or $n_{1u} < n_{1m}$, and gains otherwise. Therefore, E's objective is to choose $\alpha^*$ in the best interest of R. On the one hand, E wants to maximize the gain or minimize the loss when R updates her beliefs from $n_0$ to $n_{1m}$. On the other hand, E takes into account maximizing gains or minimizing losses from R updating from $n_{1m}$ to $n_{1u}$. First, the emotional self maximizes the following utility function with respect to $\alpha$:

$$\phi_{1E}|\text{information} = n_{1u} + v(n_{1u} - n_{1m}) + v(n_{1m} - n_0),$$

where $v(\cdot)$ is a value function for changes in self-image-relevant beliefs which satisfies the standard assumptions of prospect theory (Kahneman and Tversky, 1979). It has a kink at zero but is otherwise differentiable with $v(\Delta) < -v(-\Delta)$. In absolute terms, positive deviations from the reference point $n_0$ impact utility less than negative deviations of the same size ($v'(\Delta) < v'(-\Delta)$). I further assume that $v'(-\bar{\alpha}s) < v'|_0^+$, i.e., that the value function is sufficiently steeper for losses than for gains.[21] The value function is concave for gains ($v''(\Delta) < 0$ for $\Delta > 0$) and convex for losses ($v''(\Delta) > 0$ $\Delta < 0$).

**Proposition 1.** *Agents who experience a positive self-image shock optimally underestimate it with $\alpha^* = 1/2$.*

*Proof.* See Appendix B.1 for the proof. □

Essentially, the agent with a positive shock to her self-image will experience a gain when she updates from $n_0$ to $n_{1m}$. Depending on $\alpha$ the emotional self chooses, this gain may

[21] This assumption rules out cases where the agent who experiences a large loss might find it optimal to overestimate it to have a gain later on. This assumption assures that the marginal benefits from the smallest gain are lower than the marginal costs from the largest loss.

be relatively small if the agent's beliefs underestimate the shock and relatively large if her beliefs overestimate the shock. Additionally, she may experience a gain or a loss when she updates from $n_{1m}$ to $n_{1u}$. Her emotional self wants her to avoid this potential loss and hence keeps her motivated beliefs conservative ($\alpha < 1$) to avoid disappointment when updating from $n_{1m}$ to $n_{1u}$. Then, the trade-off she faces is how large the gain from $n_0$ to $n_{1m}$ and from $n_{1m}$ to $n_{1u}$ should be. Since the value function for self-image is concave for gains, it is better to experience two small gains than one large one. It means that the agent optimally updates exactly halfway from $n_0$ to $n_{1m}$ and then updates the remaining half from $n_{1m}$ to $n_{1u}$ resulting in optimal $\alpha^* = 1/2$.

Then, conditional on receiving information about their performance, agents' utility with the optimal degree of optimism is

$$\phi_1^{Gain}|\text{information}, \alpha^* = (1+s)n_0 + 2v(\frac{1}{2}sn_0).$$

Next, I examine how agents with a negative self-image shock update their beliefs. I summarize my findings in Proposition 2.

**Proposition 2.** *Agents who experience a negative self-image shock perceive it neutrally with* $\alpha^* = 1$.

*Proof.* See Appendix B.2 for the proof. □

The intuition behind this finding is as follows. The agent with a negative shock to her self-image will have a loss. She prefers to experience this loss all at once, rather than having two smaller losses because her value function in self-image is steeper for small losses and flatter for large ones. As $\alpha > 0$, she cannot fully neglect the loss when updating from $n_0$ to $n_{1m}$, so it is optimal to maximize it instead, such that she does not have to experience an additional loss when updating from $n_{1m}$ to $n_{1u}$. Therefore, she optimally perceives the shock neutrally, i.e., with $\alpha^* = 1$.

Then, conditional on receiving information about their performance, agents' utility with the optimal degree of optimism is

$$\phi_1^{Loss}|\text{information}, \alpha^* = (1-s)n_0 + v(-sn_0).$$

**Incentives to acquire information** The rational self R learns about the possibility of choosing whether to acquire or avoid information. She chooses to acquire information whenever her expected utility from acquiring information is higher than from avoiding it, conditional on an optimal degree of optimism. If the agent decides to acquire information, she has to forego the utility from her optimism $\alpha$ and perceive the performance shock objectively. If the agent acquires the information, her utility in is $\phi_1|\text{information}, \alpha^*$. If she chooses to avoid information, she holds her motivated beliefs $n_{1m}$ and has the following utility:

$$\phi_1|\text{no information}, \alpha^* = n_{1m} + v(n_{1m} - n_0).$$

Agents choose to acquire information whenever

$$\phi_1|\text{information}, \alpha^* > \phi_1|\text{no information}, \alpha^*. \tag{1}$$

I analyze the optimal information avoidance for agents in *Gain* and *Loss* separately because they optimally choose different $\alpha^*$. I summarize my findings on the incentives to acquire information for agents who experience a positive shock to their self-image in Proposition 3.

**Proposition 3.** *Agents who experience a positive self-image shock have a positive willingness-to-pay to acquire information conditional on their optimal degree of optimism ($\alpha^* = 1/2$).*

*Proof.* For agents with a positive shock to their self-image, Condition 1 reads

$$(1+s)n_0 + 2v(0.5sn_0) > (1+0.5s)n_0 + v\Big((1+0.5s)n_0 - n_0\Big).$$

Opening brackets and rearranging yields

$$0.5sn_0 + v(0.5sn_0) > 0.$$

The left-hand side of this condition is always positive. Hence, agents with a positive shock always prefer to acquire information. Their willingness-to-pay for information can be defined as a difference in their utilities with and without information, i.e., $WTP^{Gain} := 0.5sn_0 + v(0.5sn_0) > 0$. $\square$

Agents who are exposed to the positive self-image shock optimally underestimate the shock. Therefore, acquiring feedback improves their utility from self-image and yields a gain due to shifting beliefs from $n_{1m}$ to $n_{1u} \geq n_{1m}$.

I proceed to analyze the incentives to acquire information for agents with a negative shock to their self-image. I show my findings in Proposition 4.

**Proposition 4.** *Agents who are exposed to the negative self-image shock are indifferent between acquiring and avoiding information conditional on their optimal degree of optimism ($\alpha^* = 1$).*

*Proof.* For agents with a negative self-image shock, acquiring or avoiding information makes no difference as they have already fully internalized the shock by choosing to update their beliefs neutrally with $\alpha^* = 1$. It means that

$$\phi_1|\text{information}, \alpha^* = \phi_1|\text{no information}, \alpha^*.$$

Therefore, they are indifferent and their willingness-to-pay for information is $WTP^{Loss} := \phi_1|\text{information}, \alpha^* - \phi_1|\text{no information}, \alpha^* = 0$. $\square$

I analyze this mechanism in a stylized setup where dual-self agents optimally underestimate the positive shock in their self-image and optimally perceive a negative shock to their self-image neutrally. These patterns in belief updating are in line with my experimental data. I observed that subjects in treatment *Gain* update their beliefs by 4.07 matrices weaker than the positive performance shock they experience. It indicates that subjects in treatment *Gain* are indeed conservative in their belief updating. Conversely, subjects in treatment *Loss* are closer to neutral belief updating. They update their beliefs by only 2.47 matrices weaker than the negative performance shock they experience.

In the proposed mechanism, the agents who experience a positive shock choose to acquire information, and the agents who experience a negative shock are indifferent about whether to acquire or avoid it. This result is driven by the fact that the shock influences the optimal degree of optimism which in turn drives the updating process. My experimental data shows that an increase in the difference between the first posterior belief (Belief 2) and the prior belief (Belief 1) indeed leads to a higher willingness to pay for information.

# 6 Conclusion

This paper sheds light on the complexity and dynamic nature of self-image concerns. Individual perception of oneself is naturally belief-driven. Thus, understanding the motivation behind updating beliefs in this domain and the channels through which beliefs shape one's self-image is crucial for all decisions where self-image plays a role.

In this paper, I analyze individuals' willingness to avoid self-image-relevant information after I expose them to positive or negative shocks in their self-image. I complement this approach with multiple elicitations of beliefs about their self-image. They allow me to observe the impact of positive and negative shocks on an individual level. In my experiment, I induce an exogenous shift in self-image by introducing treatments *Loss* and *Gain*. Then, I ask subjects whether they are willing to acquire feedback about their performance and elicit their willingness-to-pay to do so and their beliefs about their performance.

As intended by the experimental design, individuals assigned to treatment *Gain* have a positive change in beliefs driven by a positive shock to their performance. Individuals in treatment *Loss*, on the contrary, update their beliefs negatively in line with a negative exogenous performance shock they experience. Interestingly, subjects in treatment *Loss* update their beliefs stronger than subjects in treatment *Gain*. Moreover, subjects in both treatments are, on average, under-confident and pessimistic in their beliefs about their intelligence. I propose a stylized theoretical framework to analyze the underlying mechanism. A possible explanation for this pessimism in beliefs is disappointment aversion (Gollier and Muermann, 2010; Kőszegi and Rabin, 2007).

On average, subjects report a negative willingness to pay for feedback relevant to their self-image. Almost one-third of participants reported a negative willingness-to-pay to acquire feedback. Moreover, about one in ten subjects had the lowest possible willingness-to-pay that guarantees no feedback about their performance on the IQ test.

I document causal evidence for reference dependence of self-image concerns. I find that an increasing change in beliefs about the performance in the IQ test leads to a statistically significant increase in willingness to pay for feedback. Furthermore, prior beliefs about subjects' abilities do not affect willingness-to-pay for self-image-relevant feedback significantly. Hence, the difference in willingness-to-pay between participants of treatments *Gain* and *Loss* being not statistically significant is driven by asymmetric belief updating. Moreover, I find no significant difference in the effect of belief difference on willingness-to-pay for subjects with marginal gains and losses in self-image concerns.

Generally, this paper studies the implications of self-image for the demand for relevant feedback and the evolution of their self-image itself. My findings may have broad implications in various domains like health, finance, labor, prosocial and altruistic behavior, etc. While avoiding information may maximize the short-term utility of an individual, it may yield severe welfare losses in the long run or negatively affect individuals themselves as well as those around them. For example, managers may avoid helpful feedback to maintain their self-image as a professional. It hinders them from becoming better managers and potentially affects the performance of their entire team. Curating effective feedback systems can therefore be vital for the well-being of the firms. Charitable donors who often prefer to remain uninformed about the actual effectiveness of donations experience a short-term warm glow from their actions. However, making a more informed choice could lead to more effective use of their resources. Another prominent and recent example comes from the rising necessity of lesson and lecture recordings. Many teachers may be reluctant to watch them back, despite the apparent benefits of improving their teaching style, to protect their ego. Detrimental effects of losses in self-image may be even more pronounced if individuals do not hold a strong prior in a particular domain. For example, new employees or students may be particularly vulnerable groups. Hence, the task allocators in the workplace and the designer of educational programs may regard the self-image effects and their possible consequences for feedback avoidance. Careful consideration of whether individuals experience gains or losses in self-image is crucial, as it can hinder individuals from acquiring relevant information.

