universität innsbruck

Research Platform
Empirical and Experimental Economics

Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

Laura Hueber, Rene Schwaiger

Working Papers in Economics and Statistics

2021-01

University of Innsbruck
https://www.uibk.ac.at/eeecon/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform "Empirical and Experimental Economics"
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

Laura Hueber[†] Rene Schwaiger[‡]

January 3, 2021

Abstract

We introduce a training intervention based on a novel tool to mitigate behavior consistent with myopic loss aversion (MLA). We present the results of a large-scale online experiment with 894 student participants. The study featured a two-step debiasing training intervention based on experience sampling and a subsequent elicitation of MLA. We found that participants at baseline exhibit behavior consistent with MLA, which was not the case for decisionmakers who underwent the debiasing training intervention. Nonetheless, we found no statistically significant difference-in-difference effect of the training intervention on the magnitude of MLA. However, when we focused on the more attentive participants by excluding participants with the 10% longest and 10% shortest processing times on the task relevant instruction screens, the magnitude of the difference-in-difference effect of the training intervention increased strongly and became statistically significant when controlling for age, gender, education, field of study, investment experience, and financial risk preferences.

JEL classification: G11 G41 G51

Keywords: online experiment, myopic loss aversion, debiasing, experience sampling

We thank Felix Holzmeister, Christoph Huber, Jürgen Huber, Michael Kirchler, Christian König, Michael Razen and Matthias Stefan for very valuable comments on previous versions of the paper. The financial support of the Netzwerk Banking, Accounting, Auditing, Finance & IT (BAFIT), the Research Platform Empirical and Experimental Economics (eeecon) and the University of Innsbruck is gratefully acknowledged. The experimental data for replication and the software are available on OSF. The are no conflicts of interest for any of the authors of this paper.

[†]University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, Austria, e-mail: laura.hueber@uibk.ac.at.

[‡]Corresponding author, University of Innsbruck, Department of Banking and Finance, Universitätsstrasse 15, 6020 Innsbruck, Austria, e-mail: rene.schwaiger@uibk.ac.at, tel: +43 (0) 512 507 73019.

1 Introduction

Loss aversion (Kahneman and Tversky, 1979) and temporal myopia, described as "the inability to consider the long-term outcomes of an action when making a choice" (Christensen and Bickel, 2010, p. 118), are two deviations from neoclassical theory that when combined form myopic loss aversion (MLA), which has a negative impact on an individual's financial decision making (Looney and Hardin, 2009). MLA is present not only among students in individual decisions (Gneezy and Potters, 1997; Thaler et al., 1997; Wendy and Asri, 2012) but also in market settings (Gneezy et al., 2003) and among teams as decisionmakers (Sutter, 2007). Furthermore, MLA-consistent behavior is confirmed among highly educated experts, such as financial professionals (Haigh and List, 2005), as well.[1] In a natural field experiment, the latter group has also been associated with MLA in asset markets–their everyday working environment (Larson et al., 2012). Therefore, this paper presents a novel interactive debiasing tool and experimentally tests whether a training intervention using this tool can successfully reduce participants' susceptibility to MLA.

Generally, MLA-consistent behavior is especially harmful for individuals who hold investments with relatively high short-term volatility, such as stocks, while following a long-term investment horizon. Stocks are often accompanied by very positive long-term return expectations; therefore, they have been an important and successful way to build up wealth in the past (Jordà et al., 2018). Nevertheless, loss-averse individuals observing lower or even negative short-term returns on their stocks compared to treasury bills or bonds, due to greater price fluctuations, are expected to be more likely than others to rebalance their portfolios–consequently reducing their equity investments. However, in the context of long-term investments, it is important to note that such a short-sighted assessment of the performance of equity portfolios is inappropriate. If investors became less inclined to evaluate their portfolio frequently and, instead, looked at aggregate returns, they would be more likely to observe positive results through statistical aggregation. On average, these would exceed the returns of other financial asset classes, such as bonds and treasury bills (Gneezy et al., 2003; Jordà et al., 2018). MLA leads to the opposite, i.e., a focus on short-term performance and a tendency to reduce risk after poor results. This is the reason why MLA is associated with reduced long term payoffs of equity portfolios (Looney and Hardin, 2009).

This can become a structural problem as the importance of an adequate private pension provision might increase in the future due to lower projected available national pensions in Europe (Hülsewig and Hülsewig, 2017) and shifts from defined benefit plans to defined contribution plans in the U.S. (U.S. Department of Labor, 2014). A representative survey in the U.S. shows that, on average, participants estimate a 45% probability that they will outlive their savings, and 41% of participants have not yet taken action against this (Northwestern Mutual, 2019). Therefore, it is of social relevance to correct behavior consistent with MLA and to contribute to improving the financial decision making of investors.[2]

[1] MLA aligns with a large number of deviations from neoclassical predictions that are empirically supported among students (Tversky and Kahneman, 1974; Svenson, 1981; Samuelson and Zeckhauser, 1988; Kahneman et al., 1990; Grosshans and Zeisberger, 2018) and among professionals from different domains (Roszkowski and Snelbecker, 1990; Haigh and List, 2005; Cipriani and Guarino, 2009; Deaves et al., 2010; Abdellaoui et al., 2013; Menkhoff and Schmeling, 2013; Pikulina et al., 2017; Kirchler et al., 2018; Sheffer et al., 2018; Huber et al., 2019; Schwaiger et al., 2020).

[2] Schooley and Drecnik Worden (2013) demonstrate that there is a positive association between a households' equity investments and the likelihood of having a pension plan. Generally, the suitability of equity investments for retirement savings also depends crucially on the investment horizon, i.e., the time until retirement. If this investment horizon is relatively short, shares may not be the best option due to their relatively high short-term volatility (Bodie et al., 1992).

In order to precisely assess the potential of the developed tool to mitigate MLA, the tools' underlying investment process is based on a lottery introduced by Gneezy and Potters (1997), the foundation for the most frequently applied measure of MLA available (e.g., Bellemare et al., 2005; Haigh and List, 2005; Fellner and Sutter, 2009). Consequently, to measure the influence of the training intervention on behavior accurately, we chose the same investment domain for the training intervention as for the elicitation procedure, i.e., the lottery. This implies that the tool is not based on risky decisions in the equity investment domain. Nevertheless, the underlying characteristics of the lottery are similar, and the tool is very flexible, so it can easily be extended to other investment areas, such as equity investments. Furthermore, the tool can potentially be used in financial consulting and planning, as well.

We set up an online experiment with student participants from the University of Innsbruck. In a between-subject design, participants were randomly assigned to either the baseline or the debiasing treatment with two experimental stages each. Only the first stage differed between the treatments. In the first stage of the baseline treatment, the participants played the game Minesweeper as a filler task. In the first stage of the debiasing treatment, the participants were confronted with a two-step training intervention based on experience sampling to familiarize themselves with the underlying properties of the lottery introduced in Gneezy and Potters (1997) and the implications of different betting decisions in this lottery. In the second stage, which was identical for the baseline and debiasing treatment, we measured MLA according to Gneezy and Potters (1997). To the best of our knowledge, this is the first study aimed at reducing MLA-consistent behavior with an interactive training intervention based on experience sampling.

In the baseline treatment, we found statistically significant evidence of behavior consistent with MLA. However, we did not find statistical evidence of MLA-consistent behavior in the debiasing treatment. When we directly tested for the difference-in-difference effect of the training intervention on MLA, we found no statistically significant reduction in MLA as a result of the training intervention, which is supported by randomization inference. Furthermore, in an exploratory approach, we excluded the participants with the 10% longest and 10% shortest processing times and repeated the main analyses with a sample comprised of more attentive participants. In contrast to the full sample, we found a stronger effect of the training intervention on MLA. It was statistically significant when controlling for age, gender, education, field of study, investment experience, and financial risk preferences. The statistical significance is also confirmed by randomization inference. Specifically, Tobit regression analyses predict that the training intervention reduces behavior consistent with MLA by 10 percentage points compared to the baseline. We conclude that the developed tool can reduce susceptibility to MLA-consistent behavior of participants with high attention and focus to the training intervention.

This study contributes to several strands in the literature. First, it contributes to the literature on MLA. MLA was introduced by Benartzi and Thaler (1995) as an explanation for the equity premium puzzle (Mehra and Prescott, 1985); ever since, MLA has been manifoldly studied (e.g., Gneezy and Potters, 1997; Gneezy et al., 2003; Haigh and List, 2005; Wendy and Asri, 2012). The two pillars of MLA—loss aversion and myopia—have long been established empirically. Students (Abdellaoui et al., 2007; Morrison and Oxoby, 2014) and members from the general population (Gächter et al., 2007) have been shown to exhibit loss aversion. Furthermore, empirical studies have shown that economic decisionmakers tend to act myopically, i.e., they have a tendency to frequently evaluate financial outcomes. Based on survey data, Lee and Veld-Merkoulova (2016) report that 44% of investors from the general Dutch population examine their stock portfolio at least once a month. This trend can be explained by the strong preference of individuals for immediate and frequent outcome feedback (Fellner and Sutter, 2009) and by mental accounting (Kahneman and Tversky, 1984) in an inter-temporal context. Specifically, behavior depends on how people aggregate outcomes and decisions, whereby people who suffer from myopia frame them narrowly (Benartzi and Thaler, 1995). We specifically add to the literature by measuring MLA according to (Gneezy and Potters, 1997) among student participants in an online experiment, resulting in slight deviations from the original instructions and a less controlled environment. Therefore, we offer a robustness check for the findings of MLA among students in laboratory settings.

Secondly, the study contributes to a nascent stream in the literature on systematic debiasing of existing cognitive biases. The literature distinguishes between three main categories of debiasing approaches namely (i) changing underlying incentives, (ii) improving the framing and the elicitation of decisions, and (iii) reducing biases through training (Morewedge et al., 2015). This study specifically adds to the last category. Generally, the literature on training interventions to improve decision making is mixed. Kaustia and Perttula (2012) present evidence that the better-than-average type of overconfidence might be reduced by communicating explicit warnings to participants. Nevertheless, this approach does not work regarding overconfidence in probability assessments. Kučera (2020) shows that confirmation bias can be statistically significantly reduced by presenting a video on confirmation bias and its impact and mitigation strategies. Fong and Nisbett (1991) provide evidence for successfully improving statistical reasoning over a longer period of time by providing example problems in a training intervention. Morewedge et al. (2015) achieve medium to large reductions of biases such as blind spot, confirmation bias, fundamental attribution error, anchoring, representativeness, and social projection. We expressly contribute by introducing an interactive debiasing tool to mitigate behavior consistent with the cognitive bias MLA.

Thirdly, the study contributes to the small but growing strand on experience sampling in finance. Prominently, Kaufmann et al. (2013) show that experience sampling and a risk tool combining experience sampling with graphical illustrations and numerical descriptions influence risk preferences. In particular, the authors find that participants increase the allocation of funds in the risky asset after being able to sample from the distribution of the risky asset. Cason and Samek (2015) report that mispricing in experimental asset markets is reduced when participants are confronted with passive pre-market training and visual representations of trade prices before actively engaging in trading. Lusardi et al. (2017) provide evidence that financial literacy and/or confidence in financial decision making improves when information is provided via videos or visual interactive tools using experience sampling. Nevertheless, Bradbury et al. (2019) only find weak support for persistent changes in investor behavior due to risk simulations. The authors argue that experience sampling might only influence the initial investment decision. We specifically contribute by exploring the role of an experience sampling based tool in financial behavior.

2 Experimental Design and Procedure

The experiment consisted of two main stages. In the first stage, participants were randomly assigned either to the training treatment, i.e., treatment DEBIASING, or the baseline treatment, i.e., treatment BASELINE. Participants in the treatment DEBIASING group underwent a training intervention tailored to mitigate or eliminate behavior consistent with MLA. Participants in the treatment BASELINE group played the game Minesweeper as an independent filler task for at least 10 minutes and 5 repetitions, which corresponded to the planned time for the training intervention.[3] This was to ensure that the expected processing time for the filler task is comparable to the expected processing time of the training intervention. Similar to the training intervention, the filler task required a certain amount of cognition and attention. In both treatments we informed the participants that, in contrast to the second stage, their decisions from the first stage of the experiment are not relevant to the payoff. In the second stage of the experiment, which was identical for participants in treatments DEBIASING and BASELINE, we measured whether participants' behavior was consonant with the theory of MLA according to Gneezy and Potters (1997) to investigate the effectiveness of the intervention.

In an exit questionnaire, we asked participants to provide information on their financial risk preferences; their individual experience with financial investments; and their demographic and socio-economic characteristics, such as age, gender, education, and field of study.[4,5] The training intervention in the DEBIASING treatment and the subsequent examination of MLA in both treatments were based on the following lottery in Gneezy and Potters (1997):

> *With a probability of one-in-three (33%) you win 2.5 times the amount you bet and with a probability of two-in-three (67%) you lose the amount you bet*.[6]

2.1 Debiasing Training Intervention

According to Muradoglu and Harvey (2012) the presentation of aggregated outcome diagrams of investment processes with otherwise frequent outcome feedback could reduce the susceptibility of individuals to MLA by distracting from myopic decisions. Furthermore, Bradbury et al. (2019) argue that easy-to-read graphical representations such as histograms are important for risk communication. In addition to this finding, experience sampling and a risk instrument that combines experience sampling with graphical representations and numerical descriptions influence risk preferences (Kaufmann et al.,

[3]After 10 minutes and 5 repetitions, a "Next" button appeared and participants could continue the experiment.

[4]To check the consistency of the answers on the question on financial risk, risk preferences in general were also examined. We find a statistically significant and strongly positive correlation between financial and general risk preferences (Spearman's rho = 0.6235, $p < 0.005$, $N = 894$). Henceforth, we use the domain relevant financial risk preferences in the analyses.

[5]The self-reported risk preferences were measured using the German SOEP questionnaire (Dohmen et al., 2011) on Likert scales from 0 to 10.

[6]As reported in the pre-registration, we initially started the experiment with the following adapted lottery properties (Charness et al., 2019): *With a probability of one-in-two (50%) you win 2.5 times the amount you bet and with a probability of one-in-two (50%) you lose the amount you bet.* This was done to ensure a higher expected value of the lottery, to clearly distinguish the prospects of a constant investment from an investment reduction (more in *Chosen_Bet* and *Reduced_Bet* in Section 2.1). We deviated from these lottery characteristics after piloting the software and finding that the student subjects in the online experiment did not exhibit behavior consistent with MLA in the baseline treatment. Since MLA-consistent behavior is a prerequisite for measuring the effectiveness of the novel debiasing training intervention, we performed a robustness check and applied another pilot for the baseline treatment with the original properties in Gneezy and Potters (1997), finding behavior consistent with MLA. Thus, this paper is based on the original lottery properties in Gneezy and Potters (1997).

2013). As a result we integrated these components into our intervention presented in this paper. Consequently, our developed tool is based on experience sampling, with its insights communicated through easy-to-read, aggregating histograms and numerical-descriptive tables.

