

## Behavioral patterns and reduction of sub-optimality: an experimental choice analysis

Daniela Di Cagno<sup>1</sup> · Arianna Galliera<sup>1,4</sup> ·  
Werner Güth<sup>1,2</sup> · Noemi Pace<sup>3</sup>

© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** This paper attempts to identify behavioral patterns and compare their average success considering several criteria of bounded rationality. Experimentally observed choice behavior in various decision tasks is used to assess heterogeneity in how individual participants respond to 15 randomly ordered portfolio choices, each of which is experienced twice. Treatments differ in (not) granting probability information and in (not) eliciting aspirations. Since in our setting neither other regarding concerns nor risk attitude matter and probability of the binary chance move is (optimal) choice irrelevant, categorizing decision types relies on parameter dependence and choice adaptations. We find that most participants reduce systematically sub-optimality when following the identified criteria.

---

The research presented in this paper was financed by the Max Planck Institute for Research on Collective Goods of Bonn.

---

✉ Daniela Di Cagno  
ddicagno@luiss.it

Arianna Galliera  
arianna.galliera@polimi.it

Werner Güth  
gueth@coll.mpg.de

Noemi Pace  
n.pace@unive.it

<sup>1</sup> Luiss, Rome, Italy

<sup>2</sup> Max Planck Institute, Bonn, Germany

<sup>3</sup> Ca' Foscari University of Venice, Venice, Italy

<sup>4</sup> Politecnico di Milano, Milan, Italy

**Keywords** (Un)bounded rationality · Satisficing · Experiments · Heterogeneity

## 1 Introduction

Accepting that sub-optimality is nearly universal (see Conlisk 1996; Selten 2001; and Gigerenzer 2006 for a review of the literature), we attempt to identify behavioral patterns to improve decision success, by reducing the distance between individuals' actual choices and the optimal ones. In this paper, the individual data previously used by Di Cagno et al. (2017) via aggregating across individuals are analyzed to assess individual heterogeneity in decision-making using several criteria of behavioral improvement and of reasonable parameter dependence (Erev and Haruvy 2013; Conte et al. 2015).

The experimental choice tasks are designed to unambiguously distinguish optimal and sub-optimal choices and to avoid various idiosyncratic characteristics by which one can rationalize such individual heterogeneity in decision-making. Specifically, the decision tasks experimentally induce risk neutrality, have only individual consequences and focus on numerical parameter constellations letting optimal choices not depend on probability (information). We additionally distinguish optimality with and without intra-personal payoff aggregation and partly elicit and incentivize aspiration data.<sup>1</sup> Moreover, since participants confront successively twice fifteen different and randomly ordered choice tasks, we also can assess experience effect.

The experimental setup is designed to explore fundamental aspects of decision theory. The decision tasks are framed as portfolio choices<sup>2</sup> and employ binary lottery incentives.<sup>3</sup> Given an initial endowment, each subject allocates it between a risky asset and a risk-free asset; both investment returns determine the probability (in points) of winning the higher payoff (€14), rather than the lower payoff (€4), in both the good and the bad state. The investment is set-optimal when one cannot obtain more in one state without having to reduce what one obtains in the other state. Since deriving the optimal investment, i.e., the investment maximizing the expected utility across both states, may be cumbersome, participants could repeatedly use a slider before committing to one slider position.

Satisficing is testable in the experimental treatments eliciting aspiration levels and incentivizing optimal aspiration formation. Rather than only one aspiration in expected utility terms, participants form payoff aspirations for each realization of the binary random event (the good and the bad state), which allows to test set-optimal satisficing. Satisficing, however, must not be (set)-optimal: an investment choice is satisficing when its success in each random state guarantees the respective aspiration, i.e., when the investment return in each random state is sufficient for what one aspires.

---

<sup>1</sup> For an analysis of the concepts of aspiration and satisficing, see the seminal contributions of Simon (1955), Siegel (1957) and Manski (2017) Sauermann and Selten (1962) and Selten (1998) for the adaptation theory (AAT). For experimental analyses, see Selten et al. (2012) and Hey et al. (2017).

<sup>2</sup> To avoid other relevant concerns, see Harrison and Johnson (2006).

<sup>3</sup> This induces risk neutrality in decision-making (see Roth and Malouf 1979, for an early use in bargaining).

Our previous aggregate data analysis concluded that, in spite of the weak rationality requirements, optimality and optimal satisficing are rare, and there is striking heterogeneity in the individual behavior of participants within and across treatments.

The aim of this study is to investigate whether and how individual participants are able to improve their investment decisions via reducing the distance between their actual choices and the optimal ones. Since the familiar tools for rationalizing individual heterogeneity in behavior are experimentally controlled or excluded, we essentially categorize boundedly rational behavior via criteria of behavioral improvement to confirm that participants learn and thereby enhance their success. Although we are not born optimizers, we may still improve our success by learning. To achieve this aim, we assess three basic criteria of behavioral improvement (improving slider use, reducing antimonotonicity, improving when experiencing the same task again). Improving according to these criteria should enhance individual success. In the treatments eliciting aspiration levels, we additionally investigate improvement in aspiration formation. Altogether, we find that our criteria of behavioral improvement are correlated and enhance average success in terms of investment choices and the consequent payoffs.

The rest of the paper is organized as follows. Section 2 formally describes the choice tasks. Section 3 presents the experimental protocols. Section 4 specifies the improvement criteria, and Sect. 5 reports the evidence for the selected criteria, their correlations and their use in identifying different behavioral types. Section 6 assesses how improvement affects average success. Section 7 concludes.

## 2 Choice tasks and treatments

In each choice task (or case), the decision-maker is endowed with an integer amount  $e (> 0)$  and decides which amount  $i$ , with  $0 \leq i \leq e$ , to invest in a risky asset, whose variable repayment rate is  $r(i) = e - i$  in case of the good state (boom) and 0 in case of the bad state (doom). The remaining amount  $e - i$  has a constant repayment rate  $c > 0$ , irrespective of good or bad state. The probabilities of boom and doom are  $p$  and  $1 - p$  with  $0 < p < 1$ . Participants are subjected to binary lottery incentives,<sup>4</sup> specifically, the choice of  $i$  determines the probability of earning €14 rather than only €4.

The return from investing  $i$  in the risky asset and  $e - i$  in the safe one is  $(e - i)c$  in case of doom and  $(e - i)(c + i)$  in case of boom. Since, due to the binary lottery incentives, expected utility is the probability  $P(i)$  of earning more (€14) rather than less (€4) and  $P(i)$  is restricted to  $0 \leq P(i) \leq 1$ , the expected utility is given by

$$P(i) = (1 - p) \cdot \min \{1, (e - i)c\} + p \cdot \min \{1, (e - i)(c + i)\} \quad (1)$$

Across all cases, one has  $(e - i)c < 1$  via  $ec < 1$ . Thus, constrained optimization requires to determine the level  $i$  for which

<sup>4</sup> For a critical assessment, see Selten et al. (1999) which, however, questions expected utility theory, not binary lottery incentives. Rejecting binary lottery incentives but maintaining expected utility theory is impossible.

