Uncovering Rationality

Thomas F. Epper[§] and Ivan Mitrouchev[†]

[§]CNRS, IESEG School of Management, Univ. Lille, UMR 9221 – LEM - Lille Economie Management, F-59000 Lille, France. thomas.epper@cnrs.fr. ORCID n°0000-0002-0826-4997.
[†]Université Lumière Lyon 2, CNRS, UJM Saint-Etienne, emlyon business school, GATE, 69007 Lyon, France. ivan.mitrouchev@univ-lyon2.fr. ORCID n°0000-0002-8960-4550.

January 28, 2026*

Preliminary version: please do not quote or share without permission

Abstract

We study individuals' adherence to rationality principles by eliciting their risk, time, and social preferences, and by giving them the opportunity to reconsider their choices. We also investigate subjects' rationales, focusing on the reasons they provide for changing or maintaining their choices. Our main results show that roughly 75% of subjects behave in accordance with the rationality principles under study. Among those who revise their choices, revisions tend to lead to a higher degree of rationality than is observed among subjects who uphold their initial choices. However, this greater adherence to rationality principles is not driven by subjects correcting apparent mistakes. Overall, fewer than 10% of subjects across all domains report having made any mistakes. Contrary to the hypothesis of true preference in behavioral welfare economics, our results provide no empirical support for the idea that reconsideration primarily reveals pre-existing true preferences. Instead, the hypothesis of *constructed preferences*—according to which preferences are shaped through learning and the process of choice itself—is more consistent with our results.

Keywords. *choice – rationality – risk – social – time*

JEL codes. D63, D91

Statements and declarations. Thomas F. Epper acknowledges funding from the Métropole Européenne de Lille (MEL). There are no competing interests.

*This project was presented at the 2024 Workshop on Social Choice and Fairness in Vienna, the 2024 Meeting of the Society for Social Choice and Welfare in Paris, the 2024 International Conference of the French Association of Experimental Economics in Grenoble, the 2024 Economic Philosophy International Conference in Reims, the "Behavioral and Experimental Economics" seminar at the Centre for Environmental Economics in Montpellier, and the 2025 Society for the Advancement of Behavioral Economics Annual Conference in Trento. We are grateful to the audiences for their comments. We thank Kene Boun My and Solal Azoulay for their assistance in conducting the lab sessions at the Laboratory for Experimental Economics in Strasbourg. We are particularly thankful to Paolo Crosetto, Dimitri Dubois, Marc Fleurbaey, Gerd Gigerenzer, Nicolas Jacquemet, Brice Magdalou, Rustam Romaniuc and Karl Schlag for their valuable feedback. All remaining errors are our own.

"A mode of behavior is rational for a given person if this person feels comfortable with it, and is not embarrassed by it, even when it is analyzed for him." (Gilboa, 2010, p. 5)

1 Introduction

Rationality is typically defined in decision theory as a set of preference rules about what constitutes a "good" choice. The axiomatic characterization of rationality has been influential in models of choice under risk, especially through expected utility theory (von Neumann and Morgenstern, 1944; Savage, 1954), where classic empirical challenges (Allais, 1953; Ellsberg, 1961) motivated early experimental investigations of whether individuals adhere to rationality principles (MacCrimmon, 1968; Moskowitz, 1974; Slovic and Tversky, 1974; MacCrimmon and Larsson, 1979). This line of research has been recently renewed with incentivized experiments in the domains of choice under risk (Benjamin et al., 2020; Nielsen and Rehbeck, 2022; Breig and Feldman, 2024; Herweg et al., 2024) and social preferences (Andersson et al., 2023). While these studies differ with respect to the rationality principles being investigated, the domains of choice (risk vs. social), and the methods for measuring adherence to the rationality principles, a common feature of these works is that revisions aligning with rationality principles are interpreted as evidence of initial mistakes.