In the DEBIASING treatment group, we implemented the training tool, which consisted of two steps to familiarize participants with the inherent characteristics of the lottery and the implications of different decisions regarding the bet amount in the lottery. Participants were endowed with 200 tokens for each of nine rounds and had to choose an amount x $(0 < x < 200)$ in tokens at the beginning of the first round, which was used to illustrate the characteristics of the lottery.[7]

Importantly, we introduced two fictitious scenarios, *Chosen_Bet* and *Reduced_Bet*, which served as a basis for the illustration.[8] The two scenarios differed only in terms of the amount of tokens actually bet in the lottery in the specific rounds of the illustration to demonstrate the implications of a behavior, consistent with the theory of MLA. In the first scenario *Chosen_Bet*, the amount x in tokens chosen by participants to illustrate the lottery was kept constant and bet in all nine rounds, regardless of previous earnings in the lottery. In the second scenario *Reduced_Bet*, the chosen amount x was reduced by 20% of the amount originally chosen by participants after a first loss was incurred.[9] This reduced amount was then bet in each subsequent round following that loss. Once a second loss occurred, the amount was further reduced by 20% of the originally chosen amount and bet in subsequent rounds following the second loss. This procedure was applied after each iterative loss in the lottery until five losses occurred. Then, an amount of zero was bet in all subsequent rounds. Consequently, the two scenarios, *Chosen_Bet* and *Reduced_Bet*, yielded identical results if no loss occurred in any of the nine rounds, or if the participants chose to bet an amount of zero to illustrate the lottery. The idea underlying the introduction of the two scenarios is directly derived from the basic problem caused by MLA. Participants suffering from MLA tend to deviate from their originally chosen risk level on the basis of outcome feedback and to reduce their risk level after losses. Nevertheless, due to the nature of the lottery presented, i.e., the underlying expected value, wealth after nine rounds is on average higher in the *Chosen_Bet* scenario than in the *Reduced_Bet* scenario, even though the former entails higher volatility.

For a simple, understandable, and direct comparison of the two scenarios, participants were presented *Simulation A* in the experiment, i.e., a dynamic bar chart with bars showing the cumulative wealth in tokens over nine rounds. The bar chart showed one bar for scenario *Chosen_Bet* and one bar for scenario *Reduced_Bet*. The bars developed gradually over nine rounds and represented the cumulative wealth in tokens after each of the nine lottery draws. Therefore, each bar after nine rounds showed the

[7] We use tokens as our experimental currency unit in the paper. Note that we used the term "Taler" in the software to tailor the wording to the German speaking participants.

[8] At first, we planned a third scenario called "No-bet", which showed the consequences of betting zero in all rounds of the lottery, which simply equals total wealth per round corresponding to the original endowment. After receiving feedback from students in a pilot of the software that the instructions were too long and cumbersome to read, we decided to discard this scenario altogether, as it is the least important in targeting MLA.

[9] The 20% reduction was based on the average periodic percentage point difference in bet amounts as a percentage of the endowment of 200 (≈ 17 percentage points) between the high and low feedback frequency groups in Gneezy and Potters (1997). We rounded this number up to 20 to facilitate any calculations by participants. Example: If a participant enters an amount of 120 tokens to illustrate the lottery over nine rounds, 100% of this amount will be bet in the first round. After a first loss, the amount bet in the following rounds will be reduced to 80% of the originally chosen amount (96 tokens), which corresponds to a reduction of 20 percentage points. This again corresponds to a reduction of 20% of the original amount of 120 tokens (24 tokens) and so on. As percentage points might be confusing for some participants, we have expressed the reduction as a percentage of the initially chosen amount.

cumulative wealth in tokens after all nine rounds in each scenario. In parallel, subjects were provided with a simultaneously evolving table that numerically displayed wealth, wealth differences between *Chosen_Bet* and *Reduced_Bet*, and the random draw of the lottery determined by the computer, i.e., win or loss, in each of the nine rounds and in each scenario. The lottery results were highlighted in green and red, depending on whether a win (green) or a loss (red) was determined by the computer. In addition, wealth differences between the two scenarios were highlighted in green (or red) when, in a given round and accumulated over nine rounds, wealth in tokens in the *Chosen_Bet* scenario was higher (or lower) than wealth in tokens in the *Reduced_Bet* scenario. Thus, participants were provided with the specific lottery results by both graphical and numerical representation, which also addresses differences in learning preferences (Fleming and Mills, 1992; Caligaris et al., 2015). There were possibilities to pause the process at any time to get an overview of the outcomes so far and also to go through the process without many single clicks. Specifically, the lottery simulation over nine rounds could be carried out either step by step by clicking on the respective button for each lottery draw individually, or continuously, by clicking once on the respective button to initialize automatic lottery draws over nine rounds in one run. Importantly, executing nine lottery draws corresponded to one iteration of *Simulation A*. Participants were required to perform at least 15 iterations of *Simulation A*.

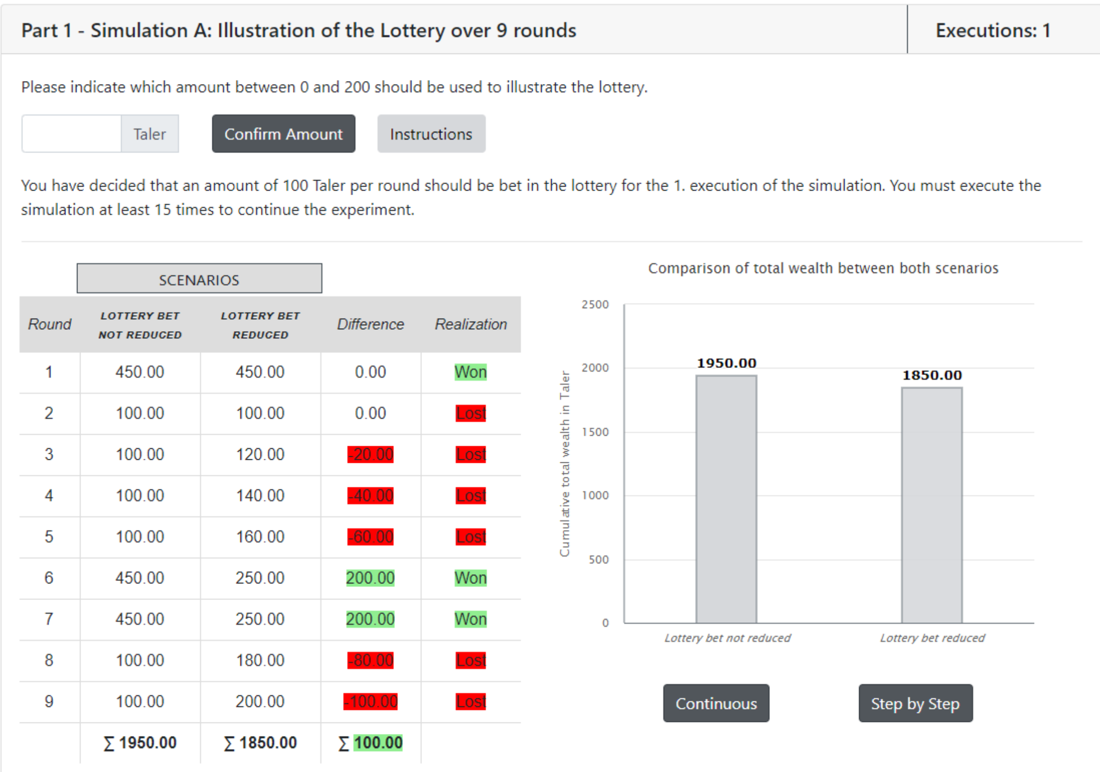


Figure 1: Experimental screen of Simulation A (English translation). The figure shows the experimental screen of the first step of the debiasing training intervention. The right side of the screen shows a gradually evolving bar chart presenting cumulative wealth in tokens in each of the nine rounds. The left bar in this chart shows cumulative wealth in tokens in scenario *Chosen_Bet* while the right bar in this chart shows cumulative wealth in tokens in scenario *Reduced_Bet*. The numbers at the top of the bars display cumulative wealth after each of the nine rounds. On the left side, an additional table is displayed, which presents the numbers processed in the bar chart. In particular, wealth in each of the nine rounds is displayed in scenario *Chosen_Bet* and *Reduced_Bet*, the numerical difference between both scenarios and the lottery realizations drawn by the computer in the respective rounds are additionally shown and colored in green or red depending on which scenario resulted in higher wealth and whether the lottery realized a loss or a win in the respective round.

This requirement was established to provide a reasonable understanding of the link between the realizations of the lottery and the accumulated wealth after nine rounds between the two scenarios. After 15 iterations, a pop-up window displayed the average cumulative wealth in tokens after nine rounds over all 15 iterations in the *Chosen_Bet* and the *Reduced_Bet* scenario and a “Next” button appeared. Figure 1 shows an example of *Simulation A*.

In a second step, participants were presented with *Simulation B* containing two simultaneously evolving histograms, each showing the distribution of 15,000 draws of cumulative wealth in tokens in the lottery over nine rounds in the *Chosen_Bet* and *Reduced_Bet* scenario, respectively. In both scenarios, the gradual evolution of the distribution of cumulative wealth after nine rounds was based on a hypothetical amount x $(1 < x < 200)$ in tokens to be chosen by the participants.

Figure 2: Experimental screen of Simulation B (English translation). The figure shows the experimental screen of the second step of the debiasing training intervention. Both graphs represent a gradually evolving histogram showing the distributions of cumulative wealth after nine rounds based on 15,000 draws in scenario *Chosen_Bet* (left graph), and scenario *Reduced_Bet*, (right graph). The vertical dashed lines display average cumulative wealth after nine rounds based on 15,000 draws for both scenarios.

After simulation of the 15,000 draws, both histograms showed final distributions and mean values of realized cumulative wealth after nine rounds for both scenarios *Chosen_Bet* and *Reduced_Bet*, respectively. As the simulations in both scenarios were based on 15,000 draws of cumulative wealth after nine rounds, the average values in the *Chosen_Bet* and *Reduced_Bet* scenarios approached the respective expected values of cumulative wealth after nine rounds for the initially chosen level of risk. Thus, the presentation of the distribution provided the participants with the expected wealth difference between maintaining the initially chosen risk level (i.e., *Chosen_Bet*) and reducing the initially chosen risk level after the occurrence of losses (i.e., *Reduced_Bet*). Furthermore, the histograms provided graphical information about the overall dispersion of cumulative wealth after nine rounds in each of the two scenarios. The second step had to be performed at least twice to continue the experiment. In addition to *Simulation A*, which should give participants an impression of the associations between the lottery results and cumulative wealth after nine rounds, Simulation B conveyed the broader picture communicating the theoretical properties of the lottery. Figure 2 shows an example of *Simulation B*.

The simulation-based learning described above was intended to show participants that given an initially chosen amount to be bet in the lottery, it is, on average, sub-optimal economically to reduce the initially chosen risk level due to realized losses in the lottery. Thus, we demonstrated MLA-consistent behavior and its negative effects on wealth in order to implicitly make participants less sensitive to realized short-term losses by repeatedly drawing from the underlying lottery distribution and, at the same time, introducing a more aggregated graphical and numerical representation.[10]

2.2 Elicitation of MLA

The second stage of the experiment was identical for the treatments DEBIASING and BASELINE and was concerned with the measurement of MLA-consistent behavior. Participants were told that the decisions in stage two of the experiment are payoff relevant. Specifically, in each of the nine consecutive rounds, participants chose an amount x $(0 < x < 200)$ in tokens of an endowment per round of 200 tokens to be bet in the described lottery. Within treatments, participants were randomly assigned either to sub-treatment H or sub-treatment L, which only differed in terms of decision and feedback frequency. In the H sub-treatment, participants in each of the nine rounds decided how much they wanted to bet in the lottery and were informed after each round about the lottery result drawn by the computer and the payoff for that round. In sub-treatment L, participants decided on their preferred bet in rounds 1, 4 and 7 for the three consecutive rounds. In this sub-treatment, the amount bet remained unchanged for three consecutive rounds. After three rounds, participants were informed about the results of the lottery for each of the three rounds and were notified about their total payoff for these three consecutive rounds (i.e., round 1–3 in round 3, round 4–6 in round 6, round 7–9 in round 9).

2.3 Implementation

Based on the variations described above, we obtained a 2×2 factorial experimental design. For both treatments, (BASELINE and DEBIASING), there are two sub-treatments, (H and L), which are implemented to examine the presence and magnitude of MLA in both treatments. To assess the success of the training intervention in reducing MLA-consistent behavior, we followed a difference-in-difference comparison of the H and L sub-treatments between the BASELINE and DEBIASING treatments.

We conducted online experiments with 894 student participants from the University of Innsbruck.[11] The average age of the participants was 24 years and 59% were female. The average payoff was EUR 4.97 (sd: EUR 1.52) across treatments for an expected processing time of approximately 25–30 minutes. The experimental online sessions took place between May and July 2020. The software was

[10] For details on the training intervention, see the screenshots in the Appendix A2 or the software.

[11] See the pre-registration for detailed power calculations. The experiments were conducted online due to the COVID-19 pandemic.

programmed using oTree (Chen et al., 2016) and participants were recruited via hroot (Bock et al., 2014). Participants received Amazon vouchers in the denomination of their experimental payoff as compensation. Screenshots of the English translation of the experiment are provided in Section A2 in the Appendix.[12]

3 Results

Result 1: *The decisionmakers in treatment* BASELINE *exhibit behavior that is consistent with MLA, which is not the case among decisionmakers in treatment* DEBIASING. *Overall, the average risk-taking of decisionmakers in* DEBIASING *is higher than that of participants in* BASELINE.

Figure 3 shows a comparison of the average lottery bet over nine rounds as a percentage of the endowment in stage 2 of the experiment between treatments BASELINE and DEBIASING and sub-treatments H and L, respectively.[13] In general, the visual impression suggests that participants in the L sub-treatment bet higher amounts in the lottery compared to the decisionmakers in H, suggesting the presence of behavior consistent with MLA. However, this difference seems to be slightly more pronounced within treatment BASELINE than within treatment DEBIASING. Furthermore, it is visible that general risk-taking, measured over both sub-treatments, seems to be higher among participants in treatment DEBIASING than in treatment BASELINE.

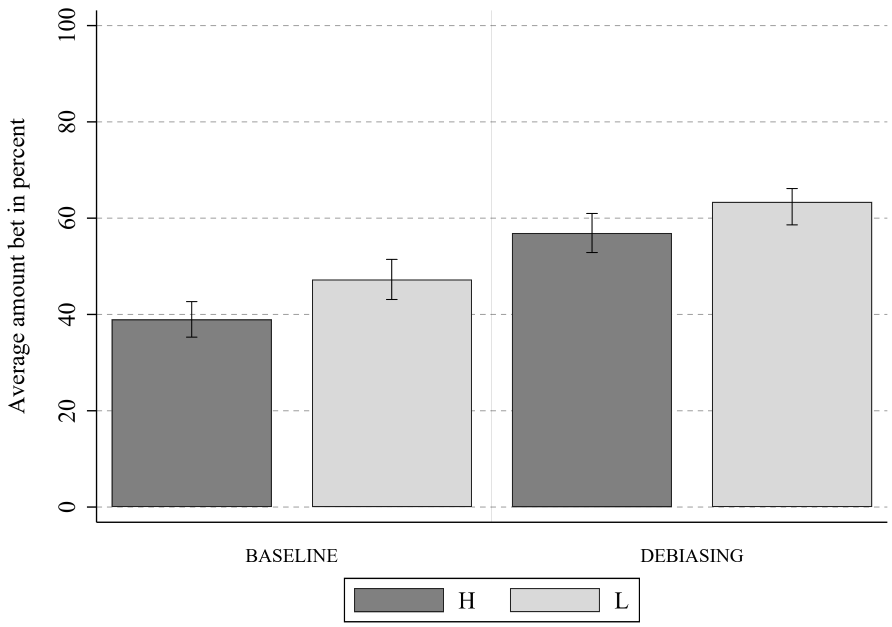


Figure 3: Average bet over nine rounds in percent of endowment. The graph shows the average amounts bet in the lottery over nine rounds as a percentage of the endowment of 200 talers for each treatment (BASELINE and DEBIASING) and sub-treatment (H and L). The whiskers represent 95% confidence intervals.