My findings offer several avenues for future research. First, motivated belief updating relies strongly on the subjects' status quo in a given environment. Therefore, influencing subjects' perceptions of the status quo may shed more light on the formation of motivated beliefs. Furthermore, individuals tend to internalize negative feedback weaker than positive one (Zimmermann, 2020). Combining this finding with the evidence from my experimental data on stronger belief updating in the presence of a negative signal but *without feedback* may be insightful. Moreover, I focus on intelligence as a self-image-relevant domain in this paper.

However, many studies have previously documented that individuals derive self-image utility from a wide range of characteristics, e.g., beauty (Eil and Rao, 2011) or morality (Gneezy et al., 2020). Investigating whether individual behavior in case of gains and losses in other self-image-relevant domains follows similar patterns might be the next step towards a deeper understanding of how individuals perceive themselves.

## References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2019). "Preferences for Truth-Telling". *Econometrica* 87.4, 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.

Aquino, Karl and Americus Reed II (2002). "The self-importance of moral identity." *Journal of Personality and Social Psychology* 83.6, 1423.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak (1964). "Measuring utility by a single-response sequential method". *Behavioral Science* 9.3, 226–232.

Bénabou, Roland and Jean Tirole (2002). "Self-confidence and personal motivation". *The Quarterly Journal of Economics* 117.3, 871–915.

Bleichrodt, Han, Jose Luis Pinto, and Peter P Wakker (2001). "Making descriptive use of prospect theory to improve the prescriptive use of expected utility". *Management Science* 47.11, 1498–1514.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini (2013). "Overconfidence and social signalling". *The Review of Economic Studies* 80.3, 949–983.

Castagnetti, Alessandro and Renke Schmacker (2020). "Protecting the Ego: Motivated Information Selection and Updating".

Chen, Daniel L, Martin Schonger, and Chris Wickens (2016). "oTree—An open-source platform for laboratory, online, and field experiments". *Journal of Behavioral and Experimental Finance* 9, 88–97.

Chen, Si and Hannah Schildberg-Hörisch (2019). "Looking at the bright side: The motivational value of confidence". *European Economic Review* 120, 103302.

Coutts, Alexander (2019). "Good news and bad news are still news: Experimental evidence on belief updating". *Experimental Economics* 22.2, 369–395.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". *Economic Theory* 33.1, 67–80.

Danz, David, Lise Vesterlund, and Alistair J Wilson (2020). *Belief elicitation: Limiting truth telling with information on incentives.* Tech. rep. National Bureau of Economic Research.

Deshpande, Rohit and Ajay K Kohli (1989). “Knowledge disavowal: structural determinants of information-processing breakdown in organizations”. *Knowledge* 11.2, 155–169.

Eil, David and Justin M Rao (2011). “The good news-bad news effect: asymmetric processing of objective information about yourself”. *American Economic Journal: Microeconomics* 3.2, 114–38.

Ewers, Mara and Florian Zimmermann (2015). “Image and misreporting”. *Journal of the European Economic Association* 13.2, 363–380.

Falk, Armin, Anke Becker, Thomas J Dohmen, David Huffman, and Uwe Sunde (2016). “The preference survey module: A validated instrument for measuring risk, time, and social preferences”.

Fehr, Ernst and Lorenz Goette (2007). “Do workers work more if wages are high? Evidence from a randomized field experiment”. *American Economic Review* 97.1, 298–317.

Fein, Steven and Steven J Spencer (1997). “Prejudice as self-image maintenance: Affirming the self through derogating others.” *Journal of Personality and Social Psychology* 73.1, 31.

Fudenberg, Drew and David K Levine (2006). “A dual-self model of impulse control”. *American Economic Review* 96.5, 1449–1476.

Gächter, Simon, Eric J Johnson, and Andreas Herrmann (2007). “Individual-level loss aversion in riskless and risky choices”. *IZA Discussion Paper*.

Gneezy, Uri, Silvia Saccardo, Marta Serra-Garcia, and Roel van Veldhuizen (2020). “Bribing the self”. *Games and Economic Behavior* 120, 311–324.

Gollier, Christian and Alexander Muermann (2010). “Optimal choice and beliefs with ex ante savoring and ex post disappointment”. *Management Science* 56.8, 1272–1284.

Golman, Russell, David Hagmann, and George Loewenstein (2017). “Information avoidance”. *Journal of Economic Literature* 55.1, 96–135.

Greiff, Matthias (2019). “Team production and esteem: A dual selves model with belief-dependent preferences”. *Games* 10.3, 33.

Greiner, Ben (2015). “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE”. *Journal of the Economic Science Association* 1.1, 114–125.

Grossman, Zachary and Joel J Van Der Weele (2017). “Self-image and willful ignorance in social decisions”. *Journal of the European Economic Association* 15.1, 173–217.

Gul, Faruk (1991). “A theory of disappointment aversion”. *Econometrica: Journal of the Econometric Society*, 667–686.

Heck, Patrick R, Daniel J Simons, and Christopher F Chabris (2018). “65% of Americans believe they are above average in intelligence: Results of two nationally representative surveys”. *PloS one* 13.7, e0200103.

Hossain, Tanjim and Ryo Okui (2013). “The binarized scoring rule”. *Review of Economic Studies* 80.3, 984–1001.

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler (1990). “Experimental tests of the endowment effect and the Coase theorem”. *Journal of Political Economy* 98.6, 1325–1348.

Kahneman, Daniel and Amos Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk”. *Econometrica* 47.2, 263–292.

Karlsson, Niklas, George Loewenstein, and Duane Seppi (2009). “The ostrich effect: Selective attention to information”. *Journal of Risk and Uncertainty* 38.2, 95–115.

Kőszegi, Botond (2006). “Ego utility, overconfidence, and task choice”. *Journal of the European Economic Association* 4.4, 673–707.

Kőszegi, Botond and Matthew Rabin (2007). “Reference-dependent risk attitudes”. *American Economic Review* 97.4, 1047–1073.

Li, Jiawei, Stephen Leider, Damian R Beil, and Izak Duenyas (2020). “Running online experiments using web-conferencing software”. *Available at SSRN 3749551.*

Niehaus, Paul (2014). “A theory of good intentions”. *San Diego, CA: University of California and Cambridge, MA: NBER* 111.

Olafsson, Arna, Michaela Pagel, Olivier Coibion, and Yuriy Gorodnichenko (2018). “The ostrich in us: Selective attention to personal finances”. *NBER working paper* 23945.

Oster, Emily, Ira Shoulson, and E Dorsey (2013). “Optimal expectations and limited medical testing: evidence from Huntington disease”. *American Economic Review* 103.2, 804–30.

Petrishcheva, Vasilisa, Gerhard Riener, and Hannah Schildberg-Hörisch (2022). “Loss aversion in social image concerns”. *Experimental Economics (forthcoming).*

Raven, John C (1983). "Manual for Raven's progressive matrices and vocabulary scales". *Standard Progressive Matrices.*

Ropka, Mary E, Jennifer Wenzel, Elayne K Phillips, Mir Siadaty, and John T Philbrick (2006). "Uptake rates for breast cancer genetic testing: a systematic review". *Cancer Epidemiology and Prevention Biomarkers* 15.5, 840–855.

Santos-Pinto, Luis and Joel Sobel (2005). "A model of positive self-image in subjective assessments". *American Economic Review* 95.5, 1386–1402.

Schulz-Hardt, Stefan, Dieter Frey, Carsten Lüthgens, and Serge Moscovici (2000). "Biased information search in group decision making." *Journal of Personality and Social Psychology* 78.4, 655.

Serra-Garcia, Marta and Nora Szech (2022). "The (in) elasticity of moral ignorance". *Management Science* 68.7, 4815–4834.

Sweeny, Kate, Darya Melnyk, Wendi Miller, and James A Shepperd (2010). "Information avoidance: Who, what, when, and why". *Review of General Psychology* 14.4, 340–353.

Thornton, Rebecca L (2008). "The demand for, and impact of, learning HIV status". *American Economic Review* 98.5, 1829–63.

Zaltman, Gerald (1983). "Construing knowledge use". In: *Realizing Social Science Knowledge*. Springer, pp. 236–251.

Zimmermann, Florian (2020). "The dynamics of motivated beliefs". *American Economic Review* 110.2, 337–61.

# A Additional tables

Table A1: Differences in individual characteristics in treatments *Gain* and *Loss*

| Individual characteristics | Variable type | Min | Max | *Loss* | *Gain* | Difference |
|---|---|---|---|---|---|---|
| Age | Continuous | 18 | 49 | 25.844 | 24.985 | 0.362 |
| Gender: 1 if female | Binary | 0 | 1 | 0.656 | 0.522 | 0.156 |
| Gender: 1 if diverse | Binary | 0 | 1 | 0.000 | 0.015 | 1.000 |
| Occupation: 1 if student | Binary | 0 | 1 | 0.891 | 0.925 | 0.555 |
| Field of study: 1 if economics | Binary | 0 | 1 | 0.344 | 0.343 | 1.000 |
| Field of study: 1 if psychology | Binary | 0 | 1 | 0.016 | 0.030 | 1.000 |
| Lab experience | Continuous | 1 | 500 | 18.313 | 7.627 | 0.353 |
| Current GPA | Continuous | 1 | 4 | 2.207 | 2.230 | 0.969 |
| High school GPA | Continuous | 1 | 3.7 | 2.097 | 2.260 | 0.210 |
| Monthly budget | Continuous | 0 | 4000 | 532.359 | 505.299 | 0.434 |
| Monthly spending | Continuous | 0 | 1500 | 328.297 | 299.179 | 0.467 |
| Control questions (# correct) | Continuous | 1 | 3 | 2.781 | 2.896 | 0.155 |
| Risk aversion | Continuous | 1 | 10 | 4.969 | 5.448 | 0.309 |
| Overconfidence | Continuous | -6 | 10 | 2.063 | 1.000 | 0.066 |
| Time preferences | Continuous | 1 | 10 | 7.250 | 7.090 | 0.757 |
| Social image concerns | Continuous | 0 | 10 | 4.938 | 5.254 | 0.524 |
| Self-image concerns | Continuous | 0 | 60 | 38.000 | 38.612 | 0.967 |
| Loss aversion | Continuous | 0 | 6 | 3.531 | 3.493 | 0.861 |
| N | | | | 64 | 67 | 131 |

*Note*: I show summary statistics for subjects' individual characteristics in treatments *Loss* and *Gain*. I report the mean, minimal and maximal values of each variable. I also display p-values for treatment comparison for each corresponding variable. I compare the distributions of the variables marked "Continuous" using two-sided MWU tests. I compare the distributions of the variables marked "Binary" using two-sided Fisher's exact tests. Gender is a categorical variable (m/f/d). I test the differences between treatments by category. A detailed description of how I measure all individual characteristics is provided in Appendix C.7 and C.8 in English and German (original), respectively. Subjects' occupation was originally elicited as binary and indicates if an individual is a student. Field of study is a categorical variable and contains multiple fields, namely, mathematics or science, medicine, psychology, law or social sciences, economics, other and "I do not study". Following Abeler et al., 2019, I focus on economics and psychology. Lab experience indicates a self-reported number of economic experiments the subject has participated in. Please note that, despite the maximum of 500, 95% of subjects participated in 30 or fewer experiments. 79% of all subjects participated in 10 or fewer experiments. Current GPA and high school GPA reflect the standardized German grading system, with 1.0 corresponding to the best possible grade and 4.0 to the worst passing grade. Monthly budget and spending are measured in Euro, with fixed costs like rent already subtracted. Variable "Control questions" indicates the number of correctly answered control questions about the instructions of the current experiment (out of 3). Risk aversion, time preferences, and social image concerns are measured on an 11-point Likert scale (0-10). Larger reported values correspond to having a higher willingness to take risks, being more patient, and having stronger social image concerns, respectively. Overconfidence may vary between -11 and 11. Negative values of overconfidence correspond to under-confidence; Larger values imply stronger overconfidence. Self-image concerns is a measure that varies between -30 and 60 and indicates the intensity of self-image concerns, with larger values indicating stronger self-image concerns. Loss aversion may vary between zero and 6, and larger values mean stronger loss aversion.