**Table 1** Cases and optimal choices

Case	$e$	$c$	$p$	$i^*$	$P(i^*)$	$\overline{P}(i^*)$	$\underline{P}(i^*)$
1	2	0.22	0.66	0.41	0.78	1	0.35
2	2	0.34	0.51	0.22	0.81	1	0.60
3	2	0.40	0.40	0.14	0.85	1	0.75
4	2	0.43	0.32	0.09	0.88	1	0.82
5	2	0.45	0.27	0.07	0.90	1	0.87
6	3	0.15	0.29	0.21	0.59	1	0.42
7	3	0.23	0.23	0.12	0.74	1	0.66
8	3	0.27	0.18	0.07	0.83	1	0.79
9	3	0.29	0.14	0.05	0.88	1	0.86
10	3	0.30	0.12	0.04	0.90	1	0.89
11	4	0.11	0.17	0.15	0.52	1	0.42
12	4	0.17	0.13	0.09	0.71	1	0.67
13	4	0.20	0.10	0.05	0.81	1	0.79
14	4	0.22	0.08	0.03	0.88	1	0.87
15	4	0.22	0.07	0.03	0.88	1	0.87

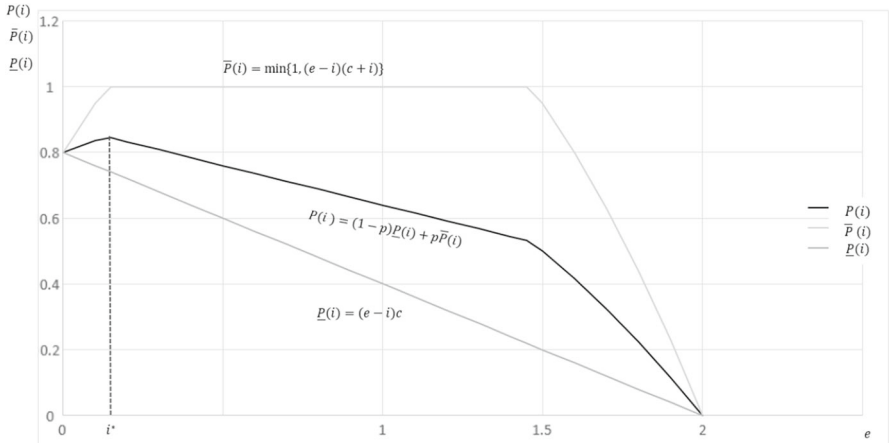
$$P(i) = (1 - p)(e - i)c + p \cdot \min\{1, (e - i)(c + i)\} \quad (2)$$

is maximal.

So expected utility depends on the choice of  $i$  via  $[1 - P(i)]u(\text{€}4) + P(i)u(\text{€}14)$ , with  $u(\cdot)$  denoting the utility of money for all  $0 \leq i \leq e$ . Setting  $u(\text{€}4) = 0$  and  $u(\text{€}14) = 1$  only assumes that more money, €14, is preferred to less, €4, and shows that  $P(i)$  represents the expected utility of choice  $i$ .

Participants confronted the cases illustrated in Table 1, whose rows specify the integer endowment  $e$ , constant repayment rate  $c$ , probability  $p$ , optimal investment  $i^* = \frac{e-c}{2} - \frac{\sqrt{(e+c)^2-4}}{2}$ , expected utility  $P(i^*)$ , as well as €14-probabilities  $\overline{P}(i^*)$  and  $\underline{P}(i^*)$  in boom and doom implied by  $i^*$ . Due to  $i^*$  being a corner maximum of the utility curve, see the  $P(i)$  curve in Fig. 1, the positive probability  $p$  is not (optimal) choice relevant but only payoff relevant. Thus, probability information and transformation (see Kahneman and Tversky 1979 and Prelec 1998) do not matter for optimality.

The between-subject treatments design differs in information regarding probabilities of the good and the bad state and aspiration elicitation: in treatments 1 and 3 (T1 and T3 hereafter), participants are aware of probability  $p$ , whereas in treatments 2 and 4 (T2 and T4 hereafter), no information about probability  $p$  is provided. The known (in treatments 1 and 3) or unknown (in treatments 2 and 4) probabilities of boom and doom are  $p$  and  $1 - p$ , respectively, with  $0 < p < 1$ . In treatments T3 and T4, participants choose not only  $i$  but also their aspiration levels for the probability of earning €14 in case of boom and doom, respectively, denoted by  $\overline{A}$  and  $\underline{A}$ , with  $\overline{A} \geq \underline{A}$ . In the following, we refer to  $\overline{A}$  and  $\underline{A}$  as aspirations. Since in boom and doom the probabilities of earning €14 are  $\overline{P}(i) = \min\{1, (e - i)(c + i)\}$  and  $\underline{P}(i) = (e - i)c$ , respectively,



**Fig. 1** Illustration of payoff incentives

the choice of  $i$  satisfies the aspiration profile  $(\bar{A}, \underline{A})$  if  $\bar{P}(i) \geq \bar{A}$  and  $\underline{P}(i) \geq \underline{A}$ . With the help of the choices  $i$ ,  $\bar{A}$  and  $\underline{A}$  one can therefore test the satisficing hypothesis. Satisficing is set-optimal if  $\bar{P}(i) = \bar{A}$ ,  $\underline{P}(i) = \underline{A}$  and  $i$  satisfies  $0 \leq i \leq i^*$ . Otherwise, one could increase  $\underline{P}(i)$  without decreasing  $\bar{P}(i) = 1$ , or vice versa. To test (set-optimal) satisficing, aspiration formation is incentivized. A participant earns €14 with probability  $\bar{A}$  or  $\underline{A}$  when satisficing in boom or doom, respectively, and otherwise nothing at all (meaning they earn €4 with probability 1). In particular, he experiences “burning money” when sets an aspiration level lower than the probability implied by its own investment choice (i.e., when  $\bar{P}(i) > \bar{A}$  or  $\underline{P}(i) > \underline{A}$  in boom and doom, respectively) or experiences “committing suicide” when, on the contrary, he sets an aspiration level greater than the probability implied by its own investment choice (i.e., when  $\bar{A} > \bar{P}(i)$  or  $\underline{A} > \underline{P}(i)$  in boom and doom, respectively). In treatment T2, which does not provide probability information and does not elicit aspirations, participants are asked for their (not incentivized) subjective probability  $\hat{p}$  of boom. This subjectively stated probability  $\hat{p}$  in T2 and the aspirations  $\bar{A}$  and  $\underline{A}$  in T3 and T4 can be once revised before confirming one’s choice of  $i$  without affecting the maximum number of slider use (six) per choice task.

Figure 1 illustrates the  $i$  allocations and the related probabilities  $P(i)$ ,  $\bar{P}(i)$ ,  $\underline{P}(i)$ .

### 3 Experimental protocols

The experiment consists of two phases with 15 rounds each, corresponding to the 15 cases in Table 1 which participants confront in idiosyncratic random orders.<sup>5</sup> At the beginning of each round participants are endowed with an amount of token which can be allocated in a risk-free bond with a state-independent repayment factor ( $c$  in Table 1, and a risky asset whose repayment factor changes with the market condition and the

<sup>5</sup> Di Cagno et al. (2017) also analyze three control cases with  $c = 0$ .

**Table 2** The  $2 \times 2$ -factorial between-subject treatments

Choice Format	Probability information	
	$p$ given	$p$ unknown
$I$ -treatment (direct $i$ -choice)	T1	T2
	1. $i_1, i_2, i_3, i_4, i_5, i_6$	1. $\hat{p}$
	$\rightarrow$ final choice of $i$	$\rightarrow$ final choice of $i$
	T3	T4
$S$ -treatment (first aspiration profiles, then choice)	1. $\bar{A}, \underline{A}$	1. $\bar{A}, \underline{A}$
	2. $i_1, i_2, i_3, i_4, i_5, i_6$	2. $i_1, i_2, i_3, i_4, i_5, i_6$
	$\rightarrow$ final choice of $i$	$\rightarrow$ final choice of $i$

amount invested in it. The four treatments differ in terms of probability information and aspiration elicitation. In T1 and T3, participants are informed of the probabilities of boom and doom ( $p$  and  $1 - p$ ), about which they are unaware in T2 and T4. Only in T2, they report their beliefs  $\hat{p}$  and  $1 - \hat{p}$  and only in T3 and T4 participants state their aspirations  $\bar{A}$  and  $\underline{A}$  for both states, (not) knowing the probability  $p$  in T3 (T4). The four treatments are depicted in Table 2.