[12]The English version of the software can be found using the following link.

[13]For a comparison of the average lottery bets as a percentage of the endowment in stage 2 between treatments and sub-treatments in each of the nine specific rounds, see Figure A2 in the Appendix.

To test econometrically for the presence of MLA-consistent behavior in both treatments, we applied two-sided unpaired sample t-tests and reported the results in the upper half of Table 1. In the BASELINE treatment, we find a statistically significant difference in the average amount bet in the lottery as percentage of endowment between participants in L and H. Specifically, decisionmakers in sub-treatment L bet an average of 8.30 percentage points more in the lottery compared to their peers in the H sub-treatment (47.30% vs. 39.00%), a highly statistically significant difference ($p = 0.003$). Thus, we find MLA-consistent behavior in the BASELINE treatment.

Next, we analyzed the participants who underwent the training treatment, i.e., participants in treatment DEBIASING. As can be seen from the top half of Table 1, participants in the L sub-treatment do not bet statistically significantly higher fractions of their endowment compared to decisionmakers in the H sub-treatment (62.40% vs. 56.90%), although the difference approaches conventional levels of statistical significance ($p = 0.055$).

Table 1: Differences in treatments and sub-treatments. The table shows pairwise differences in the average bet amount over nine rounds in percent of the endowment between sub-treatments H and L in treatments BASELINE and DEBIASING using two-sided unpaired sample t-tests. The table also shows pairwise differences in the average bet amount over nine rounds in percent of the endowment between treatments BASELINE and DEBIASING in sub-treatments H and L, separately and jointly (H + L).

Treatments	obs	Sub-Ttreatment Difference: H-L		std. err.	pr(| T | > | t |)
BASELINE	439	-0.083***	(0.390 - 0.473)	0.028	0.003
DEBIASING	455	-0.055	(0.569 - 0.624)	0.028	0.054
Sub-Treatments	**obs**	**Treatment Difference: BASELINE-DEBIASING**		**std. err.**	**pr(| T | > | t |)**
H	431	-0.179***	(0.390 - 0.569)	0.278	0.000
L	463	-0.151***	(0.473 - 0.624)	0.285	0.000
H+L	894	-0.170***	(0.430 - 0.600)	0.020	0.000

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$

Subsequently, we tested for the overall difference in average risk inclination between BASELINE and DEBIASING (17 percentage points; 43.00% vs. 60.00%) measured over both sub-treatments and reported the result of a two-sided unpaired sample t-test in the lower half of Table 1. We find that the difference is highly statistically significant. In particular, decisionmakers in DEBIASING bet on average 17 percentage points more in the lottery compared to their peers in BASELINE.

This level effect is not surprising as participants in the training treatment would learn that on average betting higher amounts leads to a higher cumulative wealth after nine rounds when varying the amount x used to illustrate the lottery properties. This result is consistent with Kaufmann et al. (2013) who find that experience sampling increases risk-taking. The level effect does not seem to be caused by a lack of understanding of the higher standard deviation of wealth after nine rounds when betting higher amounts, as step 2 of the intervention illustrates the respective dispersion of the cumulative wealth after nine rounds in the lottery. Therefore, participants tend to behave more risk-neutrally in DEBIASING compared to BASELINE. This translates into different average payoffs between treatments. Participants in BASELINE earned an average of EUR 4.88 in stage 2 of the experiment and decisionmakers in the DEBIASING treatment earned an average of EUR 5.05. Further, this effect of increased risk-taking seems to be slightly more associated with participants in sub-treatment H (17.90 percentage points; 39.00% vs. 56.90%) than with participants in sub-treatment L (15.10 percentage points; 47.30% vs. 62.40%), which is observable from the bottom half of Table 1. In the next step, we tested for an interaction effect between treatments and sub-treatments as the mere presence of MLA-consistent behavior in treatment BASELINE but not in treatment DEBIASING is insufficient evidence of success of the training intervention. Thus, we performed multivariate Tobit regressions with the average percentage lottery bet over nine rounds as the dependent variable in Table 2 to test for a difference-in-difference treatment effect.

> **Result 2:** *Based on the difference-in-difference effect, we find no statistically significant difference in the degree of MLA between treatment* BASELINE *and treatment* DEBIASING.

As a robustness check, we tested for the general treatment effects in model (I) in Table 2. The coefficient DEBIASING is a binary dummy variable that takes the value of 0 for participants in treatment BASELINE and 1 for participants in treatment DEBIASING. L is a binary dummy variable that takes the value of 1 for decisionmakers in the low-frequency feedback sub-treatment, i.e., L, and 0 for participants in the high-frequency feedback group, i.e., H. DEBIASING#LOW_FREQUENCY(L) represents an interaction term between DEBIASING and L. AGE indicates the age of the participants in years, MALE is a binary dummy variable that takes the value of 0 for female participants and the value of 1 for male participants. STUDY_ECONOMICS is a binary variable that takes the value of 1 for participants enrolled in economics, business administration, or business law, and 0 for all other programs. INVESTMENT_EXPERIENCE is a dummy that takes the value of 1 for decisionmakers who have already invested in financial products and 0 for participants who have not yet done so. RISK_FINANCIAL is the ordinal scaled variable that represents self-reported risk preferences on a 10-point Likert scale in the financial domain and GRADUATE is a binary dummy taking a value of 1 for graduate students and 0 for undergraduate students. From the coefficient DEBIASING it can be inferred that participants who completed the training intervention in the first stage of the experiment take significantly higher risks in the lottery in the second stage of the experiment, confirming the visual impression presented in Figure 3 and Result 1. When aggregating both treatments, we can discern from the coefficient L that decisionmakers in the low frequency feedback sub-treatment (L) take significantly higher risks than participants in the high frequency feedback sub-treatment (H). This indicates the general presence of MLA-consistent behavior among the participants. Secondly, to test whether the training intervention influences the degree of MLA, we estimated the following specification in model (II) of Table 2:

$$y_i = \alpha + \beta_1 \text{DEBIASING}_i + \beta_2 \text{L}_i + \beta_3 \text{DEBIASING}_i \# \text{L}_i + \epsilon_i \tag{1}$$

Apparent by the coefficient DEBIASING#LOW_FREQUENCY(L), we find an expected negative sign, which indicates that the regression predicts the difference in risk-taking between participants in L and H to be lower in DEBIASING than in BASELINE. Nevertheless, the influence of the training intervention on the existing MLA-consistent behavior is not statistically significant. In model (III) we included the participants' financial risk preferences; their individual experience with financial investments; and their demographic and socio-economic characteristics. We find a statistically significant association between students' age and their lottery decision. In particular, older students bet slightly higher amounts.

Table 2: Multivariate Tobit regression on the average percentage amount invested over nine rounds in the lottery in stage 2 (PERCENT_BET). The variable DEBIASING is a binary dummy taking on the value 0 for participants in treatment BASELINE and 1 for participants in treatment DEBIASING. L represents a binary dummy variable taking the value 1 for decisionmakers in the low-frequency feedback sub-treatment and 0 for their peers in the high-frequency feedback group, i.e., H. DEBIASING#LOW_FREQUENCY(L) represents an interaction term between DEBIASING and L. The variable AGE indicates the participants' age in years, MALE is a binary dummy taking the value of 0 for female subjects and 1 for male participants. STUDY_ECONOMICS is a binary variable, which equals 1 for participants enrolled in economics, business, or business law and 0 for all other study programs. INVESTMENT_EXPERIENCE is a dummy taking the value of 1 for decisionmakers who had already invested in financial products and 0 for participants who had not. GRADUATE is a binary dummy taking a value of 1 for graduate students and 0 for undergraduate students. RISK_FINANCIAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the financial domain. "Permute *p*" reported the *p* values of the corresponding coefficient, obtained from permutation tests with 1,000 random draws.

	Model (I)	Model (II)	Model (III)
DEBIASING	0.185***	0.194***	0.205***
	(0.023)	(0.033)	(0.032)
LOW_FREQUENCY(L)	0.078***	0.087**	0.108***
	(0.023)	(0.033)	(0.031)
DEBIASING#LOW_FREQUENCY(L)		−0.018	−0.044
		(0.047)	(0.044)
AGE			0.006*
			(0.003)
MALE			0.125***
			(0.025)
GRADUATE			0.015
			(0.036)
STUDY_ECONOMICS			0.004
			(0.024)
RISK_FINANCIAL			0.030***
			(0.005)
INVESTMENT_EXPERIENCE			0.025
			(0.026)
Constant	0.402***	0.398***	−0.109
	(0.020)	(0.023)	(0.081)
Permute *p* DEBIASING#LOW_FREQUENCY(L)		0.728	0.346
Observations	894	894	894
Prob. > Chi^2	0.000	0.000	0.000
Pseudo R^2	0.080	0.080	0.196

$^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.005$. Dependent variable: Average amount bet over nine rounds in percent of endowment (PERCENT_BET); standard errors in parentheses.

Further, we find a statistically significant and large influence of gender on risk appetite. In particular, the regression predicts that male participants bet on average 12.50 percentage points more in the lottery than female participants. This is not surprising, as the literature on financial risk-taking shows that men prefer to take higher risks than women (Charness and Gneezy, 2012). Further, it is unsurprising that participants who described themselves as risk-seeking in financial matters bet statistically significantly higher amounts in the lottery (see coefficient RISK_FINANCIAL in model (III) of Table 2).

The impact of the training intervention seems to be estimated stronger in this specification. However, also in model (III), the coefficient DEBIASING#LOW_FREQUENCY(L) is not statistically significant. In addition, we applied randomization inference and performed permutation tests with models (II) and (III) in Table 2.[14] We tested the null hypothesis that there is no effect of the training intervention on behavior consistent with MLA by simulating 1,000 draws of differences in percentage amounts bet between H and L based on ex-post randomized treatment assignments in BASELINE and DEBIASING and recording the 1,000 interaction effects. The rarer the simulated interaction effects are greater than the actual interaction effect, the lower the permutation p values for the interaction term DEBIASING#LOW_FREQUENCY(L) (line "Permute p DEBIASING#LOW_FREQUENCY(L)" in Table 2). The lower these p values, the higher the probability $(1-p)$ that the actual treatment allocation caused the observed effect. This probability is clearly lower than 95% in both specifications, which confirms the statistical insignificance of the training intervention on MLA. We tested for multicollinearity by calculating variance inflation factors (VIFs), which indicates that multicollinearity is not a primary concern (the VIFs of all independent variables in model (III) are below 3.50). Consequently, the simulation and the parametric results indicate either a null effect or a lack of statistical power or data quality to detect an effect of the training intervention on MLA of the given magnitude.

Recent evidence on the replicability of social science experiments provides an estimate of the average relative effect size of true positives, which is around 71% (Camerer et al., 2018). Nevertheless, in contrast to the original study (Gneezy and Potters, 1997), we ran the experiment online; thus, having to differ slightly from the original instructions. Therefore, for this study, we applied an even more conservative approach. We based the power calculations of the interaction term on an expected difference in risk-taking between participants in L and H, amounting to about 67% of the original difference of 16.90 percentage points in Gneezy and Potters (1997). Consequently, we ensured a sufficient number of participants to guarantee 80% power to detect an 11.30 percentage point reduction in MLA-compliant behavior through the training intervention.[15] However, the actual difference between L and H (8.30 percentage points) measured in this study in the BASELINE treatment corresponds to only about 49% of the original effect size in Gneezy and Potters (1997).

Thus, we tested whether the statistical insignificance of the interaction is due to a lack of statistical power or data quality or whether the effect is virtually equivalent to zero. To do so, we performed an equivalence test (TOST regression) to the specifications in models (II) and (III) in Table 2.[16] We set a minimum relevant effect size of $\beta = \pm 0.083$. This is rather conservative, as this minimum relevant interaction effect size corresponds to the actual difference between sub-treatments L and H in the BASELINE treatment, which the intervention was intended to correct. Nevertheless, we cannot provide strong statistical support for the null hypothesis with respect to the coefficient

[14]We used the user-written program "ritest" in Stata (Heß, 2017).

[15]See the pre-registration for details.

[16]We used the user-written program "tostregress" in Stata (Dinno, 2017).

DEBIASING#LOW_FREQUENCY(L) (model (II): $p(T > t_1) = 0.003$, $p(T > t_2) = 0.085$; model (III): $p(T > t_1) < 0.005$, $p(T > t_2) = 0.196$). We conclude that we are statistically indeterminate and would need more data or better data quality to detect a difference or an equivalence with the null (Tryon and Lewis, 2008). In the next step, we collected the last point and applied an exploratory approach by checking whether the result is driven by inattentive participants and, thus, noisy data, as the experiment was conducted online due to the COVID-19 pandemic.

4 Possible Drivers

First, we analyzed the time each participant spent on the instruction screens.[17] Sufficient attention and seriousness is a prerequisite for successfully treating participants with the training intervention, as the intervention provided the relevant information only implicitly through experience sampling. We argue that participants who did not take enough time to read the instructions for stage 1 and stage 2 of the experiment could be a source of noise in the data. This could be a possible reason for the ambiguity regarding the hypothesis and the equivalence test. On the other hand, participants who spent an excessively long time on the instruction screens could also be a problem, and we are cautious in assuming that these participants performed the experiment with the necessary diligence and without being distracted. For example, 76 out of 894 participants (8.50%) spent a total of less than 2 minutes on both instruction screens in stage 1 and stage 2 and 153 out of 894 participants (17.11%) spent a total of more than 1 hour on these screens. To get an impression of the required time to sensibly read the instructions, participants had to read on average 1,196 words in total over all treatment and sub-treatment combinations on both instruction screens. This should take a native German speaker around 6.68 minutes on average (Trauzettel-Klosinski and Dietz, 2012).

We followed Downs et al. (2010) who find that the exclusion of participants in the top decile of processing times in an Mturk sample statistically significantly distinguishes attentive from non-attentive participants. Although the authors indicate that the prediction quality of this cut-off point is far from perfect, it still provided us with a validated reference for our data cleaning process. Furthermore, Downs et al. (2010) argue that unmotivated and inattentive participants might not always click quickly, but rather act distracted and simultaneously do something else. Combined with the detection of disproportionately long processing times in the data, we decided to trim the sample symmetrically by excluding participants with the 10% shortest and 10% longest processing times on the task relevant instruction screens in stage 1 and stage 2 from the analyses. This left us with a total of 716 observations. From now on, we refer to this sample as the "high attentives". In a next step, we repeated our analyses from Section 3.