Table A2: Instrumental variable approach robustness check: Willingness-to-pay for feedback

(a) First stage

| Dependent variable: Belief difference | |
|---|---|
| Score in Part 1 | 0.686*** |
| | (0.116) |
| Constant | -13.303*** |
| | (2.263) |
| N | 131 |
| F statistic | 34.73 |

(b) Second stage

| Dependent variable: WTP | |
|---|---|
| Belief difference | 3.245** |
| | (1.511) |
| Constant | -6.920 |
| | (4.272) |
| N | 131 |

*Note*: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors are in parentheses. In these tables, I present the instrumental variable regression estimated via 2SLS. I report results of the first stage in Table (a) and results of the second stage in Table (b).

Table A3: OLS robustness check: Willingness-to-pay for feedback

| Dependent variable: WTP | (1) | (2) |
|---|---|---|
| Belief 2 | 4.747** | 4.398* |
| | (2.342) | (2.443) |
| Belief 1 | 3.027 | 3.796 |
| | (2.471) | (2.559) |
| Belief 1 × Belief 2 | -0.120* | -0.133* |
| | (0.071) | (0.075) |
| Score in Part 1 | | 2.424** |
| | | (1.187) |
| Constant | -131.082* | -174.265** |
| | (74.358) | (80.150) |
| N | 131 | 131 |

*Note*: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors are in parentheses. In this table, I present the OLS regressions. Model (1) includes no further control variables. Model (2) includes Score in Part 1 which I use as an instrumental variable in Table 3.

# B Proofs

## B.1 Proof of Proposition 1.

For an agent exposed to a positive performance shock (*Gain*), the emotional self maximizes the following utility with respect to the degree of optimism $\alpha$:

$$\phi_{1E}^{Gain}|\text{information} = (1+s)n_0 + v\Big((1+s)n_0 - (1+\alpha s)n_0\Big) + v\Big((1+\alpha s)n_0 - n_0\Big). \quad (2)$$

Importantly, $\phi_{1E}^{Gain}|$information is non-differentiable at $(1+s)n - (1+\alpha s)n = 0$ or $\alpha = 1$. It is because at $\alpha = 1$, $v(\cdot)$ switches between gain and loss domains, thus creating a kink in

Table A4: Regression discontinuity design robustness check: Willingness-to-pay for feedback

| | Bandwidth = 4 | | | Bandwidth = 6 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| RDD estimates | 3.872 | 3.827 | 1.079 | 2.255 | 1.679 | 4.946 |
| | (16.048) | (16.300) | (14.447) | (7.220) | (7.199) | (6.771) |
| Covariates | none | treatment | treatment and individual characteristics | none | treatment and individual characteristics | treatment |
| N | 73 | 73 | 73 | 92 | 92 | 92 |

*Note*: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors are in parentheses. In this table, I present kink RDD local linear estimates with Epanechnikov kernels. In columns (1)-(3), I use a bandwidth of 4. In columns (4)-(6), I use a bandwidth of 6. In column (1), there are no additional control variables. In column (2), I control for treatment assignment. In column (3), I control for treatment assignment and individual characteristics. Individual characteristics include age, gender, occupation, field of study, monthly budget and spending, experience in laboratory experiments, number of correctly answered control questions, current GPA, high school GPA and IQ test results, measures of risk aversion, time preferences, overconfidence and intensity of social and self-image concerns. The reported number of observations indicates how many observations were actually used given a particular bandwidth selection criterion. Estimations are based on all 131 observations.

the utility function. Rearranging (2) yields

$$\phi_{1E}^{Gain}|\text{information} = (1+s)n_0 + v(\underbrace{(1-\alpha)sn_0}_{\substack{\geq 0 \text{ if } \alpha \leq 1 \text{ (Gain)} \\ < 0 \text{ if } \alpha > 1 \text{ (Loss)}}}) + v(\underbrace{\alpha s n_0}_{>0}).$$

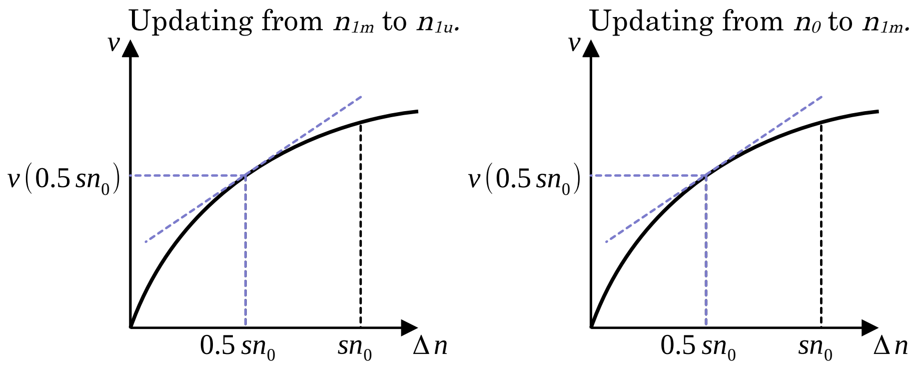
The value function is self-image is steeper for losses than for gains. Therefore, it is never optimal for the agent to choose $\alpha > 1$ that overestimates the gain because it will lead to a loss when the information is revealed. In a nutshell, the agent's decision boils down to a trade-off whether she wants to experience a gain while updating from $n_0$ to $n_1 m$ or from $n_1 m$ to $n_1 u$ as shown in Figure A1. The value function is steeper for small gains and flattens out as gains get larger. The agent prefers to go over the steepest part of the value function twice, hence, it is optimal to update exactly halfway from $n_0$ to $n_1 m$ and the remaining half from $n_1 m$ to $n_1 u$, such that $n_1 u - n_1 m = n_1 m - n_0$ or $\alpha = 1/2$. The optimal behavior is displayed in Panel A of Figure A1. In Panel B, I consider an alternative $\tilde{\alpha} \in (0, 1/2)$ and show that the slopes of the value function differ when updating from $n_0$ to $n_1 m$ or from $n_1 m$ to $n_1 u$. Hence, it would be optimal for the agent to update stronger from $n_0$ to $n_1 m$ and weaker from $n_1 m$ to $n_1 u$. The intuition is analogous for $\tilde{\alpha} \in (1/2, \bar{\alpha})$.

## B.2 Proof of Proposition 2

For an agent exposed to a negative performance shock (*Loss*), the emotional self maximizes the following utility with respect to the degree of optimism $\alpha$:

$$\phi_{1E}^{Loss}|\text{information} = (1-s)n_0 + v\Big((1-s)n_0 - (1-\alpha s)n_0\Big) + v\Big((1-\alpha s)n_0 - n_0\Big). \quad (3)$$

Panel A. $\alpha = -1/2$.




Panel B. $\alpha = \tilde{\alpha} \in (-1, -1/2)$.

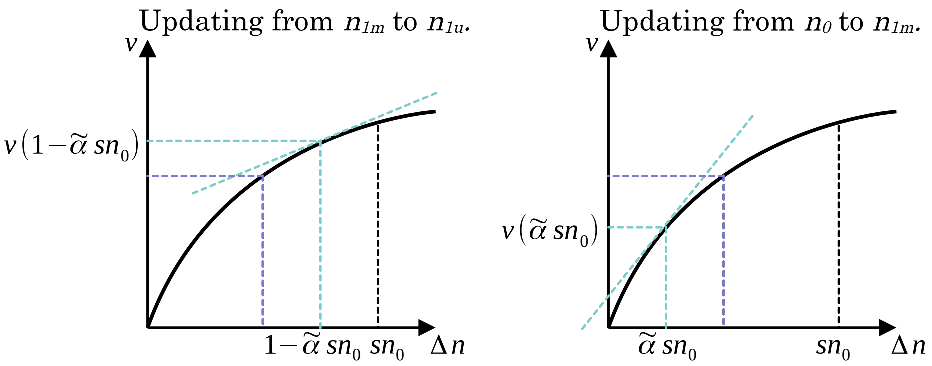



Figure A1: Graphical representation of the proof of Proposition 1.

$\phi_{1E}^{Loss}$|information is non-differentiable if $(1-s)n_0 - (1-\alpha s)n_0 = 0$ or $\alpha = 1$. Rearranging (3) yields

$$\phi_{1E}^{Loss}|\text{information} = (1-s)n_0 + v(\underbrace{-(1-\alpha)sn_0}_{\substack{\geq 0 \text{ if } \alpha \geq 1 \text{ (Gain)} \\ < 0 \text{ if } \alpha < 1 \text{ (Loss)}}}) + v(\underbrace{-\alpha sn_0}_{<0}).$$

The value function is steeper for small losses and flattens out as gains get larger. In contrast to the intuition in Section B.1, if the agent has a loss in self-image, step-wise belief updating ($\alpha = 1/2$) is the worst strategy because it maximizes the impact of losses, hence minimizing utility. Since $v'(-\bar{\alpha}s) < v'|_0^+$, i.e., the value function is sufficiently steeper for losses than for gains, choosing $\alpha > 1$ cannot be the agent's optimum either. The agent's decision is reduced to a trade-off when experiencing one large negative shock: while updating from $n_0$ to $n_1m$ or $n_1m$ to $n_1u$. As $\alpha > 0$, meaning the agent cannot completely neglect the shock when updating from $n_0$ to $n_1m$, it is optimal for her to experience a large loss then instead of when updating from $n_1 m$ to $n_1 u$. Formally,

$$\phi_{1E}^{Loss}|\alpha = \epsilon, \text{information} = (1-s)n_0 + v(-(1-\epsilon)sn_0) + v(-\epsilon sn_0)$$

and

$$\phi_{1E}^{Loss}|\alpha = 1, \text{information} = (1-s)n_0 + v(-sn_0),$$

with $\phi_{1E}^{Loss}|\alpha = \epsilon, \text{information} < \phi_{1E}^{Loss}|\alpha = 1, \text{information}$ for all $\epsilon \in (0,1)$.