Before choosing  $i$ , participants can test up to six choices of  $i$  by moving a cursor on the scroll bar, as depicted in Fig. 2 for T1 and T2<sup>6</sup> (and in Fig. 4, in Appendix B, for T3 and T4).<sup>7</sup> The cursor allows participants to change the endowment share ( $i/e$ ) invested in the risky asset and thus the probabilities  $\bar{P}(i)$  and  $\underline{P}(i)$  of gaining €14, which are represented by the left (right) column height in boom (left) or doom (right). After confirming their final choice, participants proceed to the next round.

Let us illustrate the example in Fig. 2 with the endowment of  $e = 3$ , the repayment factor for the risk-free bond  $c = 0.29$ , and probability of 14% ( $p$ ) and 86% ( $1 - p$ ), respectively, for boom and doom.

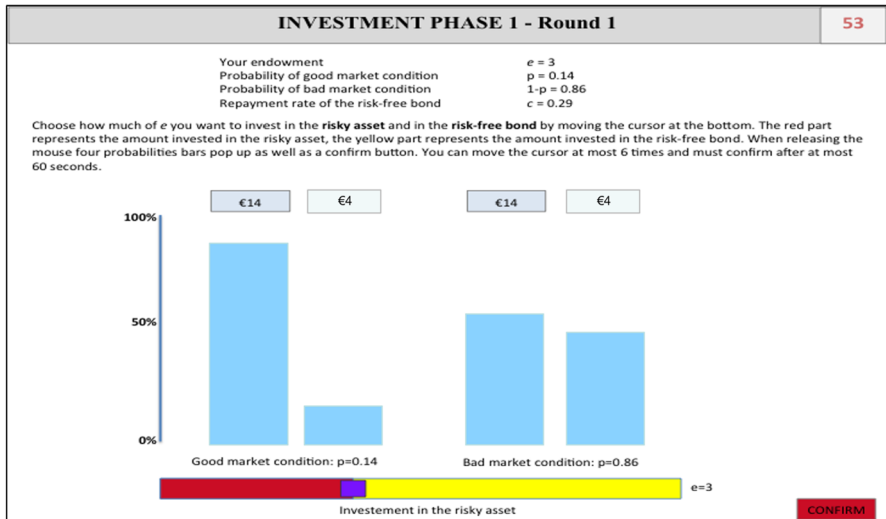
In the example, the participant invests in the risky asset  $i = 1$  which leaves for the risk-free bond  $(e - i) = 2$ . This determines the probability of earning €14 and the complementary probability of earning €4, which both depend on good or bad state.

In the good state, the probability of earning €14 is given by the sum of repayment of the risk-free bond,  $c(e - i)$ , and the repayment of the risky asset,  $(e - i)i$ . In the screenshot, the probability of earning €14, calculated by the computer, is represented by the height of the column corresponding to the good market condition. In this example, the probability of earning €14 is approximately 80%.

In the bad state, the probability of earning €14 is given only by the repayment of the risk-free bond,  $c(e - i)$ . In the screenshot, the probability of earning €14, calculated by the computer, is represented by the height of the column corresponding to the bad market condition. In the example, the probability of earning €14 is approximately

<sup>6</sup> The probability information is the only difference between T1 and T2.

<sup>7</sup> In the experimental setting of T3 and T4 the aspiration level is set on the €14 bar probabilities and we avoid confusion with the second (complementary) bar.



**Fig. 2** Investment choice in T1

50%. The translated instructions (see Appendix B) provide more details regarding how treatments are described.

The altogether 30 choices  $i$  are the only choice data in T1. In T2, we additionally elicit in each round the stated probability  $\hat{p}$  by asking participants to write down their expectations regarding the probabilities of market in good or bad state, before the investment choice (see instructions for treatment T2 in Appendix B). Similarly, in T3 and T4 we elicit in each round the aspirations  $\bar{A}$  and  $\underline{A}$ , by asking participants to assess them before the investment choice (see instructions for treatments T3 and T4).

The experiment was performed in the Cesare Lab of Luiss University in Rome with 298 student participants. The number of participants per treatment varies from 71 to 78. For each treatment, we conducted three sessions with student participants, recruited using Orsee (Greiner 2004) from different fields of study (mainly Economics, Law and Political Sciences). No one participated in more than one session. The experiment was computerized using z-tree (Fischbacher 2007). The sessions lasted approximately 90 min. After each session, the participants answered a brief questionnaire, mainly to collect demographic information, before being privately paid in cash for a randomly selected round.<sup>8</sup>

## 4 Improvement criteria

Our data show that participants quite systematically behave sub-optimally. But do participants learn to improve their success? In our setting, participants may learn in three main ways:

<sup>8</sup> Paying participants only for a random round is required by binary-lottery incentives and avoids past-earnings effects.

### 1. Improving slider attempts.

For each participant, denote by  $i_{t^*}$  the last confirmed slider position. For participants with repeated slider attempts in the same task, one can investigate the last slider adjustment  $i_{t^*} - i_{t^*-1}$ , and, if available, the second-last ones,  $i_{t^*-1} - i_{t^*-2}$ , etc.

Based on this notation, slider use is improving for  $i_{t^*} \geq i^*$  when

$$(1.1) \quad i_{t^*} < i_{t^*-1} \text{ or}$$

$$(1.2) \quad i_{t^*} \leq i_{t^*-1} \leq i_{t^*-2}$$

with at least one weak inequality being strict.<sup>9</sup>

### 2. Reducing antimonotonicity

We define two types of antimonotonicity, timewise and casewise. More specifically, denote by  $k = 2, \dots, 15$  the round in a given phase in which a given participant encounters a specific case.

Timewise monotonicity requests for two successive choices  $i_k$  and  $i_{k-1}$  that  $(i_k - i_{k-1})$  has the same sign as  $(i_k^* - i_{k-1}^*)$ , for the optimal choices  $i_k^*$  and  $i_{k-1}^*$ , i.e., if

$$(2.i) \quad \text{sign}(i_k - i_{k-1}) \times \text{sign}(i_k^* - i_{k-1}^*) > 0.$$

Casewise monotonicity (within a given phase) similarly postulates

$$(2.ii) \quad \text{sign}(i_k - i_{k'}) \times \text{sign}(i_k^* - i_{k'}^*) > 0 \text{ with } k \neq k'.$$

The latter requirement is more demanding by allowing to compare every case  $k$  with all 14 other cases  $k'$ , often separated by several intermediate cases between  $k$  and  $k'$ . We, thus, measure as an improvement the reduction of the individual frequencies of cases in which 2.i and 2.ii are not met.

### 3. Reduction of non-improving across phases

When confronting each case once again in a new random order, the second investment choice  $i$  being closer to  $i^*$  in the sense of smaller  $\max\{0, i - i^*\}$  is another form of learning.

For a given case, we compare the two  $i$ -choices  $i_1$  in phase 1 and  $i_2$  in phase 2 of a specific participant to assess individual improvement due to experience via reducing the frequency of improving choices across phases, specifically by identifying non-improving for each case via

$$(3) \quad \max\{i_2 - i^*, 0\} > \max\{i_1 - i^*, 0\} \geq 0,$$

i.e., when for the given case, the 2nd-phase choice  $i_2$  is further away than  $i_1$  from the optimal choice  $i^*$ . We thus measure experience by reducing the individual frequency of cases with no improvement in the sense of (3).