> **Result 3:** *When excluding the participants with the 10% shortest and longest processing times on the instruction screens from the analyses, we find a stronger corrective effect of the training intervention on MLA-consistent behavior, which is statistically significant when controlled for age, gender, education, field of study, investment experience, and financial risk preferences.*

Figure A1 in the Appendix again shows the comparison of the average lottery bet in percent of

[17] Because of considerations between data quality and statistical power, we did not include the processing times on the screens that were not directly relevant to the main tasks in both stages of the experiment in the data quality checks. For the main tasks themselves, we implemented minimum time requirements or a minimum number of iterations, as discussed previously.

the endowment in stage 2 of the experiment between treatments BASELINE and DEBIASING and sub-treatments H and L. The overall patterns remain similar to Figure 3, representing the full sample. Nevertheless, among the more attentive participants, the difference in risk appetite between L and H in the DEBIASING treatment is greatly reduced compared to the BASELINE treatment.

To test statistically for the presence of MLA-consistent behavior under both treatments among the more attentive participants, we again applied two-sided unpaired sample t-tests and reported the results in the upper half of Table 3. In the BASELINE treatment, we find even more statistically significant evidence for MLA-consistent behavior than in the full sample. Here, decisionmakers in sub-treatment L bet on average 11.90 percentage points more in the lottery than decisionmakers in sub-treatment H (48.40% vs. 36.50%). As also shown in the upper half of Table 3, participants in treatment DEBIASING and sub-treatment L do not bet statistically significantly higher amounts compared to their counterparts in the H sub-treatment (62.30% vs. 57.90%, $p = 0.158$), which constitutes an even more insignificant difference compared to the full sample; thus, clearly indicating no evidence of MLA-consistent behavior.

In addition, to test for the overall difference in average risk appetite between BASELINE and DEBIASING among the more attentive participants, we repeated the previous analyses and reported the results in the lower half of Table 3. Overall, we find a similarly large and as large a statistically significant difference as in the full sample. Decisionmakers in the DEBIASING treatment bet on average 17.50 percentage points more in the lottery compared to participants in the BASELINE treatment (60.20% vs. 42.70%). Among the more attentive participants, the positive effect of the training intervention on risk-taking is clearly more strongly associated with participants in sub-treatment H (21.46 percentage points; 36.45% vs. 57.91%) than with participants in sub-treatment L (13.86 percentage points; 48.43% vs. 62.29%), which is observable from the bottom half of Table 3. This is an expected result of the training intervention and suggests a statistically significant difference-in-difference effect.

Table 3: Differences in treatments and sub-treatments among more attentive participants. The table shows pairwise differences in the average bet amount over nine rounds in percent of the endowment between sub-treatments H and L in treatments BASELINE and DEBIASING using two-sided unpaired sample t-tests. The table also shows pairwise differences in the average bet amount over nine rounds in percent of the endowment between treatments BASELINE and DEBIASING in sub-treatments H and L, separately and jointly (H + L).

Treatments	obs	Sub-Ttreatment Difference: H-L		std. err.	pr(| T | > | t |)
BASELINE	363	-0.119***	(0.365 - 0.484)	0.031	0.000
DEBIASING	353	-0.044	(0.579 - 0.623)	0.031	0.158
Sub-Treatment	**obs**	**Treatment Difference: BASELINE-DEBIASING**		**std. err.**	**pr(| T | > | t |)**
H	340	-0.215***	(0.365 - 0.579)	0.031	0.000
L	376	-0.139***	(0.484 - 0.623)	0.031	0.000
H + L	716	-0.175***	(0.427 - 0.602)	0.022	0.000

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.005$

To test for the difference-in-difference effect among the more attentive participants, we repeated the multivariate Tobit regression analyses from Table 2 with the sample of high attentives. Model (I) in Table 4 yields results regarding the coefficients DEBIASING and L that are consistent with the results in Table 2, (the results with the full sample). We tested whether the training intervention has a statistically significant effect on MLA and again estimated the specification in equation 1 in model (II) of Table 4. As can be seen from the coefficient DEBIASING#LOW_FREQUENCY(L), there is a stronger estimated MLA reducing effect of the training intervention than in the full sample. However, the effect is not statistically significant. When adding in the participants' reported financial risk preferences; their individual experience with financial investments; and their demographic and socio-economic characteristics, we find a statistically significant, corrective effect of the training intervention on behavior consistent with MLA, even though we lost statistical power when we trimmed the sample. Specifically, the regression predicts that the training intervention reduces the H vs. L difference in risk-taking by about 10 percentage points in treatment DEBIASING compared to treatment BASELINE. Additionally, this significant parametric result is confirmed by the results of randomization inference (line "Permute *p* DEBIASING#LOW_FREQUENCY(L)" in Table 4).[18] Interestingly, the main difference between the sample of more attentive participants and the full sample appears to be the difference-in-difference effect, but not the level effect in overall risk appetite between treatments, as the difference in overall risk appetite between BASELINE and DEBIASING amounts to 17.00 percentage points in the full sample and 17.50 percentage points in the sample with more attentive participants. We conclude that a fairly high level of attention is required to successfully correct MLA-consistent behavior through the developed training tool as participants who did not read the instructions carefully and with focus or were distracted might be a source of noisy data.

Furthermore, we conducted a detailed analysis of the impact of percentile cut-off points on processing times other than 10% on the results. Specifically, we calculated corresponding effect sizes and *p* values of the variable DEBIASING#LOW_FREQUENCY(L), which represents the difference-in-difference effect, for all symmetric percentile cut-off points starting with 99/1 and ending with 55/45. As can be seen from Figure A3 and Figure A4 in the Appendix, we find that symmetrically trimming the sample based on the processing times clearly has a positive effect on the strength of the effect size. Additionally, we find a U-shaped relationship between the corresponding *p* values and the cut-off points.[19] In summary, this suggests that the associated results in model (III) in Table 4 are not limited to the specific cut-off point of 10%, and that there are structural differences in behavior between more and less attentive participants.

[18] We re-tested for multicollinearity by considering variance inflation factors (VIFs) that suggest that multicollinearity is not a primary concern (the VIFs of all independent variables in model (III) are again below 3.50).

[19] Consequently, trimming the sample increases the effect sizes, which, *ceteris paribus*, would reduce the *p* values. This is, however, simultaneously accompanied by a loss of statistical power, and this seems to counteract the *p* value-lowering effect of increasing effect sizes at some point.

Table 4: Multivariate Tobit regression on the average percentage amount invested over nine rounds in the lottery in stage 2 (PERCENT_BET) among the high attentives. The variable DEBIASING is a binary dummy taking on the value 0 for participants in treatment BASELINE and 1 for participants in treatment DEBIASING. L represents a binary dummy variable taking the value 1 for decisionmakers in the low-frequency feedback sub-treatment and 0 for their peers in the high-frequency feedback group, i.e., H. DEBIASING#LOW_FREQUENCY(L) represents an interaction term between DEBIASING and L. AGE indicates the participants' age in years, MALE is a binary dummy taking the value of 0 for female subjects and 1 for male participants. STUDY_ECONOMICS is a binary variable, which equals 1 for participants enrolled in economics, business, and business law and 0 for all other study programs. INVESTMENT_EXPERIENCE is a dummy taking the value of 1 for decisionmakers who had already invested in financial products and 0 for participants who had not. GRADUATE is a binary dummy taking a value of 1 for graduate students and 0 for undergraduate students. RISK_FINANCIAL is an ordinal variable representing self-reported risk preferences on a 10-point Likert scale in the financial domain. "Permute *p*" reports the *p* values of the corresponding coefficient, obtained from permutation tests with 1,000 random draws.

	Model (I)	Model (II)	Model (III)
DEBIASING	0.194***	0.231***	0.249***
	(0.025)	(0.036)	(0.034)
LOW_FREQUENCY(L)	0.094***	0.128***	0.145***
	(0.025)	(0.035)	(0.033)
DEBIASING#LOW_FREQUENCY(L)		−0.071	−0.101*
		(0.051)	(0.047)
MALE			0.129***
			(0.027)
AGE			0.006*
			(0.003)
GRADUATE			0.013
			(0.040)
STUDY_ECONOMICS			0.024
			(0.026)
RISK_FINANCIAL			0.030***
			(0.005)
INVESTMENT_EXPERIENCE			0.007
			(0.028)
Constant	0.388***	0.370***	−0.135
	(0.022)	(0.025)	(0.087)
Permute *p* DEBIASING#LOW_FREQUENCY(L)		0.177	0.047
Observations	716	716	716
Prob. > Chi^2	0.000	0.000	0.000
Pseudo R^2	0.097	0.100	0.245

*$p < 0.05$, **$p < 0.01$, ***$p < 0.005$. Dependent variable: Average amount bet over nine rounds in percent of endowment (PERCENT_BET); Standard errors in parentheses.

5 Conclusion

In this paper, we presented a novel tool to reduce or eliminate behavior consonant with the theory of myopic loss aversion (MLA) in a training intervention. Specifically, we conducted a large-scale online experiment with 894 student participants which consisted of two main stages. In the first stage, participants in the debiasing treatment underwent the training intervention. We used experience sampling as well as graphic and numerical illustrations as a means of communication of lottery results to convey the basic characteristics of the lottery originally introduced in Gneezy and Potters (1997). In the baseline treatment, participants played the game Minesweeper as a filler task. In the second experimental stage, in which treatments did not differ, the susceptibility of participants to MLA was determined. We found behavior consistent with MLA in the baseline treatment, whereas we did not find behavior consistent with MLA in the debiasing training treatment. Nonetheless, we found no statistically significant difference-in-difference effect of the training intervention on participants' susceptibility to MLA. This result was also supported by randomization inference. Nevertheless, we found that the training intervention increased overall risk-taking among participants.

In an exploratory approach, we analyzed whether the (in)attention of participants was driving the result. We found that a considerable number of participants spent a disproportionately long or short period of time on the instruction screens. Therefore, we trimmed the sample and excluded participants with the 10% shortest and 10% longest processing times on the instruction screens from the analyses. Based on this sample of more attentive participants, we found a statistically significant effect of the training intervention on the susceptibility of participants to MLA when we controlled for age, gender, field of study, education investment experience, and financial risk preferences. We can conclude that experience sampling with the help of graphical and numerical representations corrects behavior consistent with MLA only in relatively attentive participants. We consider these findings to be important because, especially in some European countries, a shift from public pension savings to private pension savings might be foreseeable and the relevance of investments in financial products with higher volatility for private individuals might increase due to the minimum/negative interest rate policy in the most important financial markets. The presented debiasing tool can play a role in improving the quality of people's decisions.

The results in this paper also emphasize the importance of shrewd attention and a lack of long interruptions among participants in training interventions to ensure a full understanding of the implications conveyed by the intervention. Future research should rely on highly controlled environments when training interventions are conducted that are tailored to mitigate cognitive biases, i.e., laboratory settings where high attention and absence of distraction are guaranteed. Future research on the effectiveness of training interventions in reducing MLA-consistent behavior could include experiments with different pools of participants. In particular, financial professionals, as well as individuals from the general population, could be invited to participate in the experiments. Another worthwhile research approach would be to investigate the effect of the training intervention when it is separated in time from the measurement of MLA. Ideally, the training intervention has long-term alleviating effects on the susceptibility of participants to MLA.

References

Abdellaoui, Mohammed, Han Bleichrodt, Hilda Kammoun. 2013. Do financial professionals behave according to prospect theory? An experimental study. *Theory and Decision* **74**(3) 411–429.

Abdellaoui, Mohammed, Han Bleichrodt, Corina Paraschiv. 2007. Loss aversion under prospect theory: A parameter-free measurement. *Management Science* **53**(10) 1659–1674.

Bellemare, Charles, Michaela Krause, Sabine Kröger, Chendi Zhang. 2005. Myopic loss aversion: Information feedback vs. investment flexibility. *Economics Letters* **87**(3) 319–324.

Benartzi, Shlomo, Richard H Thaler. 1995. Myopic loss aversion and the equity premium puzzle. *The quarterly journal of Economics* **110**(1) 73–92.

Bock, Olaf, Ingmar Baetge, Andreas Nicklisch. 2014. hroot: Hamburg registration and organization online tool. *European Economic Review* **71** 117–120.

Bodie, Zvi, Robert C. Merton, William F. Samuelson. 1992. Labor supply flexibility and portfolio choice in a life cycle model. *Journal of Economic Dynamics and Control* **16**(3-4) 427–449.

Bradbury, Meike A.S., Thorsten Hens, Stefan Zeisberger. 2019. How persistent are the effects of experience sampling on investor behavior? *Journal of Banking & Finance* **98** 61–79.

Caligaris, Marta, Georgina Rodríguez, Lorena Laugero. 2015. Learning styles and visualization in numerical analysis. *Procedia - Social and Behavioral Sciences* **174** 3696 – 3701.

Camerer, Colin F., Anna Dreber, Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A. Nosek, Thomas Pfeiffer, Adam Altmejd, Nick Buttrick, Taizan Chan, Yiling Chen, Eskil Forsell, Anup Gampa, Emma Heikensten, Lily Hummer, Taisuke Imai, Siri Isaksson, Dylan Manfredi, Julia Rose, Eric-Jan Wagenmakers, Hang Wu. 2018. Evaluating the replicability of social science experiments in nature and science between 2010 and 2015. *Nature Human Behaviour* **2** 637–644.

Cason, N., Timothy, Anya Samek. 2015. Learning through passive participation in asset market bubbles. *Journal of the Economic Science Association* **1** 170–181.

Charness, Gary, Thomas Garcia, Theo Offerman, Marie Claire Villeval. 2019. Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *GATE Working Paper* URL https://ideas.repec.org/p/gat/wpaper/1921.html.

Charness, Gary, Uri Gneezy. 2012. Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization* **83**(1) 50–58.

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

Christensen, Darren R., Warren K. Bickel. 2010. Temporal myopia. *In: Stolerman I.P. (eds) Encyclopedia of Psychopharmacology. Springer, Berlin, Heidelberg* URL https://doi.org/10.1007/978-3-540-68706-1_498.

Cipriani, Marco, Antonio Guarino. 2009. Herd behavior in financial markets: An experiment with financial market professionals. *Journal of the European Economic Association* **7**(1) 206–233.

Deaves, Richard, Erik Lüders, Michael Schröder. 2010. The dynamics of overconfidence: Evidence from stock market forecasters. *Journal of Economic Behavior & Organization* **75** 793–808.

Dinno, Alexis. 2017. tostregress: Linear regression tests for equivalence. stata software package URL https://www.alexisdinno.com/stata/tost.html.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, Gert G Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association* **9**(3) 522–550.

Downs, Julie S., Mandy B. Holbrook, Steve Sheng, Lorrie Faith Cranor. 2010. Are your participants gaming the system?: screening mechanical turk workers. *Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI* 2399–2402.

Fellner, Gerlinde, Matthias Sutter. 2009. Causes, consequences, and cures of myopic loss aversion–an experimental investigation. *The Economic Journal* **119**(537) 900–916.

Fleming, Neil D., Colleen Mills. 1992. Not another inventory, rather a catalyst for reflection. *To Improve the Academy, the Professional and Organizational Development Network in Higher Education* **11**.

Fong, Geoffrey T., Richard E. Nisbett. 1991. Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology* **120**(1) 33–45.

Gächter, Simon, Eric J. Johnson, Andreas Herrmann. 2007. Individual-level loss aversion in riskless and risky choices. *IZA Discussion Paper* (2961).

Gneezy, Uri, Arie Kapteyn, Jan Potters. 2003. Evaluation periods and asset prices in a market experiment. *The Journal of Finance* **58**(2) 821–837.

Gneezy, Uri, Jan Potters. 1997. An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics* **112**(2) 631–645.