For example, Nielsen and Rehbeck (2022) presume that "when an individual revises their axiom and lottery choices to be consistent, this reveals that the original choice was a mistake" (p. 15). Similarly, Breig and Feldman (2024) presume that a modified choice, absent informational changes, reveals the original choice to have been mistaken. Benjamin et al. (2020) adopt a related stance, suggesting that revised choices are closer to individuals' "true" preferences, as if made under idealized conditions of deliberation and free from cognitive error. The idea that people make mistakes because of limited information, cognitive ability, and self-control, relates to behavioral paternalism (Camerer et al., 2003; Thaler and Sunstein, 2003, 2009) and behavioral welfare economics (Bernheim and Rangel, 2007, 2009).[1] Much of this work rests on the notion of a "true" preference—what individuals would have chosen absent cognitive distortions.[2] Although this hypothesis is practically appealing and well-aligned with the standard paradigm of rational choice, it has been questioned for being both psychologically and philosophically problematic (Infante et al., 2016; Sugden, 2022). More concerning to us, experimental setups—which constitute artificial environments fully controlled by the experimenter—have, with the exception of Benjamin et al. (2020), not incorporated subjects' explicit acknowledgment of mistakes. Rather, they rely on definitions of error inferred from observed behavioral data, without directly asking individuals whether they consider themselves to have made mistakes.

[1] See also Köszegi and Rabin (2007); Salant and Rubinstein (2008); Dalton and Ghosal (2012); Rubinstein and Salant (2012); Manzini and Mariotti (2014); Bernheim (2016)—among others.

[2] See in particular Harsanyi (1977) who adopts this concept for welfare evaluation in his defense of utilitarianism, and Fine (1995), who distinguishes the concepts of true preference and actual choice from a social choice perspective. Some philosophers also defended the satisfaction of self-interested "informed," "rational," or "laundered" preferences as what constitutes goodness (Gauthier, 1986; Arneson, 1990; Goodin, 1992; Railton, 2003).

In this paper, we take a broader perspective on the interpretation of rational behavior by introducing a novel experimental setup in which subjects are given the opportunity to reconsider some of their choices and to explain *why* they wish—or do not wish—to modify them. In other words, rather than restricting our analysis to adherence to rationality principles, and assuming that deviations necessarily reflect initial mistakes, we examine subjects' *rationales*, that is, their reasons, justifications, or intentions underlying the choices they made. To illustrate the rather restrictive hypothesis that deviations from rationality principles necessarily reflect mistakes, consider the definition of a "mistake" in Herweg et al. (2024), which is a case in which "the respective subject changed her lottery choice to be consistent with that axiom and she additionally stated a preference in favor of the axiom" (p. 2). This interpretation does not allow for the reverse possibility—that a subject behaves consistently toward an axiom but considers it a mistake to explicitly endorse conformity with that axiom—nor does it allow for other possibilities, such as simply changing her mind, lacking any fixed rule of behavior, or updating her preferences during the choice process itself. In this sense, our approach to rationality is *subjective*, in line with Gilboa (2010), and stands in contrast to the more traditional perspectives that conceive rationality as a fixed set of objective rules governing behavior (Savage, 1954).

To explore people's subjective rationality, our experimental design combines a quantitative choice-evaluation phase with qualitative reports, enabling us to study both adherence to formal rationality principles and the different rationales subjects provide for their choices. Our main results are as follows. A large majority of subjects make choices that are consistent with the rationality principles under study. Among those who deviate, a substantial proportion revise their choices, typically achieving a higher degree of consistency than those who uphold their initial choices. This pattern indicates adherence to the rationality principles under study and is broadly consistent with the related literature. However, our key finding is that this adherence *does not* arise from subjects having made mistakes in their initial choices. Overall, fewer than 10% of subjects (all domains included) explicitly acknowledge having made any mistakes. While our study finds normative appeal for most of the rationality principles we measure across different domains (risk, time, and social), we go further by examining the sources of these revisions. Our results show that the modal reason among revising subjects is "I changed my mind," whereas "I made a mistake" is selected by only a minority of revising subjects. We provide additional results on subjects' strategies throughout the choice tasks, as well as qualitative reports in which some subjects reveal their heuristics. Overall, our results are more consistent with the constructed preference hypothesis (Slovic, 1995; Lichtenstein and Slovic, 2006), according to which preferences are formed during the elicitation process through learning, discovery and adaptation.

The rest of the paper is organized as follows. Section 2 presents our theoretical framework. Section 3 details our experimental design. Section 4 reports the results. Section 5 reviews related experiments on choice reconsideration and explains how our approach differs. Section 6 discusses the implications of our study, outline some limitations and identify directions for further research.