Behavioral improvements altogether should increase the frequency of adjusting  $i$  nearer to  $i^*$  in the range  $i$  greater than  $i^*$  and consequently reducing the difference between  $P(i^*)$  and  $P(i)$ , the expected loss due to sub-optimality. Addi-

<sup>9</sup> In a similar manner, one could check slider positions for  $i_{t^*-\tau} \leq i^*$  with  $\tau = 1, 2$  to determine whether they return to the range  $[0, i^*]$  via  $i_{t^*} \leq i^*$ ; however, such an analysis is omitted because only too few data are available.



tionally, behavioral improvement in T3 and T4 should reduce  $\min \{ \bar{P}(i) - \bar{A}, 0 \}$  and  $\min \{ \underline{P}(i) - \underline{A}, 0 \}$ , “burning money”, and decrease the frequency of non-satisficing via  $\bar{A} > \bar{P}(i)$  or  $\underline{A} > \underline{P}(i)$ , “committing suicide”.

### 5 Improvement evidence

Participants differ in terms of how often they comply with improvement criteria. Obviously, some criteria apply only to some between subjects treatments, e.g., improving slider use and antimonotonicity, are verified between phases while reduction of non-improving is analyzed across phases. In the following, it will be shown that most participants improve their behavior and that this is payoff enhancing.

#### 5.1 Slider attempts, antimonotonicity and reduction of non-improving

According to Fig. 3, comparing earlier and later successive slider adjustments from left to right in phases 1 (top) and 2 (bottom), slider adjustments on average become smaller when approaching the final attempt  $i_{t^*}$ .

In particular, the last slider adjustments significantly reduce the distance of the actual investment from the optimal one (see Table 13 for average  $i_{t^*} - i^*$ ,  $i_{t^*-1} - i^*$ ,  $i_{t^*-2} - i^*$  in phases 1 and 2 in Appendix A). Moreover, in phase 2, the distances of the actual investments from the optimal investments are always smaller than in phase 1 (based on the Wilcoxon Signed Rank test—WSR hereafter—the difference between

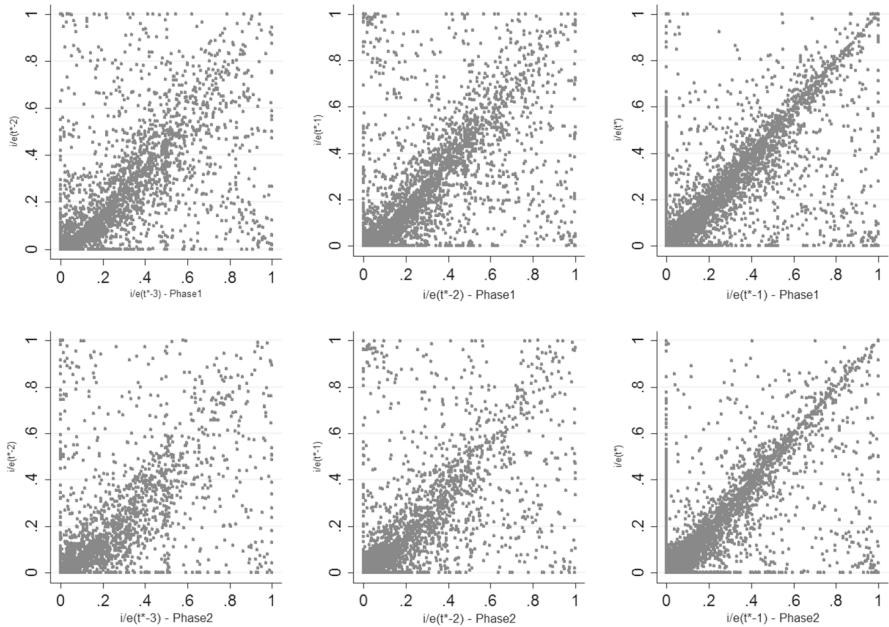


Fig. 3 Adjustments in slider attempts for all treatments

phase 1 and phase 2 is significant,  $p$  value  $< 0.001$ ). Participants adjust mainly towards the optimal investment in successive slider attempts and begin to adjust the slider more carefully and more deliberately, before confirming their choice of  $i$ . Table 3 reveals that (later) slider trials (more) significantly (at the 1% level) reduce the differences between the actual and optimal  $i$ -choice across all cases.

**Result 1** *Aggregate tendencies suggest improvement on slider use.*

**Table 3** Correlations between slider adjustment trials and deviations from the optimal investment

	Phase 1			Phase 2		
	$i_{t^*} - i^*$	$i_{t^*-1} - i^*$	$i_{t^*-2} - i^*$	$i_{t^*} - i^*$	$i_{t^*-1} - i^*$	$i_{t^*-2} - i^*$
Try	-0.08*** (0.01)	-0.06*** (0.01)	-0.03*** (0.01)	-0.09*** (0.01)	-0.06*** (0.01)	-0.00 (0.01)
Constant	0.88*** (0.03)	0.87*** (0.03)	0.81*** (0.05)	0.72*** (0.02)	0.70*** (0.03)	0.50*** (0.05)
Observations	4530	4516	3484	4530	4523	3209

Standard errors are reported in parentheses

\*\*\* significance at 1% level

**Table 4** Correct slider use: movement to the optimal  $i^*$  from the penultimate to the ultimate slider use

	Phase 1	Phase 2	$P$ value test mean difference
Treatment 1 (2340 investment choices)			
Choices	72.22%	76.58%	0.016
# subjects (rule > 50%)	69 (88.46%)	68 (87.18%)	0.343
# subjects (rule > 60%)	64 (82.05%)	62 (79.49%)	0.116
# subjects (rule > 70%)	43 (55.13%)	54 (69.23%)	0.000
Treatment 2 (2340 investment choices)			
Choices	70.68%	74.87%	0.023
# subjects (rule > 50%)	70 (89.74%)	72 (92.31%)	0.030
# subjects (rule > 60%)	63 (80.77%)	67 (85.90%)	0.001
# subjects (rule > 70%)	48 (61.54%)	49 (62.82%)	0.523
Treatment 3 (2250 investment choices)			
Choices	52.98%	48.98%	0.058
# subjects (rule > 50%)	44 (58.67%)	37 (49.33%)	0.000
# subjects (rule > 60%)	29 (38.67%)	31 (41.33%)	0.197
# subjects (rule > 70%)	15 (20.00%)	19 (25.33%)	0.003
Treatment 4 (2130 investment choices)			
Choices	57.18%	57.65%	0.827
# subjects (rule > 50%)	46 (64.79%)	46 (64.79%)	1.000
# subjects (rule > 60%)	37 (52.11%)	53.52 (38%)	0.515
# subjects (rule > 70%)	16 (22.54%)	22 (30.99%)	0.000

**Table 5** Timewise antimonotonicity

	Phase 1	Phase 2	<i>P</i> value test mean difference
Treatment 1			
Choices	39.65%	28.19%	0.000
# subjects (rule > 50%)	30 (38%)	11 (14%)	0.000
# subjects (rule > 60%)	15 (19%)	7 (9%)	0.000
# subjects (rule > 70%)	7 (9%)	6 (8%)	0.262
Treatment 2			
Choices	44.51%	37.72%	0.001
# subjects (rule > 50%)	39 (50%)	27 (35%)	0.000
# subjects (rule > 60%)	20 (26%)	16 (21%)	0.000
# subjects (rule > 70%)	9 (12%)	6 (8%)	0.002
Treatment 3			
Choices	48.29%	43.64%	0.033
# subjects (rule > 50%)	40 (53%)	33 (44%)	0.000
# subjects (rule > 60%)	21 (28%)	18 (24%)	0.031
# subjects (rule > 70%)	13 (17%)	14 (19%)	0.411
Treatment 4			
Choices	59.05%	51.79%	0.001
# subjects (rule > 50%)	53 (75%)	46 (65%)	0.000
# subjects (rule > 60%)	36 (51%)	30 (42%)	0.000
# subjects (rule > 70%)	26 (37%)	15 (21%)	0.262

The analysis by treatment reveals large heterogeneity of improving slider use choices in phases 1 and 2. In more detail, Table 4 indicates that the share of improving slider use significantly increases from phase 2 to phase 1 only in treatments T1 and T2.