Grosshans, Daniel, Stefan Zeisberger. 2018. All's well that ends well? on the importance of how returns are achieved. *Journal of Banking and Finance* **87** 397–410.

Haigh, Michael S., John A. List. 2005. Do professional traders exhibit myopic loss aversion? An experimental analysis. *Journal of Finance* **60**(1) 523–534.

Heß, Simon. 2017. Randomization inference with stata: A guide and software. *The Stata Journal* **17**(3) 630–651.

Huber, Christoph, Jürgen Huber, Laura Hueber. 2019. The effect of experts' and laypeople's forecasts on others' stock market forecasts. *Journal of Banking & Finance* **109** 105662.

Hülsewig, Anja, Oliver Hülsewig. 2017. Das österreichische Rentensystem im Blickpunkt: Rentenparadies oder eine Belastung für zukünftige Generationen? *ifo Schnelldienst* **70**(07) 31–39.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, Alan M. Taylor. 2018. The rate of return on everything, 1870-2015. *CESifo Working Paper Series* (6899).

Kahneman, Daniel, Jack Knetsch, Richard H. Thaler. 1990. Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy* **98** 1325–1348.

Kahneman, Daniel, Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* **47** 263–291.

Kahneman, Daniel, Amos Tversky. 1984. Choices, values, and frames. *American Psychologist* **39**(4) 341–350.

Kaufmann, Christine, Martin Weber, Emily Celia Haisley. 2013. The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science* **59**(2) 323–340.

Kaustia, Markku, Milla Perttula. 2012. Overconfidence and debiasing in the financial industry. *Review of Behavioral Finance* **4**(1) 46–62.

Kirchler, Michael, Florian Lindner, Utz Weitzel. 2018. Rankings and risk-taking in the finance industry. *The Journal of Finance* **73**(5) 2271–2302.

Kučera, Tomáš. 2020. Cognitive bias mitigation: How to make decision-making rational? *IES Working Papers, IES FSV. Charles University* **1/2020**.

Larson, Francis, John A. List, Robert D. Metcalfe. 2012. Can myopic loss aversion explain the equity premium puzzle? evidence from a natural field experiment with professional traders. *NBER WORKING PAPER SERIES* (22605).

Lee, Boram, Yulia Veld-Merkoulova. 2016. Myopic loss aversion and stock investments: An empirical study of private investors. *Journal of Banking and Finance* **70** 235–246.

Looney, Clayton Arlen, Andrew M Hardin. 2009. Decision support for retirement portfolio management: Overcoming myopic loss aversion via technology design. *Management Science* **55**(10) 1688–1703.

Lusardi, Annamaria, Anya Samek, Arie Kapteyn, Lewis Glinert, Angela Hung, Aileen Heinberg. 2017. Visual tools and narratives: new ways to improve financial literacy. *Journal of Pension Economics Finance* **16**(3) 297 – 323.

Mehra, Rajnish, Edward C. Prescott. 1985. The equity premium a puzzle. *Journal of Monetary Economics* **15**(2) 145–161.

Menkhoff, Lukas, Maik Schmeling. 2013. Are all professional investors sophisticated? *German Economic Review* **11**(4) 418–440.

Morewedge, Carey K., Haewon Yoon, Irene Scopelliti, Carl W. Symborski, James H. Korris, Karim S. Kassam. 2015. Debiasing decisions: Improved decision making with a single training intervention. *Behavioral and Brain Sciences* **2**(1) 129–140.

Morrison, William, Robert Oxoby. 2014. Loss aversion in the laboratory. *IZA Discussion Paper* (8309).

Muradoglu, Gulnur, Nigel Harvey. 2012. Behavioural finance: the role of psychological factors in financial decisions. *Review of Behavioral Finance* **4**(2) 68–80.

Northwestern Mutual. 2019. 2019 planning progress study. `https://news.northwesternmutual.com/planning-and-progress-2019`. Accessed: 2020-12-18.

Pikulina, Elena, Luc Renneboog, Philippe N. Tobler. 2017. Overconfidence and investment: An experimental approach. *Journal of Corporate Finance* **43** 175–192.

Roszkowski, Michael J., Glenn E. Snelbecker. 1990. Effects of "framing" on measures of risk tolerance: Financial planners are not immune. *Journal of Behavioral Economics* **19**(3) 237–246.

Samuelson, William, Richard Zeckhauser. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* **1**.

Schooley, Diane K., Debra Drecnik Worden. 2013. Accumulating and spending retirement assets: A behavioral finance explanation. *Financial Services Review* **22** 173–186.

Schwaiger, Rene, Michael Kirchler, Florian Lindner, Utz Weitzel. 2020. Determinants of investor expectations and satisfaction. a study with financial professionals. *Journal of Economic Dynamics and Control* **110**(103675).

Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave. 2018. Nonrepresentative representatives: An experimental study of the decision making of elected politicians. *American Political Science Review* **112**(2) 302–321.

Sutter, Matthias. 2007. Are teams prone to myopic loss aversion? an experimental study on individual versus team investment behavior. *Economics Letters* **97**(2) 128–132.

Svenson, Ola. 1981. Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica* **47**(2) 143–148.

Thaler, Richard H., Amos Tversky, Daniel Kahneman, Alan Schwartz. 1997. The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics* **112**(2) 647–661.

Trauzettel-Klosinski, Susanne, Klaus Dietz. 2012. Standardized assessment of reading performance: The new international reading speed texts irest. *Investigative Ophthalmology Visual Science* **53**(9).

Tryon, Warren W., Charles Lewis. 2008. An inferential confidence interval method of establishing statistical equivalence that corrects tryon's (2001) reduction factor. *Psychological Methods* **13**(3) 272–277.

Tversky, Amos, Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* **185** 1124–1131.

U.S. Department of Labor. 2014. Private pension plan bulletin. http://op.bna.com.s3.amazonaws.com/pen.nsf/r%3FOpen%3dsfos-9ppmh6. Accessed: 2020-02-25.

Wendy, Wendy, Marwan Asri. 2012. Psychological biases in investment decisions: An experimental study of myopic behavior in developing capital markets. *Journal of Indonesian Economy and Business* **27**(2) 143–158.

Appendix

A1 Additional figures and tables

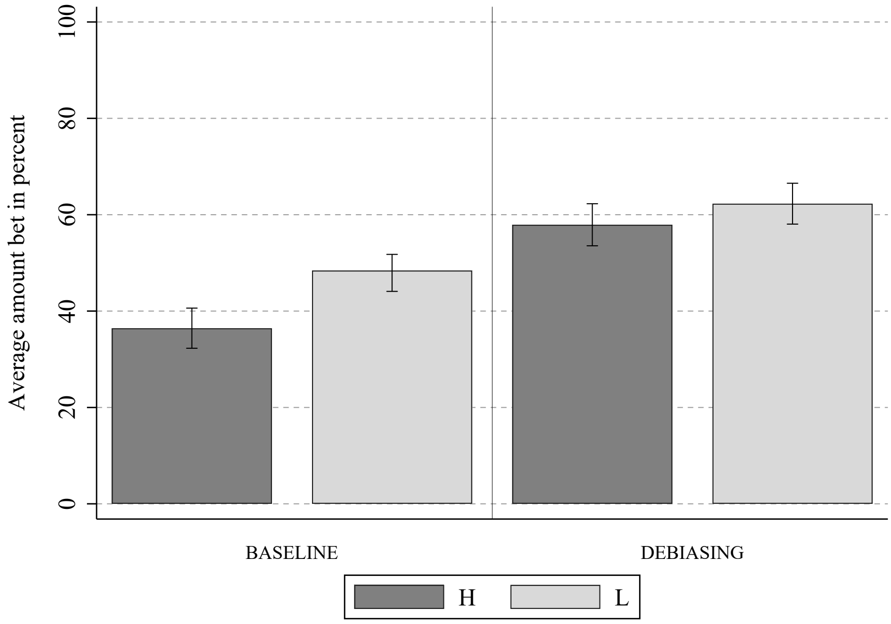


Figure A1: Average amount bet in percent of endowment (high attentives). The graph shows the average amounts bet in the lottery over nine rounds as a percentage of the endowment of 200 talers for each treatment (BASELINE and DEBIASING) and sub-treatment (H and L)and sub-treatment. The whiskers represent 95% confidence intervals.

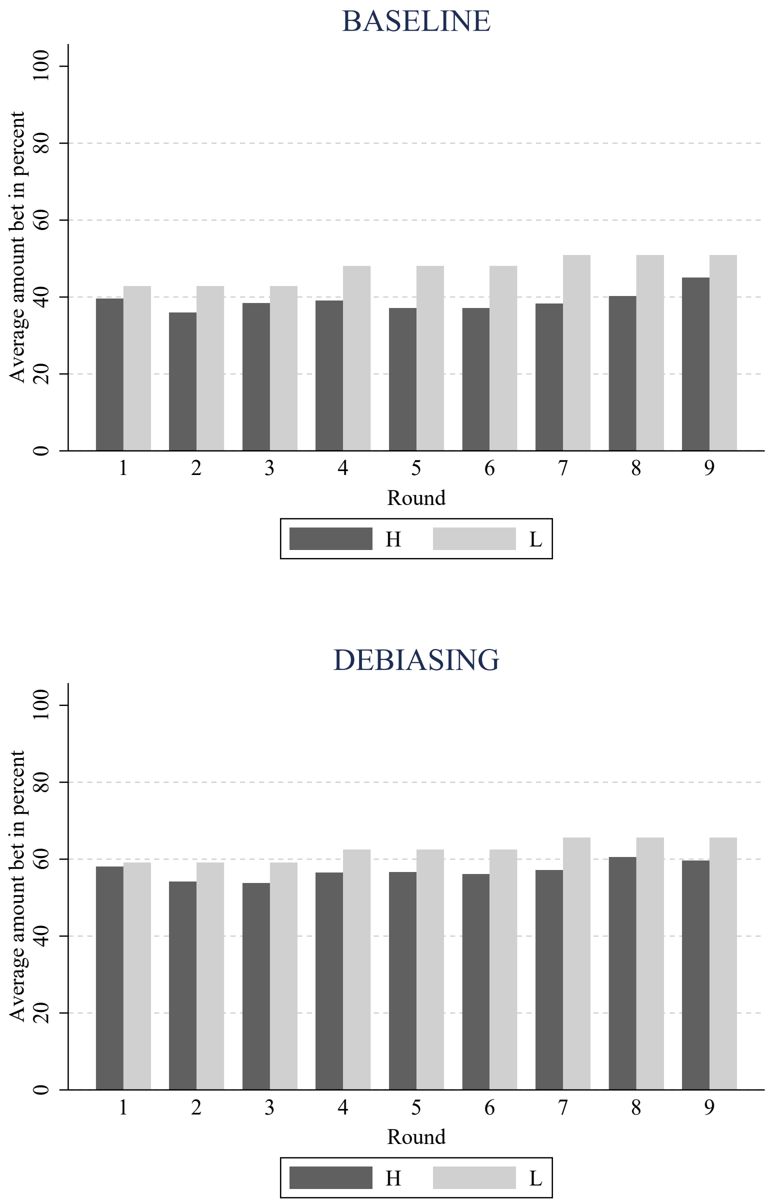


Figure A2: Percentage amount bet in each round. The graph shows the average amounts bet in the lottery in each of the nine rounds as a percentage of the endowment of 200 talers in sub-treatments H and L. The upper graph displays data for treatment BASELINE and the lower graph shows data for treatment DEBIASING.

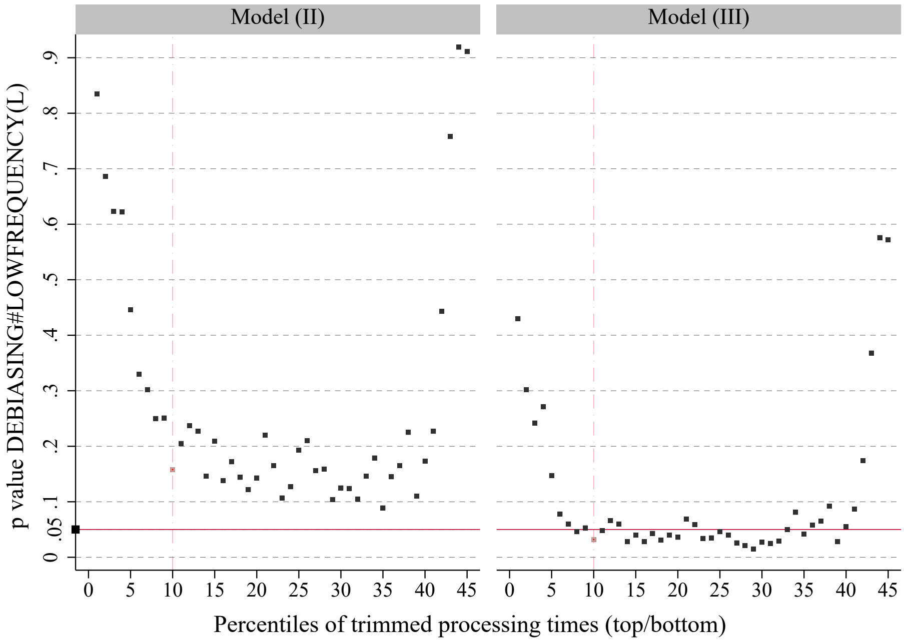


Figure A3: *p* values for different percentiles of trimmed processing times (top/bottom). The graph shows the relationship between the *p* value of the coefficient DEBIASING#LOW_FREQUENCY(L) and different cut-off points. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times on the two task relevant instruction screens in both treatments BASELINE and DEBIASING. Model (II) and model (III) show the respective results for the according specifications in Table 2 and 4.

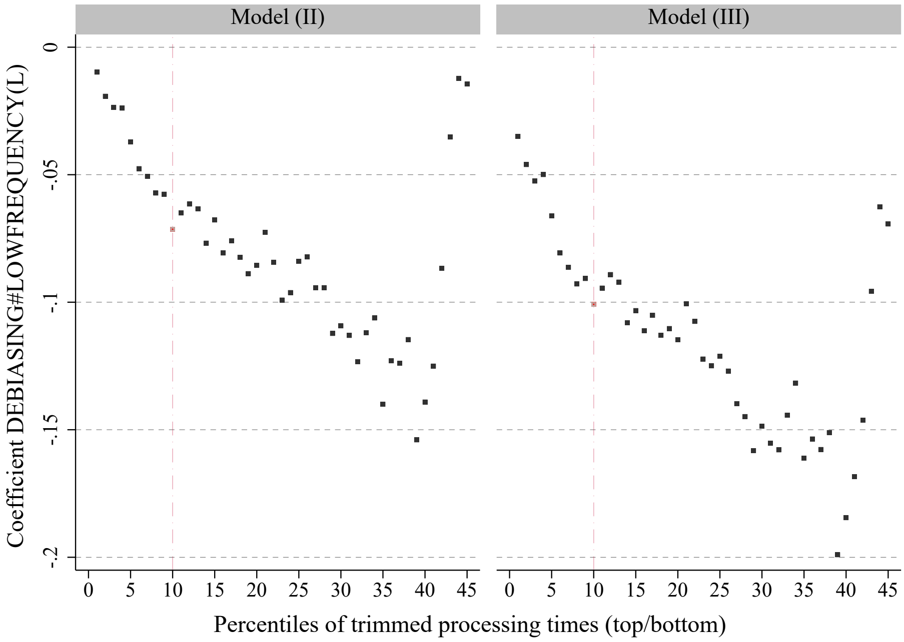


Figure A4: Coefficients for different percentiles of trimmed processing times (top/bottom). The graph shows the relationship between the coefficient DEBIASING#LOW_FREQUENCY(L) and different cut-off points. For example, for the 20th percentile, we excluded the participants with the 20% slowest and 20% fastest processing times on the two task relevant instruction screens in both treatments BASELINE and DEBIASING. Model (II) and model (III) show the respective results for the according specifications in Table 2 and 4.