# C Instructions of the Experiment

## C.1 General instructions: English

Please read the following instructions carefully! The amount of money you earn in this experiment strongly depends on your decisions. If you have any questions, please write a message to the experimenters in the chat. We will reply as soon as we can. During the experiment, it is not allowed to talk to other participants of the experiment or other people, use mobile phones or start other programs on the computer. Non-compliance with these rules will result in exclusion from the experiment and all payments. On the following pages we describe the exact procedure of the experiment.

In this experiment, we calculate your earnings using experimental currency units (talers). At the end of this experiment, all your earnings will be converted from talers to euro using the following exchange rate:

**1 taler = 5 cents**.

You will receive **a fixed payment of 74 talers** for participating in this experiment, which will be paid at the end of the experiment independent of your decisions in the experiment. Additionally, you receive **an endowment of 100 talers** which you might use in the course of the experiment. Please note that you receive your payments only upon completion of the entire experiment. In the following, there is a description of the exact experimental procedure.

### Overview of the Experiment

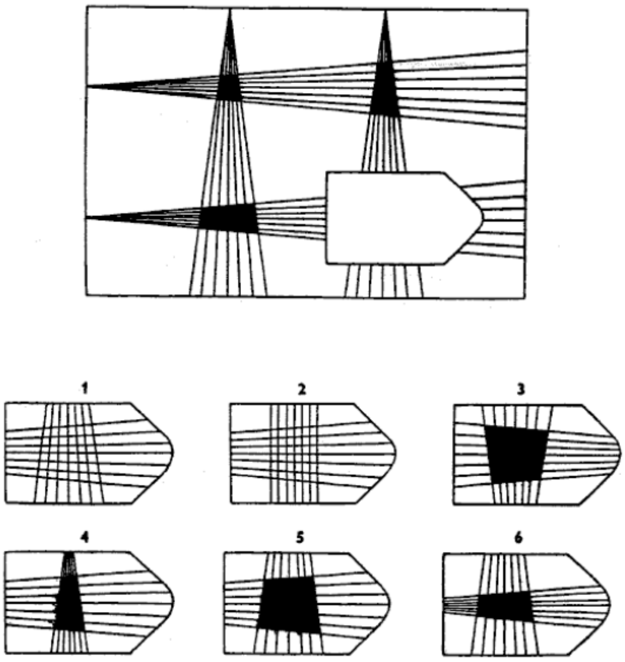
This experiment consists of **48 tasks** (24 tasks in Part 1 and 24 tasks in Part 2), which are often used **to measure so-called fluid intelligence of a person**. The fluid intelligence is an important part of the general intelligence of humans. These or similar tasks are also often used by companies in the context of recruitment procedures. **Each task corresponds to a picture puzzle.**

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.

**All participants** in the experiment **work on exactly the same 48 picture puzzles**

Example for a picture puzzle:




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. Only the order in which the picture puzzles are processed differs between group A and B, which has an influence on the relative complexity of the parts. The group membership has no further meaning. In Parts 1 and 2 you belong to the same group.

**Part 1 of the Experiment**

Before you start working on the picture puzzles, there will be some screens with questions. Then, you work on 24 picture puzzles following the rules described above (30 seconds time per puzzle, 1 point for correct answers, 0 points otherwise, etc.). After you have completed all 24 picture puzzles in Part 1, there will be some screens with questions before we proceed to Part 2.

**Part 2 of the Experiment**

Part 2 of the experiment is very similar to Part 1. 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.).

**End and Payment of the Experiment**

After Part 2 of today's experiment, there will be some more screens with information and questions before we proceed to the payment.

**If you have any questions now,**
**please write a message to the experimenters in the chat.**
**We will reply as soon as we can.**

## Control questions

1. According to which rule will your earnings be converted from the experimental currency units (talers) to euro? (correct answer - c)
   (a) 1 taler = 1 cent
   (b) 1 taler = 3 cents
   (c) 1 taler = 5 cents
   (d) 1 taler = 10 cents
2. How many tasks are you going to work on? (correct answer - c)
   (a) 24
   (b) 30
   (c) 48
   (d) 60
3. How much time do you have to work on each picture puzzle? (correct answer - b)
   (a) 15 seconds
   (b) 30 seconds
   (c) 45 seconds
   (d) 60 seconds

## C.2 General instructions: German (original)

Bitte lesen Sie die folgenden Instruktionen sorgfältig durch! Die Höhe Ihres Gewinns bei diesem Experiment hängt wesentlich von Ihren Entscheidungen ab. Wenn Sie Fragen haben, schreiben Sie bitte eine Nachricht an die ExperimentatorInnen im Chat. Wir werden so schnell wie möglich antworten. Während des Experiments ist es nicht erlaubt, mit anderen Teilnehmenden des Experiments oder anderen Personen zu sprechen, Handys zu benutzen oder andere Programme auf dem Computer zu starten. Die Nichteinhaltung dieser Regeln führt zum Ausschluss vom Experiment und sämtlicher Zahlungen. Auf den folgenden Seiten beschreiben wir den genauen Ablauf des Experiments.

In diesem Experiment berechnen wir Ihren Gewinn in Form von experimentellen Währungseinheiten (Taler). Am Ende des Experiments werden alle Ihre Gewinne unter Verwendung des folgenden Wechselkurses von Taler in Euro umgerechnet:

**1 Taler = 5 Cent**.

Sie erhalten **eine feste Zahlung von 74 Taler** für die Teilnahme an diesem Experiment, die am Ende des Experiments unabhängig von Ihren Entscheidungen im Experiment ausgezahlt wird. Zusätzlich erhalten Sie **eine Anfangsausstattung von 100 Taler**, die Sie im Laufe des Experiments verwenden können. Bitte beachten Sie, dass Sie Ihre Zahlungen erst nach Abschluss des gesamten Experiments erhalten. Im Folgenden finden Sie eine Beschreibung des genauen Versuchsablaufs.

### Überblick über das Experiment

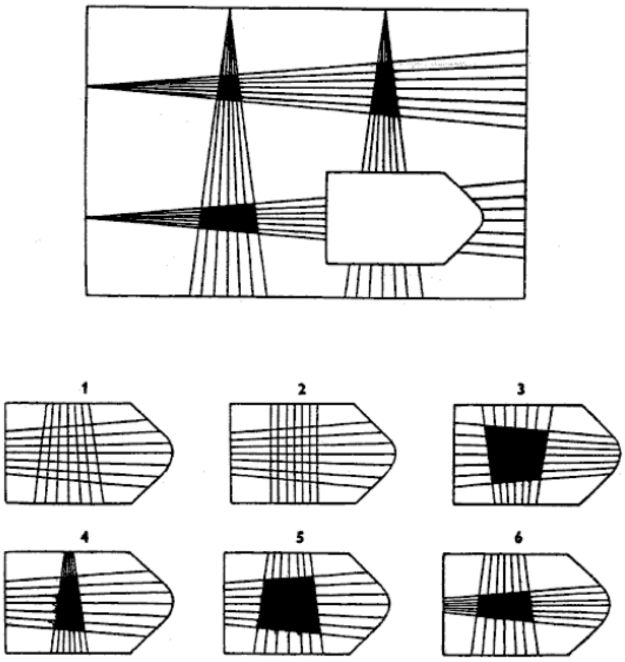
Dieses Experiment besteht aus **48 Aufgaben** (24 Aufgaben in Teil 1 und 24 Aufgaben in Teil 2), **die häufig zur Messung der** sogenannten **fluiden Intelligenz einer Person verwendet werden**. Die fluide Intelligenz ist ein wichtiger Teil der allgemeinen Intelligenz des Menschen. Diese oder ähnliche Aufgaben werden auch oft von Unternehmen im Rahmen von Einstellungsverfahren eingesetzt. **Jede Aufgabe entspricht einem Bilderrätsel.**

Jedes Bilderrätsel zeigt im oberen Teil ein Muster in einem Kasten, bei dem ein "Puzzleteil” in der unteren rechten Ecke ausgelassen ist. Ihre Aufgabe ist es, eines der unter dem Kasten aufgeführten Puzzleteile auszuwählen, das die leere untere rechte Ecke des Musters im Kasten logisch ausfüllt. **Bitte geben Sie die Nummer des Puzzleteils ein, das Ihrer Meinung nach am besten in den Rahmen passt.** Die Nummer eines Puzzleteils ist über jedem Puzzleteil angegeben. Es gibt immer genau ein Teil, das am besten passt.

Sie haben **30 Sekunden Zeit**, um die einzelnen Bilderrätsel zu lösen. Für jedes richtig ausgefüllte Bilderrätsel erhalten Sie einen Punkt. **Wie bei Intelligenztests üblich, werden richtige Antworten nicht zusätzlich vergütet.** Sie erhalten 0 Punkte für jedes falsch beantwortete Bilderrätsel oder wenn Sie nicht innerhalb von 30 Sekunden das am besten passende Teil des Rätsels auswählen.

**Alle Teilnehmenden** des Experiments **arbeiten an genau den gleichen 48 Bilder-**

Beispiel für ein Bilderrätsel:




**rätseln**, die oben beschrieben wurden. Die Teilnehmenden werden zufällig einer von zwei Gruppen zugewiesen: Gruppe A oder Gruppe B. Während des gesamten Experiments lösen alle Teilnehmenden beider Gruppen genau die gleichen 48 Bilderrätsel, 24 in Teil 1 und 24 in Teil 2. Nur die Reihenfolge, in der die Bilderrätsel bearbeitet werden, unterscheidet sich zwischen Gruppe A und B, was einen Einfluss auf die relative Komplexität der Teile hat. Die Gruppenzugehörigkeit hat keine weitere Bedeutung. Sie gehören in Teil 1 und 2 der gleichen Gruppe an.

**Teil 1 des Experiments**

Bevor Sie mit der Bearbeitung der Bilderrätsel beginnen, werden mehrere Seiten mit Fragen angezeigt. Dann bearbeiten Sie 24 Bilderrätsel nach den oben beschriebenen Regeln (30 Sekunden Zeit pro Rätsel, 1 Punkt für richtige Antworten, ansonsten 0 Punkte, usw.). Nachdem Sie alle 24 Bilderrätsel in Teil 1 gelöst haben, werden erneut ein paar Seiten mit Fragen gezeigt, bevor wir zu Teil 2 übergehen.

**Teil 2 des Experiments**

Teil 2 des Experiments ist sehr ähnlich zu Teil 1. Sie bearbeiten 24 weitere Bilderrätsel nach den gleichen Regeln (30 Sekunden Zeit pro Rätsel, 1 Punkt für richtige Antworten, ansonsten 0 Punkte, usw.).

**Ende und Bezahlung des Experiments**

Nach Teil 2 des heutigen Experiments werden noch einige Seiten mit Informationen und Fragen angezeigt, bevor wir zur Bezahlung übergehen.