In our view, this is due to the fact that participants in treatments T3 and T4 have to additionally focus on improving aspiration formation, i.e., the cognitive burden of aspiration formation weakens “improving slider use” across phases. At the individual level, one can categorize individual participants on the basis of improving slider use, by distinguishing three compliance levels of at least 50, 60 and 70%. A significant improvement from phase 1 to phase 2 applies only to subjects at the 70% compliance level in all treatments except for treatment T2 (see Table 4). These subjects, who better utilize the slider, also should improve their average performance from phase 1 to phase 2.

Timewise antimonotonicity is greatly and significantly reduced from phase 1 to phase 2 in all treatments (see Table 5). Although the percentage of anti-monotonic choices is significantly greater in the cognitively more demanding satisficing treatments, T3 and T4, the average tendencies reveal overwhelming improvement. In all treatments, some participants defined as ‘anti-monotonic’ in phase 1 become ‘monotonic’ in phase 2, i.e., learn to behave more consistently, for all compliance levels, except the 70%.

**Table 6** Casewise antimonotonicity

	Phase 1	Phase 2	<i>P</i> value test mean difference
Treatment 1 (2340 investment choices)			
Choices	40.5%	30.84%	0.002
# subjects (rule > 50%)	31 (39.74%)	19 (24.36%)	0.040
# subjects (rule > 60%)	22 (28.21%)	15 (19.23%)	0.190
# subjects (rule > 70%)	15 (19.23%)	13 (16.67%)	0.679
Treatment 2 (2340 investment choices)			
Choices	44.99%	38.25%	0.028
# subjects (rule > 50%)	41 (52.56%)	30 (38.46%)	0.078
# subjects (rule > 60%)	29 (37.18%)	24 (30.77%)	0.401
# subjects (rule > 70%)	21 (26.92%)	15 (19.23%)	0.257
Treatment 3 (2250 investment choices)			
Choices	48.50%	43.43%	0.106
# subjects (rule > 50%)	47 (62.67%)	35 (46.67%)	0.050
# subjects (rule > 60%)	36 (48.00%)	30 (40.00%)	0.327
# subjects (rule > 70%)	26 (34.67%)	24 (32.00%)	0.731
Treatment 4 (2130 investment choices)			
Choices	56.16%	51.29%	0.101
# subjects (rule > 50%)	58 (81.69%)	47 (66.20%)	0.036
# subjects (rule > 60%)	51 (71.83 )	46 (64.79%)	0.371
# subjects (rule > 70%)	38 (53.52%)	32 (45.07%)	0.317

The frequencies of casewise antimonotonicity are greatly and significantly reduced from phase 1 to phase 2 in T1 and T2 (see Table 6) but to a lesser extent in T3 and T4 where on the individual level we observe consistent improvement between phase 1 and phase 2 only at the 50% level of compliance.

As expected, also in the more cognitively demanding casewise antimonotonicity there is an overall improvement of individual investment behavior.

**Result 2** *On average timewise and casewise antimonotonicity decrease from phase 1 to phase 2.*

In contrast with other behavioral improvements, we cannot compare the reduction of non-improving between phases 1 and 2.

Overall, Table 7 shows that the non-improving frequency is the smallest in treatment T1 and is the highest in treatment T4 which requires to state the aspiration level without any information on the probability of the good or bad state.

**Table 7** Non-improving frequency from phase 1 to phase 2

	T1	T2	T3	T4
# subjects	78	78	75	71
Average	3.440	4.115	6.107	6.775
SD	2.541	2.763	3.047	3.067

Average non-improving frequency. The non-improving individual frequency is a variable from 0 to 15

**Table 8** Pairwise correlation (using the threshold at 60%)

Correlation	Treatment 1		Treatment 2		Treatment 3		Treatment 4	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
Slider use/casewise monotonic								
Coeff.	0.1523	0.4772	-0.1061	-0.1105	0.2148	0.2985	0.1616	0.2731
<i>p</i> value	(0.183)	(0.000)	(0.355)	(0.336)	(0.064)	(0.009)	(0.178)	(0.021)
Slider use/timewise monotonic								
Coeff.	0.0261	0.2848	0.0115	-0.1146	0.1293	0.2181	0.2685	0.1725
<i>p</i> value	(0.821)	(0.012)	(0.921)	(0.318)	(0.269)	(0.060)	(0.024)	(0.150)
Casewise monotonic/timewise monotonic								
Coeff.	0.7785	0.6435	0.6418	0.762	0.5302	0.6882	0.6351	0.5346
<i>p</i> value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Correlations between individual dummies for a threshold of 60%

## 5.2 Correlation among criteria

Pairwise correlating improvement in slider use and antimonotonicity reduction (see Table 8) reveals a significant positive correlation between slider use and casewise monotonicity in T1, T3 and T4 (in phase 2) but not T2. The correlation between slider use and timewise monotonicity is less clear in spite of the always significantly positive correlation between casewise and timewise monotonicity improvement.

Table 9 refines the correlation analysis between casewise and timewise monotonicity improvements by differentiating between participants with small ( $< 60\%$ ) and large ( $\geq 60\%$ ) shares of correct slider adjustments. Correlation of monotonicity improvement is less relevant (although still significant) even for the smaller share of correct slider adjustments (lower panel in Table 8).

**Result 3** *Across all four treatments, there are significant correlations between the improvement criteria regarding slider use and monotonically reacting to cases.*

## 5.3 Experience

We consider the experience based on the improvement criteria for phases 1 and 2 by combining participants into eight categories representing the average individual tendency of using correctly the slider use, time/casewise monotonicity with the non-

**Table 9** Pairwise correlation (for a threshold of 60%) by slider use

Definition of monotonic subject	Treatment 1		Treatment 2		Treatment 3		Treatment 4	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
Share of "correct" slider choices $\geq 60\%$								
Timewise and casewise monotonic correlation	0.832 (0.000)	0.689 (0.000)	0.645 (0.000)	0.768 (0.000)	0.704 (0.000)	0.713 (0.000)	0.574 (0.000)	0.531 (0.001)
Definition of antimonotonic subject								
Share of "correct" slider choices $< 60\%$								
Timewise and casewise antimonotonic correlation	0.603 (0.0224)	0.509 (0.044)	0.659 (0.008)	0.671 (0.024)	0.423 (0.003)	0.653 (0.000)	0.689 (0.000)	0.503 (0.003)

Correlations between individual dummies with a threshold of 60%

improving frequency across phases.<sup>10</sup> Table 10 reports the non-improving frequency separately for T1 and T2 and for phases 1 and 2, considering the other improving criteria.<sup>11</sup>

The majority of participants are consistent (yes, yes, yes) with improving slider use and both measures of anti-monotonicity reduction, and this category invariantly has the lowest average non-improving frequency score in both T1 and T2 and both Phases 1 and 2; in particular, non-improving frequency drops in average from 3.21 to 2.82 in T1 and from 3.46 to 2.80 in T2.<sup>12</sup> Categories with no such reduction of the average non-improving score usually contain rather few members.<sup>13</sup>

For treatments T3 and T4, we did not expect (see Table 10) the same low non-improving scores as for treatments T1 and T2 because participants in T3 and T4 face serious additional cognitive demands when forming aspirations  $\bar{A}$  and  $\underline{A}$  and trying to avoiding “burning too much money” and “committing suicide”. Actually, the non-improving scores, reported in Table 10, confirm our expectations.