A2 Screenshots of the software

A2.1 Intro

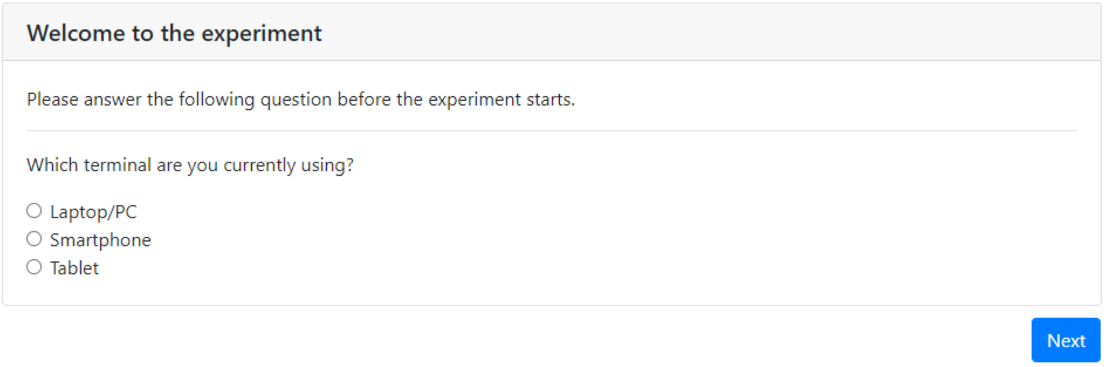


Figure A1: Question on terminal device

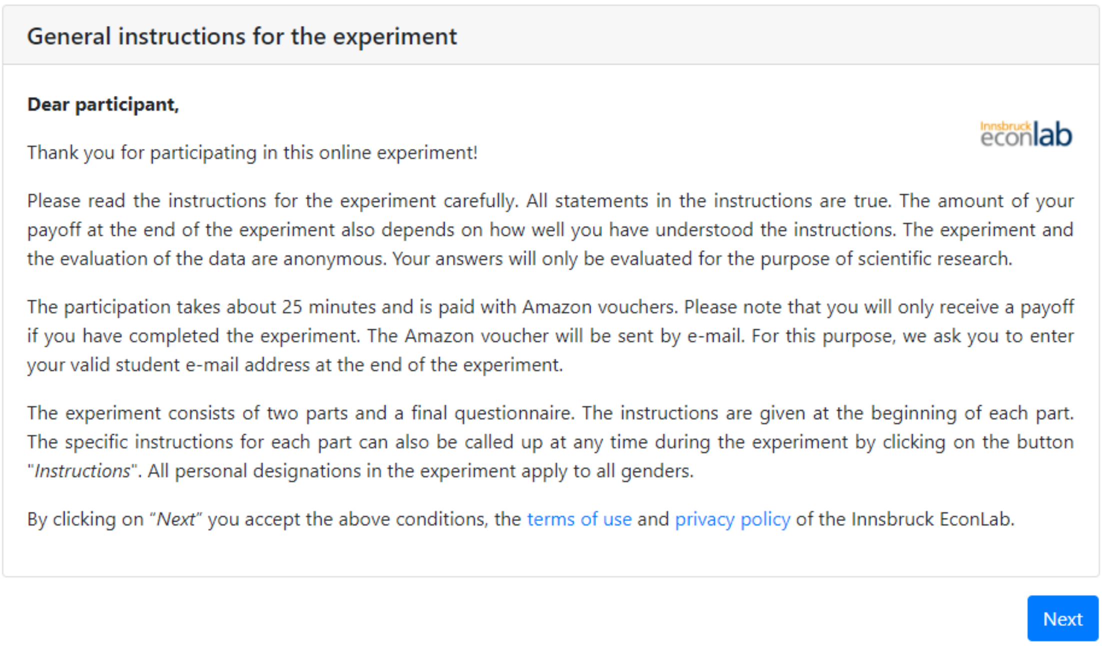


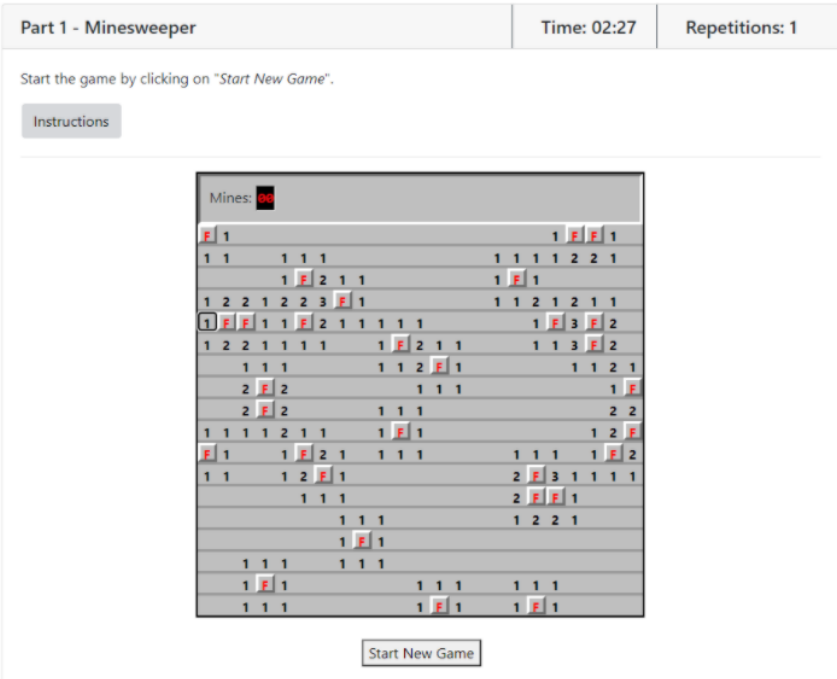
Figure A2: General Instructions

A2.2 Treatment: BASELINE

Instructions for Part 1

In Part 1 of the experiment you will play the game "Minesweeper". Note that your success in this game will **not** affect your payout. The game starts when you click on the button "*Start New Game*".
Your task is to mark all 30 fields under which there is a mine. A game ends when you uncover a field under which a mine is located or when you correctly mark all fields under which a mine is located. If you assume that a mine is hidden under a field, you can click on it with the **right mouse button** to mark the field with a flag. By right-clicking on the field again, you can unmark it.
Analyze the patterns. For example, if a 2 is displayed in a field, then there is a mine under 2 of the 8 adjacent fields. For example, if a field displays the number 8, then there is a mine under each of the adjacent 8 fields. If you do not know which field to click on, you should enter unexplored territory. It is better to click in the middle of an area of unmarked fields than in an area where you think there might be mines. During the game, you will be shown in the upper left corner how many mines are left according to your marking.
If you have solved a minefield or the game has ended because you uncovered a field with a mine, please click again on the button "*Start New Game*" to start a new game.

Regardless of your results in this game, you can continue the experiment by clicking on the "*Next*" button, which will be displayed after 10 minutes and at least 5 repetitions of the game.
Below you can see an example of the game with a solved minefield. You can trace the marking logic by looking at the individual numbers next to the fields.



Next

Figure A3: Instructions Filler Task

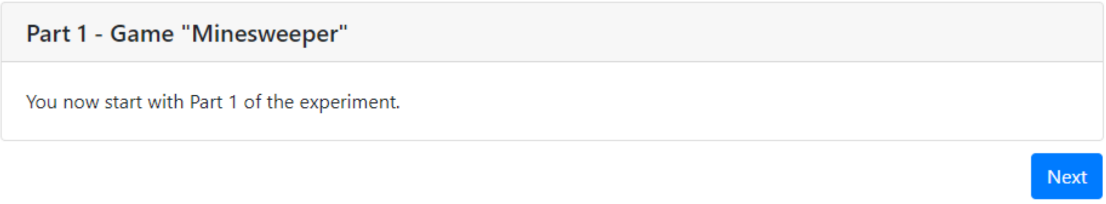


Figure A4: Transition Filler Task

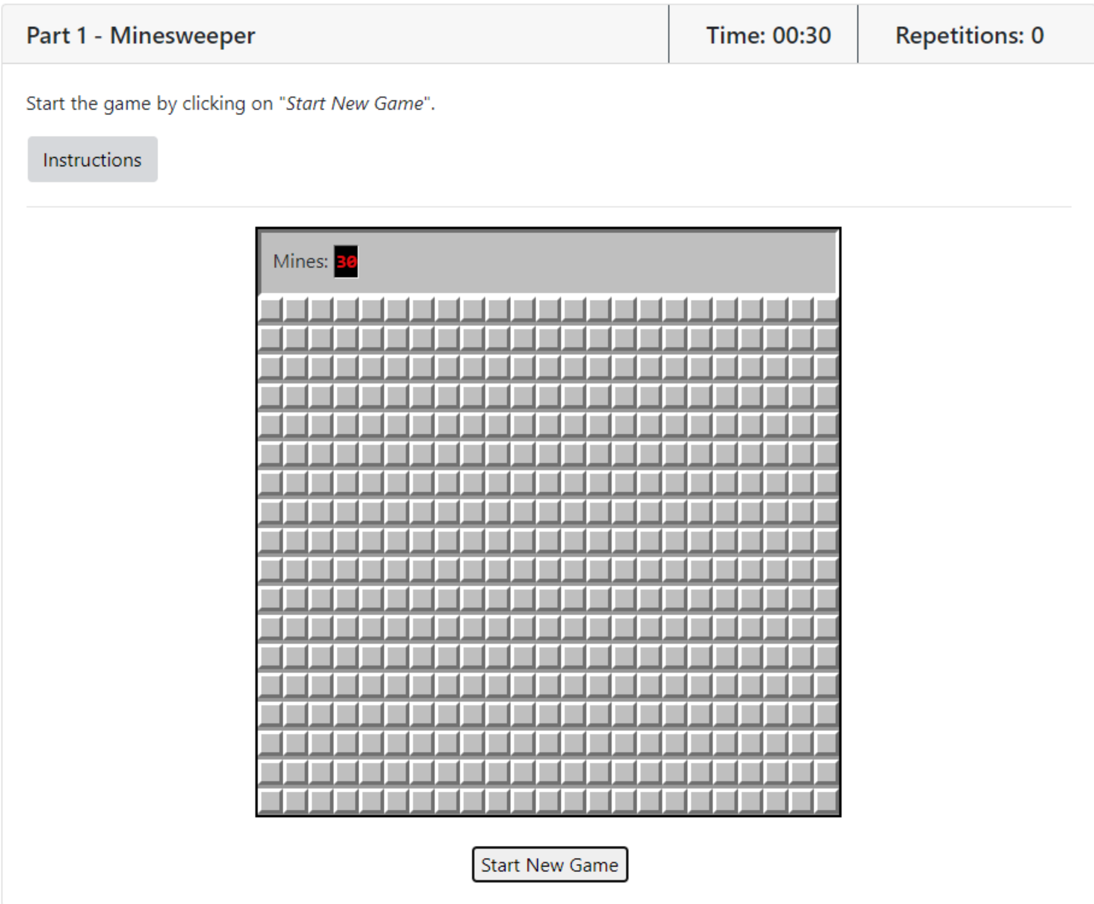


Figure A5: Filler Task

A2.3 Treatment: DEBIASING

Figure A6: General Instructions Intervention

General | Simulation A | Simulation B

Instructions for Simulation A

How do you run Simulation A?

First you are asked to enter a fictitious amount "*x*" in Taler between 0 and 200, which is to be used to simulate the lottery over 9 rounds (draws). Then click on the button "*Confirm Amount*" to confirm the entry.
Then you can choose how you want to simulate the lottery over 9 rounds (drawings). If you want to simulate the lottery automatically in one go, please click on the button "*Continuous*". If you prefer to simulate the lottery manually round by round (or draw by draw), please click on the button "*Step by step*". In both cases, it will be drawn 9 times from the lottery for each simulation.
After 9 rounds (draws), which corresponds to a one-time execution of the simulation, you will be asked again to enter a fictitious amount "*x*" in Taler between 0 and 200 with which the simulation of the lottery is to be performed and the process starts again. You must execute simulation A 15 times to continue the experiment.

What does Simulation A show?

You will see a graphic with two bars. Each of these two bars shows the accumulated total wealth in each of the 9 rounds, based on **one of the two previously described scenarios** (see instructions "*General*"). As can be seen in the example screen of Simulation A at the bottom of this page, the left bar represents the scenario "*Lottery bet not reduced*" and the right bar represents the scenario "*lottery bet reduced*".

If you click on one of the bars, you will see the total wealth from each round in that scenario. To return to the original chart, click on the button "*Back to the wealth comparison*".

In addition to the chart, you will also see a table showing the total wealth in each round and scenario. At the end of the 9 rounds, this table will show you the total wealth after 9 rounds, which is the sum of the total wealth of each round and therefore corresponds to the height of the bars. In addition, you can see in each round whether there was a win or a loss.

Below is an example of simulation A.

Part 1 - Simulation A: Illustration of the Lottery over 9 rounds | **Executions: 1**

Please indicate which amount between 0 and 200 should be used to illustrate the lottery.

Taler | Confirm Amount | Instructions

You have decided that an amount of 100 Taler per round should be bet in the lottery for the 1. execution of the simulation. You must execute the simulation at least 15 times to continue the experiment.

	SCENARIOS			
Round	LOTTERY BET NOT REDUCED	LOTTERY BET REDUCED	Difference	Realization
1	450.00	450.00	0.00	Won
2	100.00	100.00	0.00	Lost
3	100.00	120.00	-20.00	Lost
4	100.00	140.00	-40.00	Lost
5	100.00	160.00	-60.00	Lost
6	450.00	250.00	200.00	Won
7	450.00	250.00	200.00	Won
8	100.00	180.00	-80.00	Lost
9	100.00	200.00	-100.00	Lost
	∑ 1950.00	∑ 1850.00	∑ 100.00	

Continuous | Step by Step

Back | Next

Figure A7: Instructions Simulation A

General Simulation A Simulation B

Instructions for Simulation B

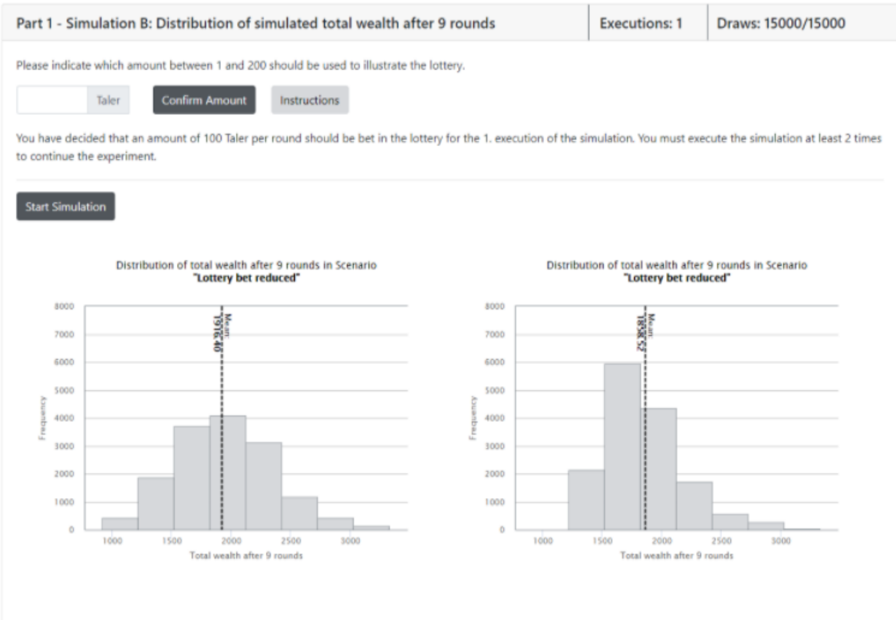
How do you run Simulation B?