**Wenn Sie jetzt noch Fragen haben,**
**schreiben Sie bitte eine Nachricht an die ExperimentatorInnen im Chat.**

**Wir werden so schnell wie möglich antworten.**

## Kontrollfragen

1. Nach welcher Regel wird Ihr Gewinn von der experimentellen Währungseinheit (Taler) in Euro umgerechnet? (richtige Antwort - c)
   (a) 1 Taler = 1 Cent
   (b) 1 Taler = 3 Cent
   (c) 1 Taler = 5 Cent
   (d) 1 Taler = 10 Cent
2. Wie viele Aufgaben werden Sie bearbeiten? (richtige Antwort - c)
   (a) 24
   (b) 30
   (c) 48
   (d) 60
3. Wie viel Zeit haben Sie für die Bearbeitung der einzelnen Bilderrätsel? (richtige Antwort - b)
   (a) 15 Sekunden
   (b) 30 Sekunden
   (c) 45 Sekunden
   (d) 60 Sekunden

## C.3 Belief elicitations: English

Please answer the question below.

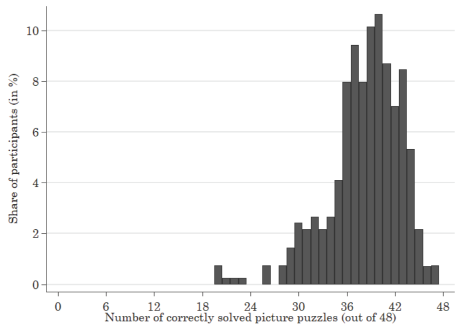
## How many picture puzzles (out of 48) do you think you will solve correctly?

**You have a chance of winning 20 thalers. Truthful reporting will maximize your chances of winning.**

In 2014, 413 people worked on exactly the same 48 picture puzzles in the DICE Lab. In the figure below, you can see how many participants gave how many correct answers.

**To summarize:**

- **None** of previous participants **solved all 48** picture puzzles.
- **An average** participant **solved 39** picture puzzles.
- **All** previous participants **solved at least 20** picture puzzles or more.




**Please note: Carefully and honestly answering the question is in your best interest. An honest answer increases the probability of earning the bonus of 20 thalers.**

The precise payment rule details are available by request at the end of the experiment. Please indicate your answer by clicking on the slider. You can adjust the position afterwards.

**How many picture puzzles (out of 48) do you think you will solve correctly?**

**No picture puzzle (0)** | **All picture puzzles (48)**

**Please click on the slider to indicate your assessment. You can still change your decision afterwards.**

Continue

Figure A2: Belief elicitation screen

Please answer the question below.

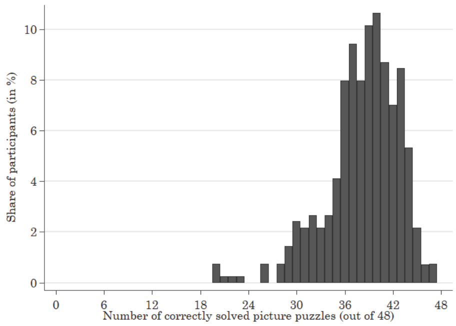
## How many picture puzzles (out of 48) do you think you will solve correctly?

**You have a chance of winning 20 thalers. Truthful reporting will maximize your chances of winning.**

In 2014, 413 people worked on exactly the same 48 picture puzzles in the DICE Lab. In the figure below, you can see how many participants gave how many correct answers.

**To summarize:**

- **None** of previous participants **solved all 48** picture puzzles.
- **An average** participant **solved 39** picture puzzles.
- **All** previous participants **solved at least 20** picture puzzles or more.




**Please note: Carefully and honestly answering the question is in your best interest. An honest answer increases the probability of earning the bonus of 20 thalers.**

The precise payment rule details are available by request at the end of the experiment. Please indicate your answer by clicking on the slider. You can adjust the position afterwards.

**How many picture puzzles (out of 48) do you think you will solve correctly?**

**No picture puzzle (0)** — **All picture puzzles (48)**

**I think, I will solve 38 out of 48 picture puzzles correctly. This means that I think I will perform better than 38.01% of previous participants.**

Continue

Figure A3: Belief elicitation screen (answered)

## C.4 Belief elicitations: German (original)

Bitte beantworten Sie die untenstehende Frage.

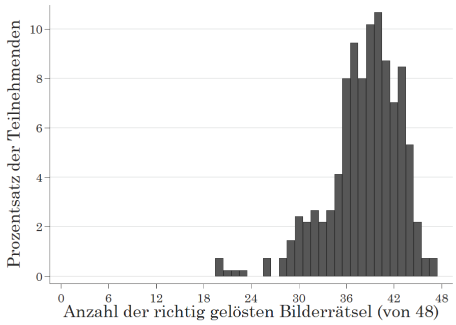
## Wie viele Bilderrätsel von insgesamt 48 denken Sie, werden Sie korrekt lösen?

**Sie haben die Chance 20 Taler zu gewinnen. Eine wahrheitsgemäße Angabe maximiert Ihre Gewinnchancen.**

Im Jahr 2014 arbeiteten 413 Personen an genau denselben 48 Bilderrätseln im DICE Lab. In der Abbildung unten können Sie sehen, wie viele Teilnehmende wie viele richtige Antworten abgegeben haben.

Zusammenfassend

- **Keine(r)** der damaligen Teilnehmenden **löste alle 48** Bilderrätsel.
- **Der/die durchschnittliche** Teilnehmende **löste 39** Bilderrätsel.
- **Alle** damaligen Teilnehmenden **lösten mindestens 20** Bilderrätsel oder mehr.




**Bitte beachten Sie: Eine sorgfältige und ehrliche Beantwortung der Frage ist in Ihrem besten Interesse. Eine ehrliche Antwort erhöht die Wahrscheinlichkeit, dass Sie den Bonus von 20 Taler verdienen.**

Die genauen Details der Zahlungsregelung sind am Ende des Experiments auf Anfrage einsehbar. Bitte geben Sie Ihre Antwort an, indem Sie auf den Schieberegler klicken. Sie können die Position anschließend anpassen.

**Wie viele Bilderrätsel von insgesamt 48 denken Sie, werden Sie korrekt lösen?**

**Gar kein Bilderrätsel (0)** — **Alle Bilderrätsel (48)**

**Bitte klicken Sie auf den Schieberegler, um Ihre Einschätzung anzugeben. Sie können Ihre Entscheidung im Anschluss noch verändern.**

Weiter

Figure A4: Belief elicitation screen

Bitte beantworten Sie die untenstehende Frage.

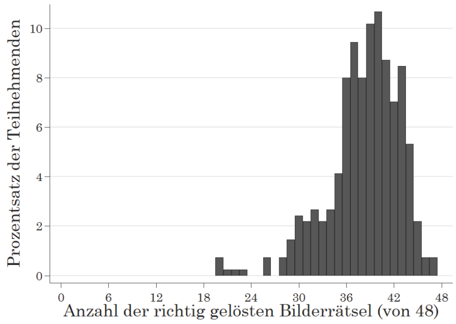
## Wie viele Bilderrätsel von insgesamt 48 denken Sie, werden Sie korrekt lösen?

### Sie haben die Chance 20 Taler zu gewinnen. Eine wahrheitsgemäße Angabe maximiert Ihre Gewinnchancen.

Im Jahr 2014 arbeiteten 413 Personen an genau denselben 48 Bilderrätseln im DICE Lab. In der Abbildung unten können Sie sehen, wie viele Teilnehmende wie viele richtige Antworten abgegeben haben.

Zusammenfassend

- **Keine(r)** der damaligen Teilnehmenden **löste alle 48** Bilderrätsel.
- **Der/die durchschnittliche** Teilnehmende **löste 39** Bilderrätsel.
- **Alle** damaligen Teilnehmenden **lösten mindestens 20** Bilderrätsel oder mehr.




**Bitte beachten Sie: Eine sorgfältige und ehrliche Beantwortung der Frage ist in Ihrem besten Interesse. Eine ehrliche Antwort erhöht die Wahrscheinlichkeit, dass Sie den Bonus von 20 Taler verdienen.**

Die genauen Details der Zahlungsregelung sind am Ende des Experiments auf Anfrage einsehbar. Bitte geben Sie Ihre Antwort an, indem Sie auf den Schieberegler klicken. Sie können die Position anschließend anpassen.

**Wie viele Bilderrätsel von insgesamt 48 denken Sie, werden Sie korrekt lösen?**

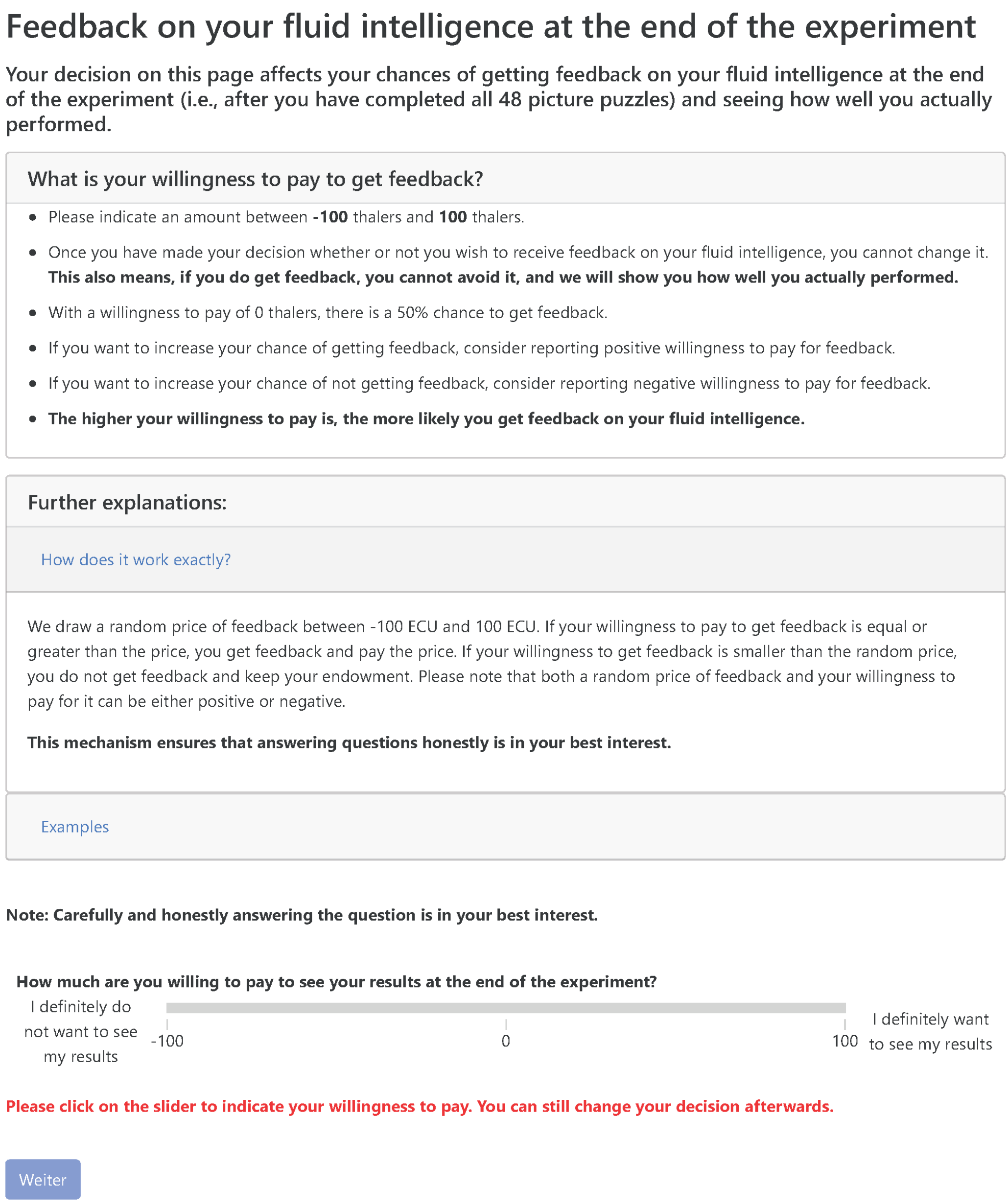
**Gar kein Bilderrätsel (0)** **Alle Bilderrätsel (48)**