Nevertheless, the non-improving frequency, separately for T3 and T4 and for phase 1 and 2, shows that all improvement criteria—“improving slider use”, “reduced anti-monotonicity” and “reduced non-improving”—are highly correlated, as in T1 and T2. Also, in this case, improving according to one criterion reliably predicts improvements in the other ones: those participants consistently in line with the improving criteria (yes, yes, yes) have the lowest non-improving score.<sup>14</sup>

**Result 4** *Participants learn to react qualitatively to parameter changes, as suggested by optimality  $i^*$ .*

Altogether, we conclude that all improvement criteria—“improving slider use”, “reduction of anti-monotonicity” and less “non-improving”—are highly correlated: improving according to one criterion reliably predicts improvements in the other ones, where, of course, the categorization is partly binary (only non-improving frequency is assessed numerically).

## 6 Success improvement

We distinguish different individual patterns based on individual compliance to improvement criteria to point out possible effects on reducing the sub-optimality  $P(i^*) - P(i)$ . We adopt the improvement criteria, i.e., improving slider attempts,

<sup>10</sup> We adopt the 60% individual level of compliance per phase.

<sup>11</sup> The reason for distinguishing phases 1 and 2 in Table 10 is that the categorization of the eight behavioral categories is independently performed for phase 1 and phase 2 data, i.e., an individual participant may belong in phase 2 to a different class than in phase 1.

<sup>12</sup> Table 14 (in the Appendix A) additionally controls for whether participants switch categories from phase 1 to 2. Again, the dominant categories are those with (yes, yes, yes)-participants in both phases (41 in T1 and 31 in T2), whose average non-improving score in phase 2 is 2.68 in T1 and 2.74 in T2, i.e. slightly lower than the respective phase 2-score in Table 10.

<sup>13</sup> An exception is phase 1 of T2 and the case (no, yes, yes), whose non-improving frequency score is 3.30 and thus smaller than the score of 3.46 for (yes, yes, yes).

<sup>14</sup> Exceptions exist in phase 1 of T4, e.g. case (no, yes, yes) with only 7 participants whose non-improving score is 5.43, less than 6.00 for (yes, yes, yes).

**Table 10** Non-improving frequency between phase 1 and phase 2

Correct slider use	Timewise monotonic		Casewise monotonic		Treatment 1				Treatment 2					
	Timewise monotonic	Casewise monotonic	Phase 1		Phase 2		Phase 1		Phase 2					
			#	Mean	S.d.	#	Mean	S.d.	#	Mean	S.d.	#	Mean	S.d.
Yes	No	No	12	3.83	2.17	3	5.67	2.08	15	4.67	2.82	15	5.20	2.76
Yes	No	Yes	0	-	-	0	-	-	1	8.00	-	0	-	-
Yes	Yes	No	4	4.25	0.96	3	4.33	1.15	10	5.20	3.26	7	6.43	2.64
Yes	Yes	Yes	48	3.21	2.65	56	2.82	2.30	37	3.46	2.65	45	2.80	1.91
No	No	No	3	3.33	2.52	4	4.25	2.63	3	5.67	3.79	1	10.00	.
No	Yes	No	3	4.33	1.53	5	5.40	3.13	1	6.00	-	1	8.00	.
No	No	Yes	0	-	-	0	-	-	1	7.00	-	0	-	-
No	Yes	Yes	8	3.50	3.51	7	5.14	3.02	10	3.30	1.70	9	6.00	2.83

Correct slider use	Timewise monotonic		Casewise monotonic		Treatment 3				Treatment 4					
	Timewise monotonic	Casewise monotonic	Phase 1		Phase 2		Phase 1		Phase 2					
			#	Mean	S.d.	#	Mean	S.d.	#	Mean	S.d.	#	Mean	S.d.
Yes	No	No	6	6.83	4.26	4	6.00	3.16	14	6.21	3.47	11	6.82	3.16
Yes	No	Yes	-	-	-	-	-	-	-	-	-	1	8.00	-
Yes	Yes	No	4	5.75	4.35	3	6.67	5.13	10	6.00	3.27	9	6.78	3.87
Yes	Yes	Yes	19	5.21	2.51	24	4.33	2.30	13	6.15	2.27	17	4.29	2.02
No	No	No	13	6.62	2.36	14	8.14	3.46	22	8.50	2.96	16	7.31	2.73
No	Yes	No	13	6.77	3.27	9	7.11	2.57	5	5.80	3.03	10	9.70	2.41
No	No	Yes	2	5.00	-	-	-	-	-	-	-	-	-	-
No	Yes	Yes	18	6.17	3.43	21	6.29	2.55	7	5.43	2.07	7	7.14	1.46

Dummies with 60% threshold. The non-improving individual frequency is a variable from 0 to 15



reducing timewise antimonicity and reducing casewise antimonicity which generate the 8 categories in Table 9, without explicitly referring to non-improving frequency<sup>15</sup> but considering the relevant correlation with the other dimensions.<sup>16</sup> Table 11, based on this  $2 \times 2 \times 2$ -factorial improvement classification, i.e., the 8 rows in the upper and lower subtable, reports the average expected success  $P(i)$  and the success loss  $P(i^*) - P(i)$ , separately for phase 1 and phase 2 (since individual participants may belong to different categories in phases 1 and 2) and treatments T1 and T2. Participants with “yes” for all three criteria represent the majority (48 or 56 in T1, and 37 or 45 in T2) with their number increasing from phase 1 to phase 2. Furthermore, the success  $P(i)$  of earning €14 of this fully compliant group is generally the highest and increases from phase 1 to phase 2. For T1 and T2, this result confirms that experience enhances success.

**Result 5** *In treatment T1 and T2, participants compliant with the three improvement criteria succeed in increasing the probability of earning €14, reducing the distance  $P(i^*) - P(i)$ . This success improvement increases from phase 1 to phase 2.*

Because success in T3 and T4 depends not only on the investment choice  $i$  but also on the aspirations  $\underline{A}$  and  $\overline{A}$ , Table 12 distinguishes for the same improvement classification (the 8 rows) the average amounts of “burned money”  $P - A$  when positive, separately for boom,  $\overline{P}(i) - \overline{A}$ , and doom,  $\underline{P}(i) - \underline{A}$ . In T4, the (yes, yes, yes)-group of participants with “yes” for all three binary improvement criteria “burn”, on average, the least amount in boom and doom and phase 2. For T3, the effect is less pronounced.

Considering the high frequencies of “burning money” cases (see column frequencies  $P > A$  in Table 12 which are almost 100% in both boom and doom) improvement in satisficing occurs when burning less money in boom and doom. The same applies to “committing suicide”, based on  $\overline{P}(i) < \overline{A}$  or  $\underline{P}(i) < \underline{A}$ , i.e. non satisficing. Nearly all participants learn to avoid non-satisficing when transforming the visually displayed column height (see Fig. 4 in Appendix A) into numerical aspirations,  $\overline{A}$  and  $\underline{A}$ .

**Result 6** *Participants compliant with the three improvement criteria succeed in reducing “burning money” ( $P - A$ ), especially in treatment T4.*

<sup>15</sup> Thus, we obviously cannot refer separately to phases 1 and 2 and note the greater success in phase 2 than in phase 1.

<sup>16</sup> In the following analysis, as in the previous, compliance with a selected criterion is calculated based on 60% average compliance.