First, you are asked to enter a fictitious amount "*x*" in Taler between 1 and 200, with which the simulation of the lottery should be run. Then click on the button "*Confirm Amount*" to confirm the entry.
Click on the button "*Start Simulation*" to start the simulation. You must run Simulation B at least 2 times so that the experiment can be continued.

What does Simulation B show?

You will see two graphics showing the distribution of simulated total wealth after 9 rounds in the lottery. There are 135,000 lottery draws, which corresponds to 15,000 repetitions of the lottery over 9 rounds and thus 15,000 cumulated total wealth realizations after 9 rounds. The left graphic shows the distribution of total wealth after 9 rounds in the first scenario ("*Lottery bet not reduced*") and the right graphic shows the distribution of total wealth after 9 rounds in the second scenario ("*Lottery bet reduced*"). Due to the large number of lottery draws, the underlying characteristics of the lottery based on the respective scenarios can be illustrated here in addition to Simulation A. The higher a bar is, the more often total wealth after 9 rounds was in the corresponding range. If you move the mouse over the individual bars, you will also see what percentage of all previous total wealth realizations after 9 rounds are in this area. At the end of the simulation, the average value of the total wealth after 9 rounds in both scenarios is shown in the respective graphic by a vertical dotted line with the text "*Mean*".

Below is an example of simulation B.



Back Next

Figure A8: Instructions Simulation B

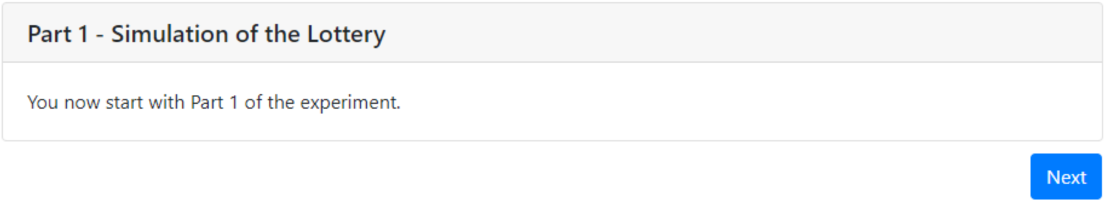
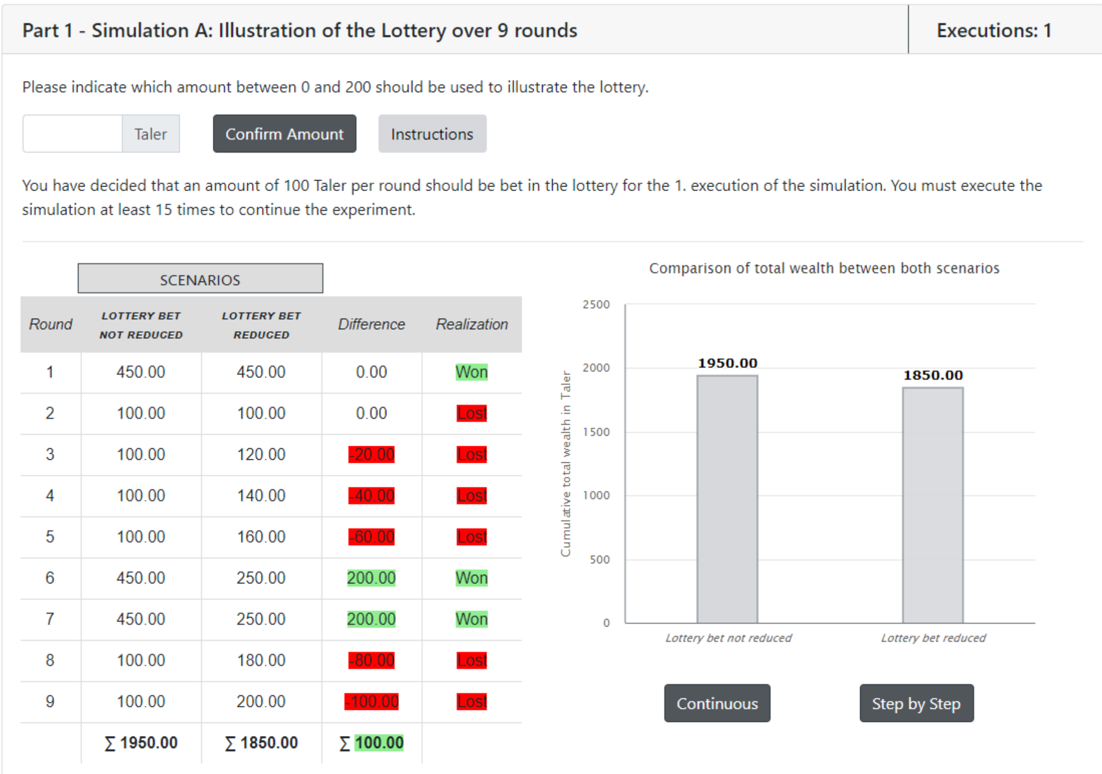


Figure A9: Transition Intervention



Round	LOTTERY BET NOT REDUCED	LOTTERY BET REDUCED	Difference	Realization
1	450.00	450.00	0.00	Won
2	100.00	100.00	0.00	Lost
3	100.00	120.00	-20.00	Lost
4	100.00	140.00	-40.00	Lost
5	100.00	160.00	-60.00	Lost
6	450.00	250.00	200.00	Won
7	450.00	250.00	200.00	Won
8	100.00	180.00	-80.00	Lost
9	100.00	200.00	-100.00	Lost
	∑ 1950.00	∑ 1850.00	∑ 100.00	

Figure A10: Example of Simulation A

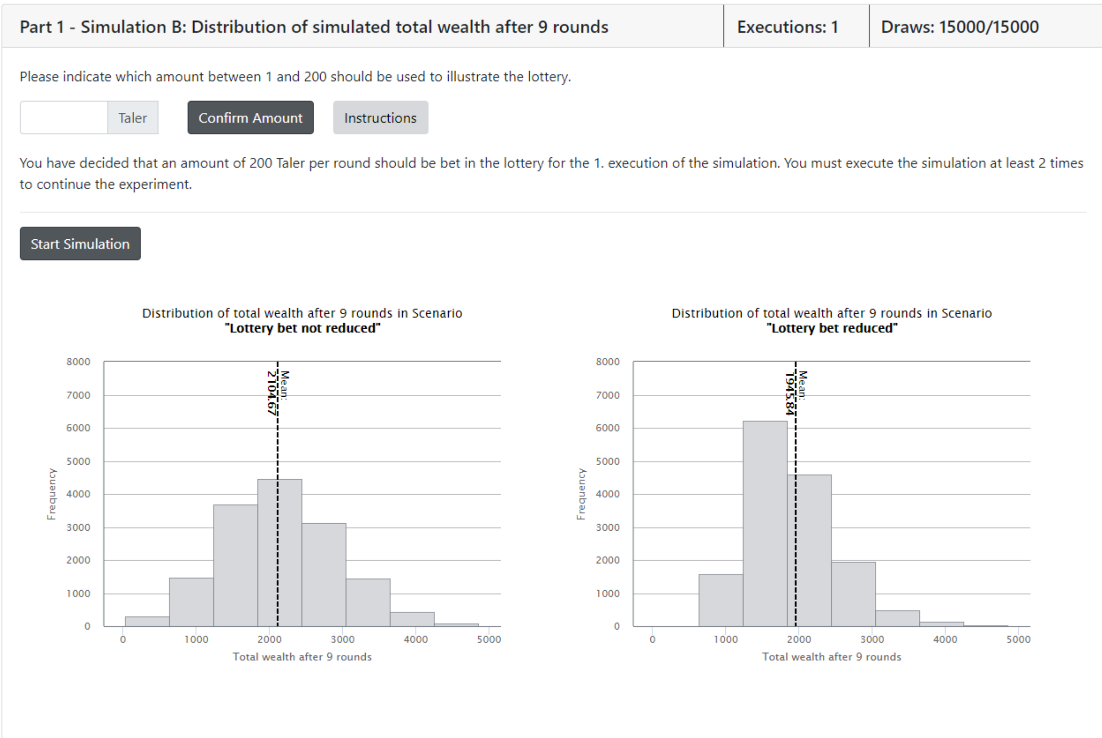


Figure A11: Example of Simulation B

A2.4 MLA Elicitation

A2.4.1 Low Frequency Feedback

Instructions for Part 2

Part 2 of this experiment consists of 9 consecutive rounds. In each of the 9 rounds, you will receive 200 Taler as initial endowment. You must decide in **every third round** what amount "x" of your initial 200 Taler you wish to bet in a lottery with the following characteristics:

With a probability of 2/3 (67%) you lose the amount you bet and with a probability of 1/3 (33%) you receive two and a half times (2.5 times) the amount you bet.

Be aware that you make your decision for the current round as well as for the 2 following rounds. So if you decide to bet an amount "x" in round 1 in the lottery described above, you automatically decide to bet the same amount "x" in this lottery in rounds 2 and 3. Therefore your decision is always valid for 3 consecutive rounds.

If you win in the lottery in one round, you will receive $+2.5 \cdot x$ from the lottery for that round. If you lose in a round in the lottery, you will receive $-x$ from the lottery for that round. Your **total wealth per round** is equal to the win ($+2.5 \cdot x$) or loss ($-x$) from the lottery plus the initial endowment of 200 Taler you receive in each round.

In the following table you will see a general example of how to calculate your total wealth for the first 3 rounds.

Bet amount	Realization of the Lottery Round 1 - Round 2 - Round 3	Total wealth after 3 rounds
x $(0 \leq x \leq 200)$	*Won-Won-Won*	$600 + 2{,}5 \cdot 3 \cdot x$
x $(0 \leq x \leq 200)$	*Won-Won-Lost*	$600 - x + 2{,}5 \cdot 2 \cdot x$
x $(0 \leq x \leq 200)$	*Won-Lost-Won*	$600 - x + 2{,}5 \cdot 2 \cdot x$
x $(0 \leq x \leq 200)$	*Won-Lost-Lost*	$600 - 2 \cdot x + 2{,}5 \cdot x$
x $(0 \leq x \leq 200)$	*Lost-Won-Won*	$600 - x + 2{,}5 \cdot 2 \cdot x$
x $(0 \leq x \leq 200)$	*Lost-Won-Lost*	$600 - 2 \cdot x + 2{,}5 \cdot x$
x $(0 \leq x \leq 200)$	*Lost-Lost-Won*	$600 - 2 \cdot x + 2{,}5 \cdot x$
x $(0 \leq x \leq 200)$	*Lost-Lost-Lost*	$600 - 3 \cdot x$

At the end of the third round, your cumulative total wealth from the last 3 rounds (1 to 3) is displayed. At the beginning of the fourth round you will be asked to make your decision for the next 3 rounds (4 to 6). Your initial endowment for each of these 3 rounds is again 200 Taler, of which you can again bet an amount in the lottery described above for rounds 4 to 6. At the end of the sixth round you will be shown your cumulative total wealth from the last 3 rounds (4 to 6). This process is repeated in the following 3 rounds (7 to 9). Your total wealth from round 4 to 6 and round 7 to 9 is also calculated according to the calculation example above (see table).

Note that your total wealth from all rounds is collected. This means that any wealth already accumulated will not be available for use in the lottery in later rounds. At the end of this part, your total wealth over all nine rounds is added up and displayed in a ratio **1:400** converted to Euro. This sum determines the value of your Amazon voucher for this experiment.

Next

Figure A12: Instructions Elicitation - Low Frequency

Part 2 - Betting in the Lottery

The experiment is now continued with Part 2. Note that the decisions in this part are **payoff relevant**.

Next

Figure A13: Transition Elicitation

Round 1 von 9

Part 2 - Betting in the Lottery

Please specify in the field below, which amount "*x*" of your initial endowment of 200 Taler per round you would like to bet for this round as well as for the next 2 rounds in a lottery with the following characteristics

With a probability of 2/3 (67%) you will lose the amount you bet and with a probability of 1/3 (33%) you will win two and a half times (2.5 times) the amount you bet.

x =

Instructions

Next

Figure A14: Example of First Three Rounds

Round 3 of 9

Part 2 - Your total wealth

The following table shows the realizations of the lottery draws in rounds 1, 2 and 3 and your total wealth in Taler from these three rounds.

Round	Realization of the lottery	Total wealth
1	Won	1000.0 Taler
2	Lost	
3	Won	

Next

Figure A15: Example of First Three Rounds (History)

A2.4.2 High Frequency Feedback

Instructions for Part 2

Part 2 of this experiment consists of 9 consecutive rounds. In each of the 9 rounds, you will receive 200 Taler as initial endowment. In each round you have to decide what amount *x* of your initial endowment of 200 Taler you want to bet into a lottery with the following characteristics:

With a probability of 2/3 (67%) you lose the amount you bet and with a probability of 1/3 (33%) you will receive two and a half times (2.5 times) the amount you bet.

If you win in a specific round in the lottery you will receive $+2{,}5 \cdot x$ from the lottery for this round. If you lose in a specific round of the lottery, you will receive $-x$ for this round from the lottery. Your **total wealth per round** corresponds to the win ($+2{,}5 \cdot x$) or loss ($-x$) from the lottery and the initial endowment of 200 Taler, which you will receive in each round.

In the following table you can see a general example for the calculation of your total wealth the first round.

Bet amount	Realization of the lottery: *Won*	Realization of the lottery: *Lost*
x $(0 \leq x \leq 200)$	$200 + 2{,}5 \cdot x$	$200 - x$

At the end of the first round your total wealth of the first round is displayed. At the beginning of the second round you will be asked to make your decision for the second round. Your initial endowment for this round is again 200 Taler, whereby you can bet any amount into the lottery described above. At the end of the second round your total wealth from the second round is displayed. This procedure is repeated in all subsequent rounds. Your total wealth will be calculated in each round according to the calculation example determined above (see table).

Note that your total wealth from all rounds is collected. This means that any wealth already accumulated will not be available for use in the lottery in later rounds. At the end of this part, your total wealth over all nine rounds is added up and displayed in a ratio **1:400** converted to Euro. This sum determines the value of your Amazon voucher for this experiment.

Next

Figure A16: Instructions stage 2

Part 2 - Betting in the Lottery

The experiment is now continued with Part 2. Note that the decisions in this part are **payoff relevant**.