**Ich denke, ich werde 38 von 48 Bilderrätseln richtig lösen. Das bedeutet, ich denke, dass ich besser als 38.01% der früheren Teilnehmenden abschneiden werde.**

Weiter

Figure A5: Belief elicitation screen (answered)

## C.5 Willingness-to-pay for feedback: English



# Feedback on your fluid intelligence at the end of the experiment

**Your decision on this page affects your chances of getting feedback on your fluid intelligence at the end of the experiment (i.e., after you have completed all 48 picture puzzles) and seeing how well you actually performed.**

### What is your willingness to pay to get feedback?

- Please indicate an amount between **-100** thalers and **100** thalers.
- Once you have made your decision whether or not you wish to receive feedback on your fluid intelligence, you cannot change it. **This also means, if you do get feedback, you cannot avoid it, and we will show you how well you actually performed.**
- With a willingness to pay of 0 thalers, there is a 50% chance to get feedback.
- If you want to increase your chance of getting feedback, consider reporting positive willingness to pay for feedback.
- If you want to increase your chance of not getting feedback, consider reporting negative willingness to pay for feedback.
- **The higher your willingness to pay is, the more likely you get feedback on your fluid intelligence.**

### Further explanations:

How does it work exactly?

We draw a random price of feedback between -100 ECU and 100 ECU. If your willingness to pay to get feedback is equal or greater than the price, you get feedback and pay the price. If your willingness to get feedback is smaller than the random price, you do not get feedback and keep your endowment. Please note that both a random price of feedback and your willingness to pay for it can be either positive or negative.

**This mechanism ensures that answering questions honestly is in your best interest.**

Examples

**Note: Carefully and honestly answering the question is in your best interest.**

**How much are you willing to pay to see your results at the end of the experiment?**

I definitely do not want to see my results -100 0 100 I definitely want to see my results

**Please click on the slider to indicate your willingness to pay. You can still change your decision afterwards.**

Weiter

Figure A6: Willingness-to-pay for feedback screen

## Feedback on your fluid intelligence at the end of the experiment

**Your decision on this page affects your chances of getting feedback on your fluid intelligence at the end of the experiment (i.e., after you have completed all 48 picture puzzles) and seeing how well you actually performed.**

### What is your willingness to pay to get feedback?

- Please indicate an amount between **-100** thalers and **100** thalers.
- Once you have made your decision whether or not you wish to receive feedback on your fluid intelligence, you cannot change it. **This also means, if you do get feedback, you cannot avoid it, and we will show you how well you actually performed.**
- With a willingness to pay of 0 thalers, there is a 50% chance to get feedback.
- If you want to increase your chance of getting feedback, consider reporting positive willingness to pay for feedback.
- If you want to increase your chance of not getting feedback, consider reporting negative willingness to pay for feedback.
- **The higher your willingness to pay is, the more likely you get feedback on your fluid intelligence.**

### Further explanations:

How does it work exactly?

We draw a random price of feedback between -100 ECU and 100 ECU. If your willingness to pay to get feedback is equal or greater than the price, you get feedback and pay the price. If your willingness to get feedback is smaller than the random price, you do not get feedback and keep your endowment. Please note that both a random price of feedback and your willingness to pay for it can be either positive or negative.

**This mechanism ensures that answering questions honestly is in your best interest.**

Examples

**Note: Carefully and honestly answering the question is in your best interest.**

**How much are you willing to pay to see your results at the end of the experiment?**

I definitely do not want to see my results | -100 | 0 | 100 | I definitely want to see my results

**My willingness to pay to get my results at the end of the experiment is -13 thalers.**

Weiter

Figure A7: Willingness-to-pay for feedback screen (answered)

## C.6 Willingness-to-pay for feedback: German (original)

# Feedback zu Ihrer fluiden Intelligenz am Ende des Experiments

**Ihre Entscheidung auf dieser Seite beeinflusst Ihre Chancen, am Ende des Experiments (d.h. nachdem Sie alle 48 Bilderrätsel bearbeitet haben) Feedback zu Ihrer fluiden Intelligenz zu erhalten und zu sehen wie gut Sie tatsächlich abgeschnitten haben.**

### Wie viel sind Sie bereit zu zahlen, um Feedback zu erhalten?

- Bitte geben Sie dazu einen Betrag zwischen **-100** Taler und **100** Taler an.
- Nachdem Sie Ihre Entscheidung getroffen haben, ob Sie Feedback über Ihre fluide Intelligenz erhalten wollen oder nicht, können Sie diese nicht mehr ändern. **Das bedeutet auch, dass Sie, wenn Sie Feedback bekommen, dieses nicht vermeiden können und wir Ihnen anzeigen werden, wie gut Sie tatsächlich abgeschnitten haben.**
- Bei einer Zahlungsbereitschaft von 0 Taler besteht eine 50%ige Wahrscheinlichkeit, Feedback zu erhalten.
- Wenn Sie die Wahrscheinlichkeit Feedback zu erhalten erhöhen möchten, sollten Sie eine positive Zahlungsbereitschaft angeben.
- Wenn Sie die Wahrscheinlichkeit erhöhen möchten, kein Feedback zu erhalten, sollten Sie eine negative Zahlungsbereitschaft für das Feedback angeben.
- **Je höher Ihre Zahlungsbereitschaft ist, desto wahrscheinlicher erhalten Sie Feedback zu Ihrer fluiden Intelligenz.**

### Weitere Erklärungen:

Wie funktioniert das genau?

Wir setzen einen zufälligen Preis für Feedback zwischen -100 Taler und 100 Taler fest. Wenn Ihre Zahlungsbereitschaft für Feedback gleich oder größer als der Preis ist, erhalten Sie Feedback und zahlen den Preis. Wenn Ihre Zahlungsbereitschaft für Feedback kleiner ist als der zufällig gesetzte Preis, erhalten Sie kein Feedback und behalten Ihre Anfangsausstattung. Bitte beachten Sie, dass sowohl der zufällig gesetzte Preis für Feedback als auch Ihre Zahlungsbereitschaft sowohl positiv als auch negativ sein können.

**Durch diesen Mechanismus wird sichergestellt, dass eine ehrliche Beantwortung der Fragen in Ihrem besten Interesse ist.**

Beispiele

**Beachten Sie: Eine sorgfältige und ehrliche Beantwortung der Fragen ist in Ihrem besten Interesse.**

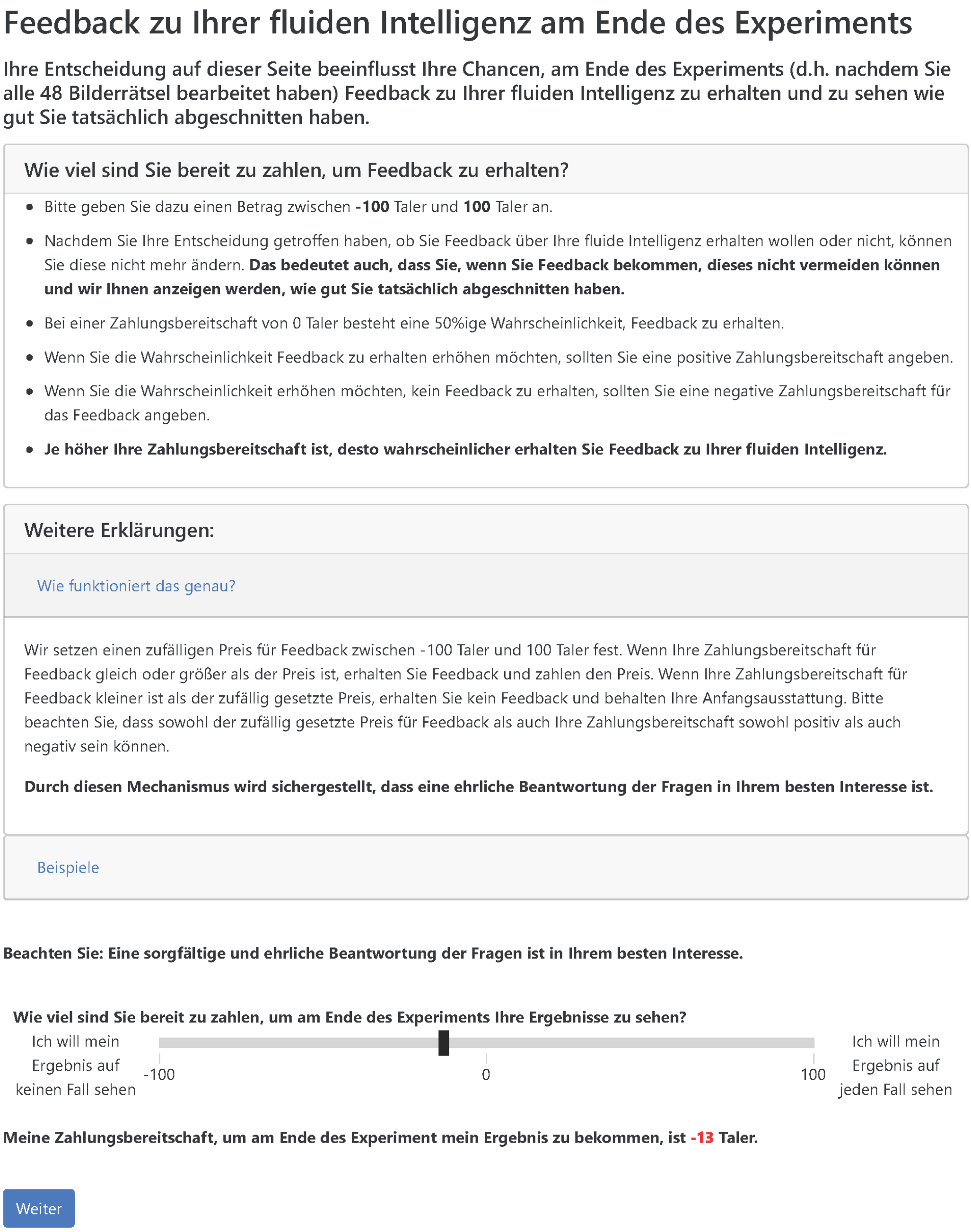
**Wie viel sind Sie bereit zu zahlen, um am Ende des Experiments Ihre Ergebnisse zu sehen?**