**Table 11** Success in T1 and T2

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 1											
			Phase 1						Phase 2					
			# subjects	% of the total	$P(i)$	SD	$P^* - P$	SD	# subjects	% of the total	$P(i)$	SD	$P^* - P$	SD
Yes	No	No	12	15.38	0.652	0.058	0.145	0.058	3	3.85	0.661	0.084	0.136	0.084
Yes	No	Yes	0	0.00	-	-	-	-	0	0.00	-	-	-	-
Yes	Yes	No	4	5.13	0.613	0.070	0.184	0.070	3	3.85	0.687	0.013	0.110	0.013
Yes	Yes	Yes	48	61.54	0.745	0.042	0.052	0.042	56	71.79	0.767	0.024	0.030	0.024
No	No	No	3	3.85	0.576	0.035	0.221	0.035	4	5.13	0.625	0.063	0.172	0.063
No	Yes	No	3	3.85	0.487	0.076	0.310	0.076	5	6.41	0.602	0.111	0.195	0.111
No	No	Yes	0	0.00	-	-	-	-	0	0.00	-	-	-	-
No	Yes	Yes	8	10.26	0.696	0.061	0.101	0.061	7	8.97	0.748	0.097	0.049	0.097
		Tot. subjects	78						78					

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 2											
			Phase 1						Phase 2					
			# subjects	% of the total	$P(i)$	SD	$P^* - P$	SD	# subjects	% of the total	$P(i)$	SD	$P^* - P$	SD
Yes	No	No	15	19.23	0.680	0.037	0.117	0.037	15	19.23	0.703	0.052	0.094	0.052
Yes	No	Yes	1	1.28	0.763		0.034		0.00	-	-	-	-	-
Yes	Yes	No	10	12.82	0.66	0.062	0.134	0.062	7	8.97	0.687	0.043	0.110	0.043
Yes	Yes	Yes	37	47.44	0.737	0.037	0.060	0.037	45	57.69	0.772	0.016	0.024	0.016

**Table 11** continued

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 2									
			# subjects	% of of the total	$P(i)$	SD	$P^* - P$	SD	# subjects	% of the total	$P(i)$	SD
No	No	No	3	3.85	0.590	0.051	0.207	0.051	1	1.28	0.576	0.221
No	Yes	No	1	1.28	0.514	0.283	0.283	1	1.28	0.590	0.292	
No	No	Yes	1	1.28	0.725	0.072	0.072	0	0.00	-	-	
No	Yes	Yes	10	12.82	0.748	0.034	0.049	0.034	9	11.54	0.752	0.045
Tot. subjects			78					78				0.073

Dummies with a threshold of 60%

**Table 12** Success in T3 and T4

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 3 (Boom)									
			Phase 1					Phase 2				
			# subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)	#subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)
Yes	No	No	6	8	0.287	0.092	98	4	5	0.092	0.161	97
Yes	No	Yes	0	0	–	–	0	0	0	–	–	0
Yes	Yes	No	4	5	0.290	0.149	98	3	4	0.205	0.210	98
Yes	Yes	Yes	18	25	0.272	0.147	94	24	33	0.287	0.192	96
No	No	No	13	18	0.322	0.067	99	14	19	0.305	0.084	100
No	Yes	No	13	18	0.320	0.096	98	9	12	0.335	0.114	98
No	No	Yes	2	3	0.270	0.069	100	0	0	–	–	0
No	Yes	Yes	17	23	0.334	0.093	98	19	26	0.33	0.108	98
			Total subjects 73									
Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 3 (Doom)									
			Phase 1					Phase 2				
			# subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)	#subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)
Yes	No	No	5	8	0.157	0.101	81	4	6	0.184	0.159	83
Yes	No	Yes	0	0	–	–	0	0	0	–	–	0
Yes	Yes	No	4	6	0.151	0.105	85	3	4	0.140	0.074	93
Yes	Yes	Yes	19	30	0.276	0.111	87	23	33	0.285	0.182	90

Table 12 continued

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 3 (Doom)									
			Phase 1					Phase 2				
			# subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)	#subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)
No	No	No	10	16	0.119	0.095	68	13	19	0.118	0.056	73
No	Yes	No	9	14	0.162	0.114	73	7	10	0.161	0.138	58
No	No	Yes	1	2	0.246	–	100	0	0	–	–	100
No	Yes	Yes	16	25	0.244	0.093	88	20	29	0.269	0.118	80
		Total subjects	64					70				
Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 4 (Boom)									
			Phase 1					Phase 2				
			# subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)	#subjects	% of the total	<i>P</i> – <i>A</i> if <i>P</i> > <i>A</i>	SD	Frequencies <i>P</i> > <i>A</i> (%)
Yes	No	No	14	19.72	0.309	0.095	98	11	15.49	0.302	0.125	100
Yes	No	Yes	0	0.00	–	–	0	1	0.322	–	–	0
Yes	Yes	No	10	14.08	0.266	0.130	99	9	12.68	0.226	0.134	98
Yes	Yes	Yes	13	18.31	0.213	0.150	99	17	23.94	0.087	0.102	99
No	No	No	22	30.99	0.339	0.076	98	16	22.54	0.336	0.139	98
No	Yes	No	5	7.04	0.309	0.053	84	10	14.08	0.362	0.131	97
No	No	Yes	0	0.00	–	–	0	0	0.00	–	–	0
No	Yes	Yes	7	9.86	0.304	0.109	99	7	9.86	0.180	0.183	99
		Total subjects	71					71				

**Table 12** continued

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 4 (Doom)									
			Phase 1					Phase 2				
			# subjects	% of the total	$P - A$ if $P > A$	SD	Frequencies $P > A$ (%)	#subjects	% of the total	$P - A$ if $P > A$	SD	Frequencies $P > A$ (%)
Yes	No	No	12	18.46	0.165	0.084	89	11	16.67	0.137	0.092	92
Yes	No	Yes	0	0.00	-	-	0	1	1.52	0.290	-	100
Yes	Yes	No	10	15.38	0.164	0.091	81	8	12.12	0.165	0.110	90
Yes	Yes	Yes	13	20.00	0.146	0.081	86	17	25.76	0.101	0.079	91
No	No	No	19	29.23	0.135	0.061	77	15	22.73	0.155	0.079	77
No	Yes	No	4	6.15	0.188	0.109	55	7	10.61	0.221	0.153	69
No	No	Yes	0	0.00	-	-	0	0	0.00	-	-	0
No	Yes	Yes	7	10.77	0.238	0.092	80	7	10.61	0.165	0.128	85
Total subjects			65					66				

Dummies with a threshold of 60%

## 7 Conclusions

Our previous analysis reported in Di Cagno et al. (2017) shows nearly exclusive sub-optimality, irrespective of whether optimality is interpreted as expected utility maximization or in the set-optimal sense of not being able to do better in boom without harming success in doom, or vice versa. Sub-optimality is revealed by positive differences  $i - i^*$  (as all  $i \leq i^*$  are at least set-optimal). Further, sub-optimality is revealed by treatments eliciting aspirations via “burning money”, which is still consistent with satisficing, and “committing suicide”, which violates satisficing.

From this starting point, we have identified different criteria for how well participants manage to reduce sub-optimality and improve their success, for example, by letting  $\max\{0, i - i^*\}$ ,  $\overline{P}(i) - \overline{A}$  and  $\underline{P}(i) - \underline{A}$ , when positive, shrink. One criterion applies to successive slider attempts before confirmation. Another improvement aspect is to react qualitatively to parameter changes in the optimal investment level  $i^*$  in the sense of timewise and casewise monotonicity. Finally, individual improving can be assessed since all 15 tasks are confronted twice, once in the first phase and once in the second phase.