Next

Figure A17: Transition Elicitation

Round 1 of 9

Part 2 - Betting in the Lottery

Please specify in the field below, which amount "*x*" of your initial endowment of 200 Taler per round you want to bet into a lottery with the following characteristics in this round:

With a probability of 2/3 (67%) you will lose the amount you bet and with a probability of 1/3 (33%) you will win two and a half times (2.5 times) the amount you bet.

x =

Instructions

Next

Figure A18: Example of First Round

Round 1 of 9

Part 2 - Your total wealth

In the following table you can see the realization of the lottery in Round 1 and your total wealth in Taler from the first round.

Round	Realization of the Lottery	Total Wealth
1	Lost	100.0 Taler

Next

Figure A19: Example of First Round (History)

A2.5 Exit Questionnaire and Payoff

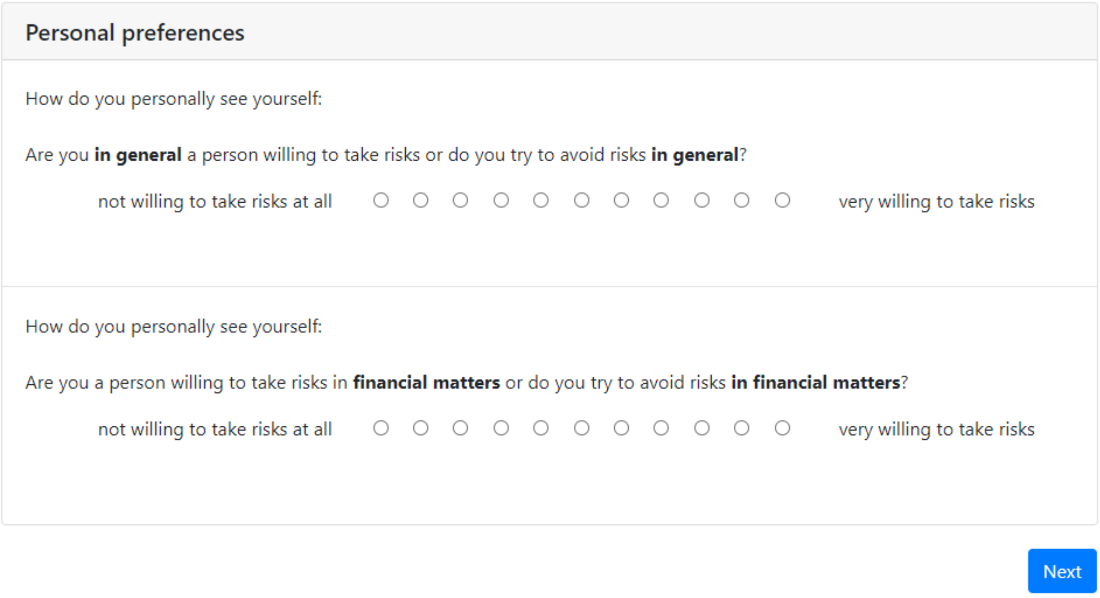
Exit Questionnaire

We would like you to fill out a final questionnaire on the following two pages. Please answer all questions truthfully.

Afterwards you will be informed about your your total payment.

Next

Figure A20: Transition Exit Questionnaire



Personal preferences

How do you personally see yourself:

Are you **in general** a person willing to take risks or do you try to avoid risks **in general**?

not willing to take risks at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ very willing to take risks

How do you personally see yourself:

Are you a person willing to take risks in **financial matters** or do you try to avoid risks **in financial matters**?

not willing to take risks at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ very willing to take risks

Next

Figure A21: Risk Preferences

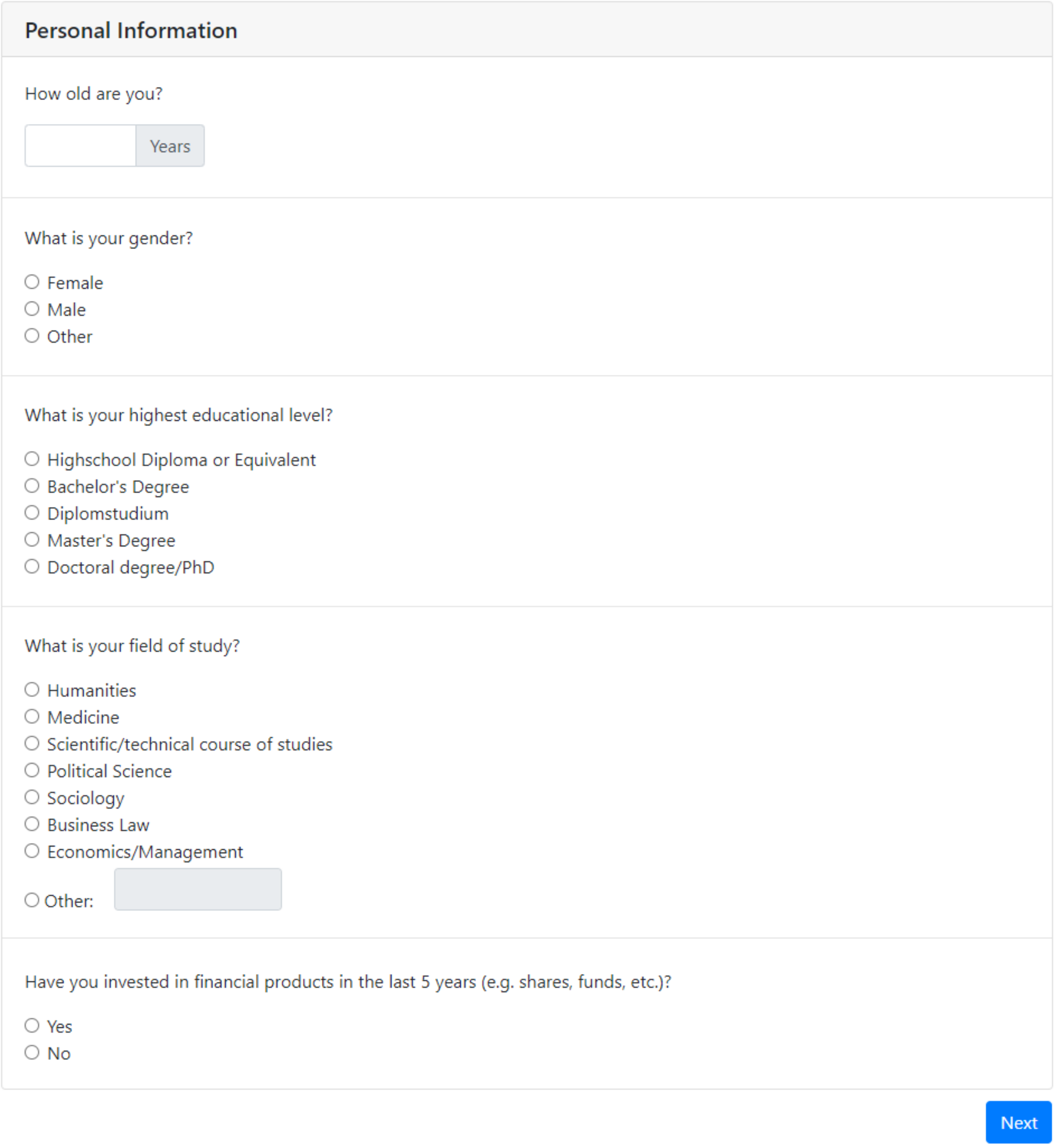
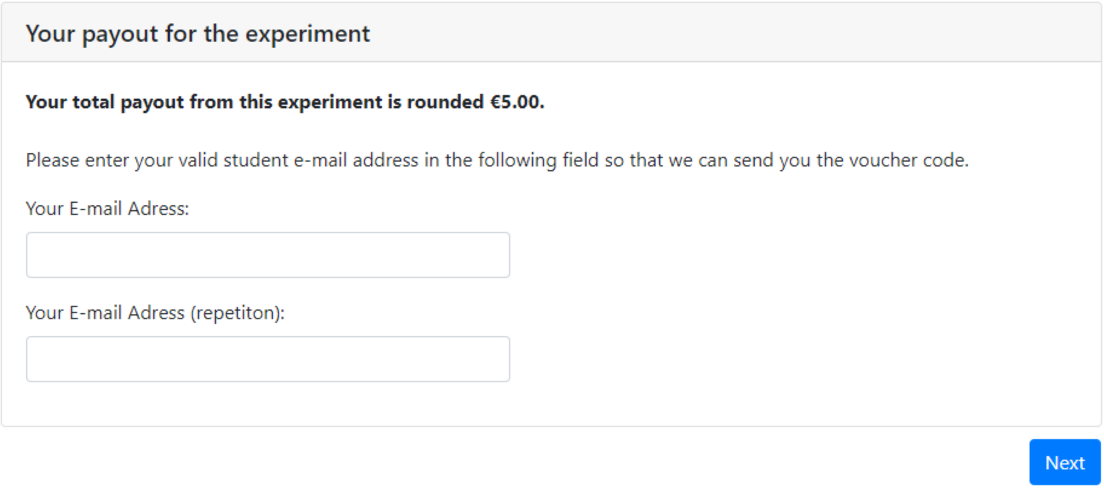


Figure A22: Personal Information



Your payout for the experiment

Your total payout from this experiment is rounded €5.00.

Please enter your valid student e-mail address in the following field so that we can send you the voucher code.

Your E-mail Adress:

Your E-mail Adress (repetiton):

Next

Figure A23: Example of Payoff Information

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2021-01 **Laura Hueber, Rene Schwaiger:** Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

2020-34 **Kai A. Konrad, Florian Morath:** The Volunteer's Dilemma in Finite Populations

2020-33 **Katharina Momsen, Markus Ohndorf:** Expressive Voting vs. Self-Serving Ignorance

2020-32 **Silvia Angerer, Daniela Glätzle-Rützler, Christian Waibel:** Monitoring institutions in health care markets: Experimental evidence

2020-31 **Jana Friedrichsen, Katharina Momsen, Stefano Piasenti:** Ignorance, Intention and Stochastic Outcomes

2020-30 **Esther Blanco, Alexandra Baier, Felix Holzmeister, Tarek Jaber-Lopez, Natalie Struwe:** Substitution of social concerns under the Covid-19 pandemic

2020-29 **Andreas Hackethal, Michael Kirchler, Christine Laudenbach, Michael Razen, Annika Weber:** On the (ir)relevance of monetary incentives in risk preference elicitation experiments

2020-28 **Andrej Gill, Matthias Heinz, Heiner Schumacher, Matthias Sutter:** Trustworthiness in the Financial Industry

2020-27 **Matthias Sutter, Michael Weyland, Anna Untertrifaller, Manuel Froitzheim:** Financial literacy, risk and time preferences - Results from a randomized educational intervention

2020-26 **Rene Schwaiger, Jürgen Huber, Michael Kirchler, Daniel Kleinlercher, Utz Weitzel:** Unequal Opportunities, Social Groups, and Redistribution

2020-25 **Roman Inderst, Martin Obradovits:** Competitive Strategies when Consumers are Relative Thinkers: Implications for Pricing, Promotions, and Product Choice

2020-24 **Martin Obradovits, Philipp Plaickner:** Price-Directed Search and Collusion

2020-23 **Helena Fornwagner, Oliver P. Hauser:** Climate action for (my) children

2020-22 **Esther Blanco, Natalie Struwe, James M. Walker:** Incentivizing public good provision through outsider transfers: experimental evidence on sharing rules and additionality requirements

2020-21 **Loukas Balafoutas, Helena Fornwagner, Rudolf Kerschbamer, Matthias Sutter, Maryna Tverdostup:** Diagnostic Uncertainty and Insurance Coverage in Credence Goods Markets

2020-20 **Anna Ulrichshofer, Markus Walzl:** Customer Disputes, Misconduct, and Reputation Building in the Market for Financial Advice

2020-19 **Anna Ulrichshofer, Markus Walzl:** Social Comparison and Optimal Contracts in the Competition for Managerial Talent

2020-18 **Martin Obradovits, Philipp Plaickner:** Searching for Treatment

2020-17 **Jun Honda:** The Gender-Punishment Gap revisited

2020-16 **Jun Honda:** The Relation between Rankings and Risk-Taking in the Labor Market for Financial Advice

2020-15 **Christina Bannier, Eberhard Feess,Natalie Packham, Markus Walzl:** Differentiation and Risk-Aversion in Imperfectly Competitive Labor Markets

2020-14 **Felix Holzmeister, Rudolf Kerschbamer:** oTree: The Equality Equivalence Test

2020-13 **Parampreet Christopher Bindra, Graeme Pearce:** The effect of priming on fraud: Evidence from a natural field experiment

2020-12 **Alessandro De Chiara, Marco A. Schwarz:** A Dynamic Theory of Regulatory Capture

2020-11 **Christoph Huber, Jürgen Huber, Michael Kirchler:** Market shocks and professionals' investment behavior - Evidence from the COVID-19 crash

2020-10 **Elisabeth Gsottbauer, Daniel Müller, Samuel Müller, Stefan T. Trautmann, Galina Zudenkova:** Social class and (un)ethical behavior: Causal versus correlational evidence

2020-09 **Parampreet Christopher Bindra, Rudolf Kerschbamer, Daniel Neururer, Matthias Sutter:** Reveal it or conceal it: On the value of second opinions in a low-entry-barriers credence goods market

2020-08 **Robert Steiger, Eva Posch, Gottfried Tappeiner, Janette Walde:** Effects of climate change on tourism demand considering individual seasonal preferences

2020-07 **Fang Liu, Alexander Rasch, Marco A. Schwarz, Christian Waibel:** The role of diagnostic ability in markets for expert services

2020-06 **Matthias Stefan, Jürgen Huber, Michael Kirchler, Matthias Sutter, Markus Walzl:** Monetary and Social Incentives in Multi-Tasking: The Ranking Substitution Effect

2020-05 **Michael Razen, Jürgen Huber, Laura Hueber, Michael Kirchler, Matthias Stefan:** Financial Literacy, Economic Preferences, and Adolescents' Field Behavior

2020-04 **Christian König-Kersting, Johannes Lohse, Anna Louisa Merkel:** Active and Passive Risk-Taking

2020-03 **Christoph Huber, Jürgen Huber:** Bad bankers no more? Truth-telling and (dis)honesty in the finance industry

2020-02 **Dietmar Fehr, Daniel Müller, Marcel Preuss:** Social Mobility Perceptions and Inequality Acceptance

2020-01 **Loukas Balafoutas, Rudolf Kerschbamer:** Credence goods in the literature: What the past fifteen years have taught us about fraud, incentives, and the role of institutions

University of Innsbruck

Working Papers in Economics and Statistics

2021-01

Laura Hueber, Rene Schwaiger

Debiasing Through Experience Sampling: The Case of Myopic Loss Aversion.

Abstract
We introduce a training intervention based on a novel tool to mitigate behavior consistent with myopic loss aversion (MLA). We present the results of a large-scale online experiment with 894 student participants. The study featured a two-step debiasing training intervention based on experience sampling and a subsequent elicitation of MLA. We found that participants at baseline exhibit behavior consistent with MLA, which was not the case for decisionmakers who underwent the debiasing training intervention. Nonetheless, we found no statistically significant difference-in- difference effect of the training intervention on the magnitude of MLA. However, when we focused on the more attentive participants by excluding participants with the 10 % longest and 10 % shortest processing times on the task relevant instruction screens, the magnitude of the difference-in-difference effect of the training intervention increased strongly and became statistically significant when controlling for age, gender, education, field of study, investment experience, and financial risk preferences.

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)