Ich will mein Ergebnis auf keinen Fall sehen -100 0 100 Ich will mein Ergebnis auf jeden Fall sehen

**Bitte klicken Sie auf den Schieberegler, um Ihre Zahlungsbereitschaft anzugeben. Sie können Ihre Entscheidung im Anschluss noch verändern.**

Weiter

Figure A8: Willingness-to-pay for feedback screen



## Feedback zu Ihrer fluiden Intelligenz am Ende des Experiments

**Ihre Entscheidung auf dieser Seite beeinflusst Ihre Chancen, am Ende des Experiments (d.h. nachdem Sie alle 48 Bilderrätsel bearbeitet haben) Feedback zu Ihrer fluiden Intelligenz zu erhalten und zu sehen wie gut Sie tatsächlich abgeschnitten haben.**

### Wie viel sind Sie bereit zu zahlen, um Feedback zu erhalten?

- Bitte geben Sie dazu einen Betrag zwischen **-100** Taler und **100** Taler an.
- Nachdem Sie Ihre Entscheidung getroffen haben, ob Sie Feedback über Ihre fluide Intelligenz erhalten wollen oder nicht, können Sie diese nicht mehr ändern. **Das bedeutet auch, dass Sie, wenn Sie Feedback bekommen, dieses nicht vermeiden können und wir Ihnen anzeigen werden, wie gut Sie tatsächlich abgeschnitten haben.**
- Bei einer Zahlungsbereitschaft von 0 Taler besteht eine 50%ige Wahrscheinlichkeit, Feedback zu erhalten.
- Wenn Sie die Wahrscheinlichkeit Feedback zu erhalten erhöhen möchten, sollten Sie eine positive Zahlungsbereitschaft angeben.
- Wenn Sie die Wahrscheinlichkeit erhöhen möchten, kein Feedback zu erhalten, sollten Sie eine negative Zahlungsbereitschaft für das Feedback angeben.
- **Je höher Ihre Zahlungsbereitschaft ist, desto wahrscheinlicher erhalten Sie Feedback zu Ihrer fluiden Intelligenz.**

### Weitere Erklärungen:

Wie funktioniert das genau?

Wir setzen einen zufälligen Preis für Feedback zwischen -100 Taler und 100 Taler fest. Wenn Ihre Zahlungsbereitschaft für Feedback gleich oder größer als der Preis ist, erhalten Sie Feedback und zahlen den Preis. Wenn Ihre Zahlungsbereitschaft für Feedback kleiner ist als der zufällig gesetzte Preis, erhalten Sie kein Feedback und behalten Ihre Anfangsausstattung. Bitte beachten Sie, dass sowohl der zufällig gesetzte Preis für Feedback als auch Ihre Zahlungsbereitschaft sowohl positiv als auch negativ sein können.

**Durch diesen Mechanismus wird sichergestellt, dass eine ehrliche Beantwortung der Fragen in Ihrem besten Interesse ist.**

Beispiele

**Beachten Sie: Eine sorgfältige und ehrliche Beantwortung der Fragen ist in Ihrem besten Interesse.**

**Wie viel sind Sie bereit zu zahlen, um am Ende des Experiments Ihre Ergebnisse zu sehen?**

Ich will mein Ergebnis auf keinen Fall sehen | -100 | 0 | 100 | Ich will mein Ergebnis auf jeden Fall sehen

**Meine Zahlungsbereitschaft, um am Ende des Experiment mein Ergebnis zu bekommen, ist -13 Taler.**

Weiter

Figure A9: Willingness-to-pay for feedback screen (answered)

## C.7 Questionnaire: English

### Please answer the following questions (Page 1 of 4)

**Now please fill in the following questions before we proceed to the payment. Please provide the following data about yourself.**

How old are you?

What is your gender?

---------

What is your occupation?

- Study
- Other

What do you study?

---------

How many experiments (approximately) have you already participated in?

What is your current average grade or that of your last degree?

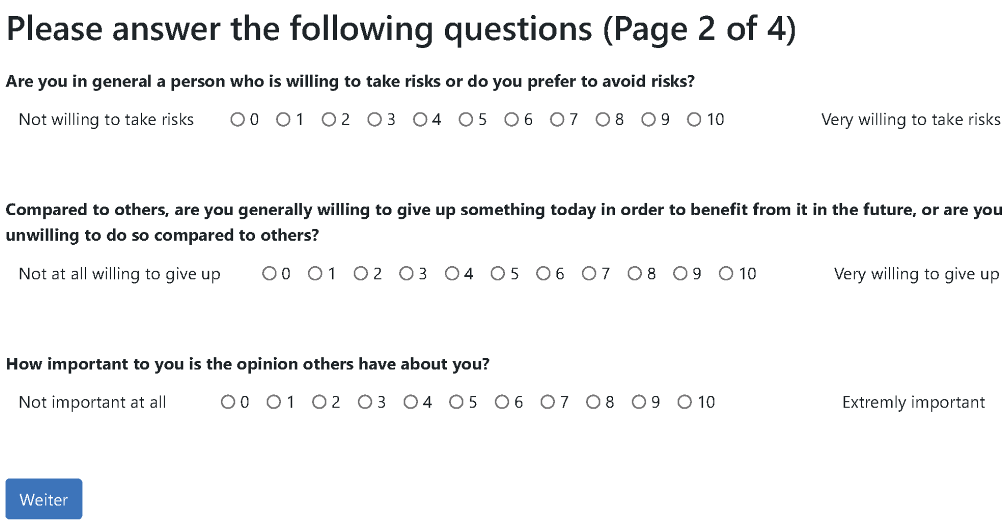
What was the final grade of your last school degree (1.0 - 4.0)?

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.)?

Weiter

Figure A10: Questionnaire: Page 1 of 4


## Please answer the following questions (Page 2 of 4)

**Are you in general a person who is willing to take risks or do you prefer to avoid risks?**

Not willing to take risks ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Very willing to take risks

**Compared to others, are you generally willing to give up something today in order to benefit from it in the future, or are you unwilling to do so compared to others?**

Not at all willing to give up ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Very willing to give up

**How important to you is the opinion others have about you?**

Not important at all ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Extremly important

Weiter

Figure A11: Questionnaire: Page 2 of 4

## Please answer the following questions (Page 3 of 4)

**How strongly do you agree with the following statements? Please answer on a scale of 0 to 5, where 0 means you do not agree with the statement at all and 5 means you agree very strongly.**

**This is about the characteristic "fairness"**

1. I would feel good if I was a person who had this quality.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

2. Being someone who has this quality is an important part of who I am.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. A big part of my emotional well-being is tied to having this quality.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. I would be ashamed to be a person who has this characteristic.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Having this characteristic is not really important to me.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Having this quality is an important part of my self-image.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

**This is about the characteristic "generosity"**

1. I would feel good if I was a person who had this quality.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

2. Being someone who has this quality is an important part of who I am.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. A big part of my emotional well-being is tied to having this quality.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. I would be ashamed to be a person who has this characteristic.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Having this characteristic is not really important to me.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Having this quality is an important part of my self-image.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

**This is about the characteristic "kindness"**

1. I would feel good if I was a person who had this quality.
○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Figure A12: Questionnaire: Page 3 of 4

2. Being someone who has this quality is an important part of who I am.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. A big part of my emotional well-being is tied to having this quality.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. I would be ashamed to be a person who has this characteristic.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Having this characteristic is not really important to me.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Having this quality is an important part of my self-image.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Weiter

Figure A13: Questionnaire: Page 3 of 4

# Lottery decisions (Page 4 of 4)

Please answer a few more questions about lotteries where you can earn or lose money once again if you decide to accept them.

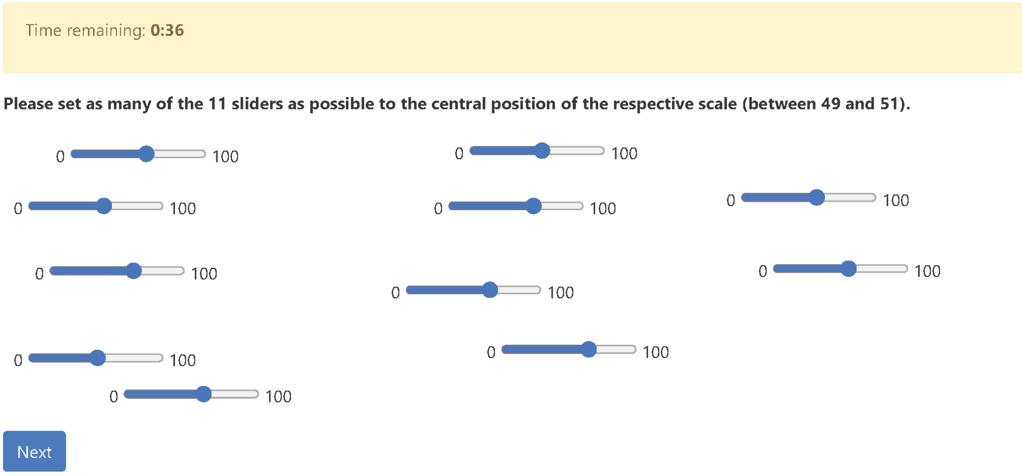
Below are listed **6 different lotteries**:

- For each of the 6 lotteries, you can choose to accept or reject the lottery.
- If you reject a lottery, your payment will remain unchanged. If you accept a lottery, you will realize an additional profit or an additional loss.
- 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 profit or loss will be realized.

| Lottery | Your decision | | | |
|---|---|---|---|---|
| With a probability of 50 % you **lose** 4 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accept** | ○ | ○ | **Reject** |
| With a probability of 50 % you **lose** 6 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accept** | ○ | ○ | **Accept** |
| With a probability of 50 % you **lose** 8 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accpet** | ○ | ○ | **Reject** |
| With a probability of 50 % you **lose** 10 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accept** | ○ | ○ | **Reject** |
| With a probability of 50 % you **lose** 12 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accept** | ○ | ○ | **Reject** |
| With a probability of 50 % you **lose** 14 thalers.<br>With a probability of 50 % you **win** 12 thalers. | **Accept** | ○ | ○ | **Reject** |

Weiter

Figure A14: Questionnaire: Page 4 of 4

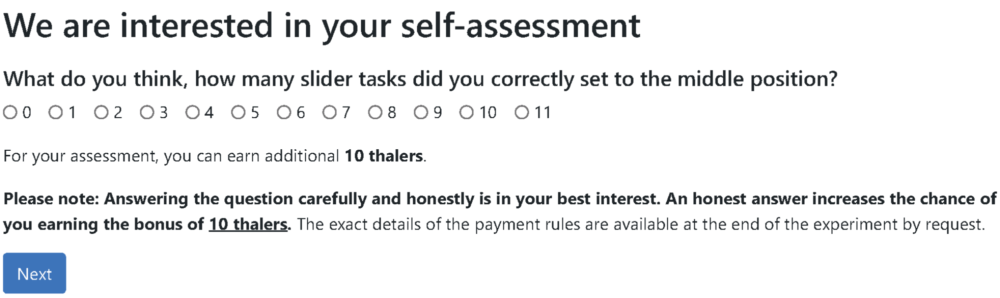

Time remaining: **0:36**

**Please set as many of the 11 sliders as possible to the central position of the respective scale (between 49 and 51).**

0 100
0 100
0 100
0 100
0 100
0 100
0 100
0 100
0 100
0 100
0 100

Next

Figure A15: Overconfidence elicitation: Slider task


## We are interested in your self-assessment

### What do you think, how many slider tasks did you correctly set to the middle position?

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 ○ 11

For your assessment, you can earn additional **10 thalers**.