We have used all criteria to demonstrate that most participants systematically reduce sub-optimality and thereby considerably enhance their success. Participants still do not behave optimally, but, being only boundedly rational, they learn how to improve decision-making. Actually nearly all of them become satisficers.<sup>17</sup> Even for those participants who still score poorly in terms of the improvement criteria, it seems possible that improving how to use the slider and listen to advice, for example, that  $\max\{0, i - i^*\}$  is sub-optimal, will also help them improve their decision-making.

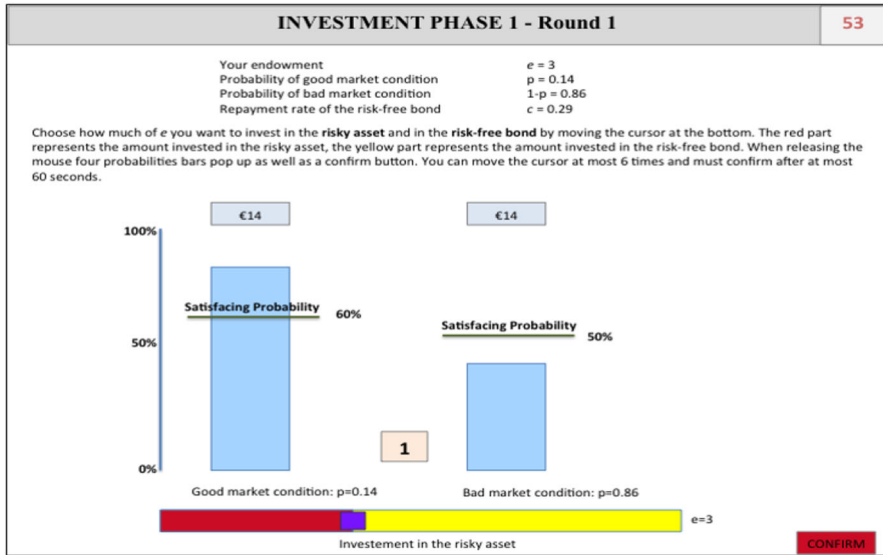
We hope to explore in future research if some sort of advice could speed up behavioral improvements and possibly avoids sub-optimality altogether.

## Appendix

See Fig. 4 and Tables 13, 14.

---

<sup>17</sup> We admit that our method of incentivizing aspiration formation by letting participants lose all chances of earning more €14, rather than only €4 is partly responsible for the striking confirmation of satisficing, at least in phase 2 of treatments T3 and T4.



**Fig. 4** Investment choice in T3 and T4

**Table 13** Summary results regarding slider adjustments

Variable	Phase 1			Phase 2		
	Obs	Mean	SD	Obs	Mean	SD
$i_{t^*} - i^*$	4530	0.551	0.745	4530	0.413	0.688
$i_{t^*-1} - i^*$	4516	0.640	0.841	4523	0.486	0.782
$i_{t^*-2} - i^*$	3484	0.659	0.824	3209	0.477	0.760
$i_{t^*-3} - i^*$	2888	0.732	0.828	2527	0.533	0.787



**Table 14** Non-improving frequency by improvement criteria in phases 1 and 2

Correct slider use	Timewise monotonic	Casewise monotonic	Treatment 1		Treatment 2		
			(No, no)	(Yes, no)	(No, no)	(Yes, no)	
No	No	No	Obs	1	1	1	2
			Mean	3.00	6.00	1.00	10.00
			SD	-	-	-	0.71
No	Yes	Yes	Obs	-	-	-	1
			Mean	-	-	-	7.00
			SD	-	-	-	-
No	Yes	No	Obs	1	1	1	1
			Mean	6.00	4.00	3.00	6.00
			SD	-	-	-	-
No	Yes	Yes	Obs	1	1	7	2
			Mean	9.00	9.00	2.71	3.50
			SD	-	-	2.93	2.12

**Table 14** continued

		Treatment 1				Treatment 2			
Correct slider use		Timewise monotonic				Casewise monotonic			
		(Correct slider use, Timewise monotonic, Casewise monotonic)							
		Phase 1				Phase 2			
		(No, no)	(No, yes)	(Yes, no)	(Yes, yes)	(No, no)	(No, yes)	(Yes, no)	(Yes, yes)
Yes	No	2	1	2	1	4	5	4	6
		3.00	3.00	4.50	5.00	4.00	5.60	7.25	2.17
		1.41	-	0.71	-	2.58	1.52	2.22	1.94
Yes	Yes						1		
							8.00		
Yes	Yes						-		
Yes	No				1	3	1	1	3
					5.00	4.00	4.20	9.00	4.67
						1.00	4.09		0.58
Yes	Yes	1	4	1		41	1		31
		8.00	5.00	8.00		2.68	7.20	7.00	2.74
			3.37	-		2.25	2.77	-	2.03

Dummies with 60% threshold

## References

- Conlisk, J. (1996). Why bounded rationality? *Journal of Economic Literature*, 34(2), 669–700.
- Conte, A., Di Cagno, D. T., & Sciubba, E. (2015). Behavioral patterns in social networks. *Economic Inquiry*, 53(2), 1331–1349.
- Di Cagno, D. T., Galliera, A., Güth, W., Marzo, F., & Pace, N. (2017). (Sub) Optimality and (non) optimal satisficing in risky decision experiments. *Theory and Decision*, 83(2), 195–243.
- Erev, I., & Haruvy, E. (2013). Learning and the economics of small decisions. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics*. Princeton University Press.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.
- Gigerenzer, G. (2006). Bounded and rational. In R. J. Stainton (Ed.), *Contemporary debates in cognitive science Oxford*. UK: Blackwell.
- Greiner, B. (2004). An online recruitment system for economic experiments. Published. In: Forschung und Wissenschaftliches Rechnen 2003. GWDG Bericht, 63, 79–93.
- Harrison, G.W., & Johnson, L.T. (2006). Identifying altruism in the laboratory. In R.M. Isaac, & D.D. Davis (Eds.) *Experiments investigating fundraising and charitable contributors* (Research in Experimental Economics, Volume 11). Emerald Group Publishing Limited.
- Hey, J. D., Permana, Y., & Rochanahastin, N. (2017). When and how to satisfice: an experimental investigation. *Theory and Decision*, 83(3), 1–17.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 47(2), 263–291.
- Manski, C. F. (2017). Optimize, satisfice, or choose without deliberation? A simple minimax-regret assessment. *Theory and Decision*, 82(2), 1–19.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.
- Roth, A. E., & Malouf, M. W. (1979). Game-theoretic models and the role of information in bargaining. *Psychological Review*, 86(6), 574–594.
- Sauermann, H. and Selten, R., (1962). Anspruchsanpassungstheorie der unternehmung. *Zeitschrift für die gesamte Staatswissenschaft/Journal of Institutional and Theoretical Economics*, (H. 4), 577–597.
- Selten, R. (1998). Aspiration adaptation theory. *Journal of Mathematical Psychology*, 42(2–3), 191–214.
- Selten, R. (2001). What is bounded rationality. In G. Gigerenzer & R. Selten (Eds.), *Bounded Rationality: The Adaptive Toolbox*. The MIT Press.
- Selten, R., Pittnauer, S., & Hohnisch, M. (2012). Dealing with dynamic decision problems when knowledge of the environment is limited: an approach based on goal systems. *Journal of Behavioral Decision Making*, 25(5), 443–457.
- Selten, R., Sadrieh, A., & Abbink, K. (1999). Money does not induce risk neutral behavior, but binary lotteries do even worse. *Theory and Decision*, 46(3), 213–252.
- Siegel, S. (1957). Level of aspiration and decision making. *Psychological Review*, 64(4), 253–262.
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118.