**Please note: Answering the question carefully and honestly is in your best interest. An honest answer increases the chance of you earning the bonus of 10 thalers.** The exact details of the payment rules are available at the end of the experiment by request.

Next

Figure A16: Overconfidence elicitation: Self-assessment

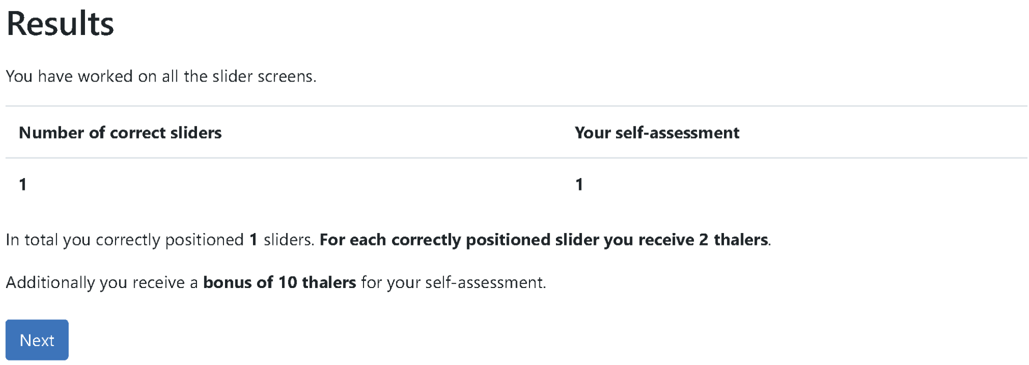



Figure A17: Overconfidence elicitation: Feedback

## C.8 Questionnaire: German (original)

# Bitte beantworten Sie die folgenden Fragen (Seite 1 von 4)

**Füllen Sie nun bitte die folgenden Fragen aus, bevor wir zur Auszahlung kommen. Bitte geben Sie die folgenden Daten zu Ihrer Person an.**

Wie alt sind Sie?

Was ist Ihr Geschlecht?

---------

Was ist Ihre Tätigkeit?

- Studium
- Anderes

Was studieren Sie?

---------

An wie vielen Experimenten haben Sie (ungefähr) bereits teilgenommen?

Was ist Ihre akutelle Durchschnittesnote bzw. die Ihres letzten Abschlusses?

Was war die Abschlussnote Ihres letzten Schulabschlusses (1,0 - 4,0)?

Wie viel Geld haben Sie monatlich (nach Abzug von Fixkosten wie Miete, Versicherungen etc.) zur Verfügung?

Wie viel Geld geben Sie monatlich aus (nach Abzug von Fixkosten wie Miete, Versicherungen etc.)?

Weiter

Figure A18: Questionnaire: Page 1 of 4

## Bitte beantworten Sie noch die folgenden Fragen (Seite 2 von 4)

**Sind Sie im Allgemeinen ein risikobereiter Mensch oder versuchen Sie, Risiken zu vermeiden?**

Gar nicht risikobereit ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Sehr risikobereit

**Sind Sie im Vergleich zu anderen im Allgemeinen bereit heute auf etwas zu verzichten, um in der Zukunft davon zu profitieren oder sind Sie im Vergleich zu anderen dazu nicht bereit?**

Gar nicht bereit zu verzichten ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Sehr bereit zu verzichten

**Wie wichtig ist Ihnen die Meinung, die andere über Sie haben?**

Überhaupt nicht wichtig ○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Extrem wichtig

Weiter

Figure A19: Questionnaire: Page 2 of 4

## Bitte beantworten Sie noch die folgenden Fragen (Seite 3 von 4)

**Wie sehr stimmen Sie den folgenden Aussagen zu? Bitte antworten Sie auf einer Skala von 0 bis 5. Dabei bedeutet 0, dass Sie der Aussage gar nicht zustimmen und 5, dass Sie sehr stark zustimmen.**

**Es geht um die Eigenschaft "Fairness"**

1. Ich würde mich gut fühlen, wenn ich eine Person wäre, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

2. Jemand zu sein, der diese Eigenschaft hat, ist ein wichtiger Teil von dem, was ich bin.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. Ein großer Teil meines emotionalen Wohlbefindens ist damit verbunden, diese Eigenschaft zu haben.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. Ich würde mich schämen, eine Person zu sein, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Diese Eigenschaft zu haben, ist für mich nicht wirklich wichtig.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Diese Eigenschaft zu haben, ist ein wichtiger Teil meines Selbstverständnisses.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

**Es geht um die Eigenschaft "Großzügigkeit"**

1. Ich würde mich gut fühlen, wenn ich eine Person wäre, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

2. Jemand zu sein, der diese Eigenschaft hat, ist ein wichtiger Teil von dem, was ich bin.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. Ein großer Teil meines emotionalen Wohlbefindens ist damit verbunden, diese Eigenschaft zu haben.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. Ich würde mich schämen, eine Person zu sein, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Diese Eigenschaft zu haben, ist für mich nicht wirklich wichtig.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Diese Eigenschaft zu haben, ist ein wichtiger Teil meines Selbstverständnisses.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

**Es geht um die Eigenschaft "Freundlichkeit"**

1. Ich würde mich gut fühlen, wenn ich eine Person wäre, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Figure A20: Questionnaire: Page 3 of 4

2. Jemand zu sein, der diese Eigenschaft hat, ist ein wichtiger Teil von dem, was ich bin.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

3. Ein großer Teil meines emotionalen Wohlbefindens ist damit verbunden, diese Eigenschaft zu haben.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

4. Ich würde mich schämen, eine Person zu sein, die diese Eigenschaft hat.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

5. Diese Eigenschaft zu haben, ist für mich nicht wirklich wichtig.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

6. Diese Eigenschaft zu haben, ist ein wichtiger Teil meines Selbstverständnisses.

○ 0 ○ 1 ○ 2 ○ 3 ○ 4 ○ 5

Weiter

Figure A21: Questionnaire: Page 3 of 4

## Lotterieentscheidungen (Seite 4 von 4)

Bitte beantworten Sie im Folgenden noch ein paar Fragen zu Lotterien, bei denen Sie noch einmal Geld verdienen oder auch verlieren können, falls Sie sich entscheiden, die Lotterien zu akzeptieren.

Unten sind **6 verschiedene Lotterien** aufgelistet:

- Sie können für jede der 6 Lotterien wählen, ob Sie die Lotterie akzeptieren oder ablehnen möchten.
- Falls Sie eine Lotterie ablehnen, bleibt Ihre Auszahlung unverändert. Falls Sie eine Lotterie akzeptieren, werden Sie einen zusätzlichen Gewinn oder einen zusätzlichen Verlust realisieren.
- Am Ende des Experiments wird zufällig eine der 6 Lotterien ausgewählt.
- Sie sollten also jede Lotterieentscheidung so treffen, als wäre es Ihre einzige Entscheidung. Die ausgewählte Lotterie wird anschließend ausgelost, um festzustellen, ob sich der zusätzliche Gewinn oder Verlust realisiert.

| Lotterie | Ihre Entscheidung | | | |
|---|---|---|---|---|
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 4 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 6 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 8 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 10 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 12 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |
| Mit einer Wahrscheinlichkeit von 50 % **verlieren** Sie: 14 Taler.<br>Mit einer Wahrscheinlichkeit von 50 % **gewinnen** Sie 12 Taler. | **Akzeptieren** | ○ | ○ | **Ablehnen** |

Weiter

Figure A22: Questionnaire: Page 4 of 4

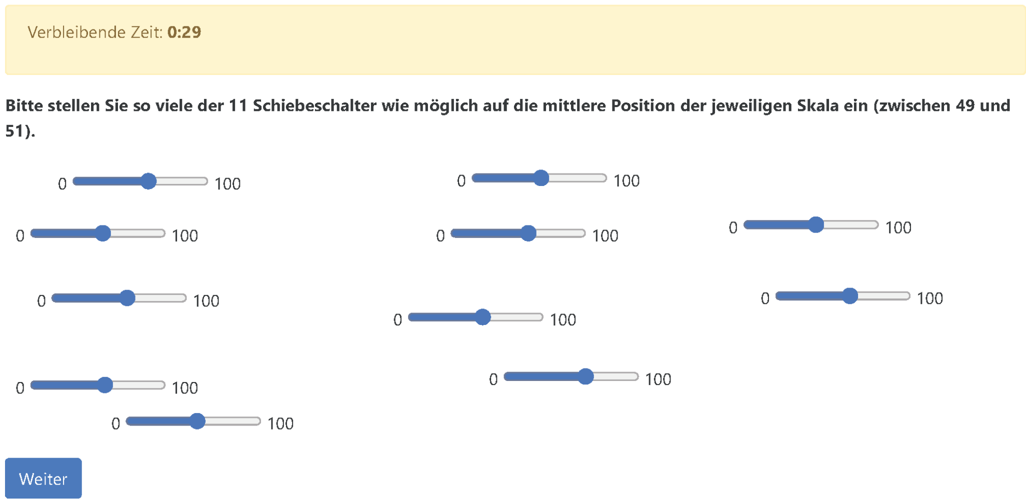



Figure A23: Overconfidence elicitation: Slider task


Figure A24: Overconfidence elicitation: Self-assessment

## Ergebnisse

Sie haben alle Schieberbildschirme bearbeitet.

| Anzahl korrekter Schieber | Ihre Selbsteinschätzung |
|---|---|
| **1** | **1** |

Sie haben insgesamt **1** Schieber korrekt positioniert. **Für jeden korrekten Schieber erhalten Sie 2 Taler**.

Zusätzlich erhalten Sie einen **Bonus von 10 Taler** für die Selbsteinschätzung.

Weiter

Figure A25: Overconfidence elicitation: Feedback