2 Theoretical Framework

We model the experimental decision problem as a choice situation C, where $\mathcal{C}$ denote the set of all possible choice situations. Each choice situation $C \in \mathcal{C}$ is represented by the tuple $C = (\Omega, r)$, where $\Omega = \{\omega, \omega'\}$ is the binary set of mutually exclusive contingencies, and $r : \Omega \to \mathbb{R}$ specifies the return (payoff) associated with each contingency. Depending on the domain, the two contingencies either correspond to: (i) the two possible states of nature of a fair coin flip (risk), $\Omega = \{\mathrm{H}, \mathrm{T}\}$ ("Heads" vs. "Tails"), (ii) two payment dates (time), $\Omega = \{\mathrm{E}, \mathrm{L}\}$ ("Early" vs. "Late"), or (iii) two persons (social), $\Omega = \{\mathrm{S}, \mathrm{O}\}$ ("Self" vs. "Other"). We allow returns to coincide or differ across contingencies (details follow in Section 3). We test four rationality principles. The first, "Test-Retest", applies across all domains (risk, time, and social). The remaining three are domain-specific: "Symmetry" (risk), "Stationarity" (time), and "Anonymity" (social).

Test-Retest (RETEST). For any two allocations $z, z' \in Z$, if $z \succsim z'$ when the choice situation C is presented on one occasion, then $z \succsim z'$ when the identical choice situation is presented again. That is, preferences should be stable across repetitions of the same choice situation. This principle is a consistency requirement for preferences across identical choice situations.

Symmetry (SYM). Consider the risk domain, with $\Omega = \{\mathrm{H}, \mathrm{T}\}$ and return vector $r = (r_{\mathrm{H}}, r_{\mathrm{T}})$. Let $r^{\mathrm{swap}} = (r_{\mathrm{T}}, r_{\mathrm{H}})$ denote the contingency-wise permutation that swaps the returns associated with H and T. For any two allocations $z, z' \in Z$, if $z \succsim^{r} z'$, then $z \succsim^{r^{\mathrm{swap}}} z'$. That is, one's preference over allocations should remain unchanged when the returns associated with the two coin-flip contingencies are swapped.

Stationarity (STAT). For any two allocations $z, z' \in Z$ and any two payment dates t_1, t_2, if $z(t_1) \succsim z'(t_2)$, then $z(t_1 + d) \succsim z'(t_2 + d)$, where $d > 0$ is a fixed delay. That is, preferences between two allocations should remain the same when both are shifted forward in time by the same delay. Equivalently, the preference for earlier versus later gratification should depend on the temporal distance between delivery dates, but not on their absolute calendar timing.

Anonymity (ANON). Consider the social domain, with $\Omega = \{\mathrm{S}, \mathrm{O}\}$ and return vector $r = (r_{\mathrm{S}}, r_{\mathrm{O}})$. Let $r^{\mathrm{swap}} = (r_{\mathrm{O}}, r_{\mathrm{S}})$ denote the permutation that swaps the returns assigned to self and other. For any two allocations $z, z' \in Z$, if $z \succsim^{r} z'$, then $z \succsim^{r^{\mathrm{swap}}} z'$. This principle mirrors SYM but is adapted to identity contingencies in the social domain. It states that one's preference over allocations should remain unchanged when the assignment of payoffs to the two individuals is swapped. In other words, aversion towards inequality should depend on the payoff difference only, but not on who is better or worse off.

The intuition behind SYM is that individuals should have no intrinsic preference over the two coin-flip contingencies, which are complementary outcomes of a fair randomization device (here, a coin). The key feature is that the probability distribution (50%, 50%) is known to the decision-maker and held fixed, while only the mapping from contingencies to returns is permuted. Since both contingencies occur with the same probability, the idea of SYM is that conditioning on either contingency should not affect one's pref-

erence.[3] STAT is one of the central principles in intertemporal choice models, as defined by Koopmans (1960), and is violated in hyperbolic or quasi-hyperbolic models (Laibson, 1997).

The intuition behind ANON is less obvious in practice but theoretically appealing. It is inspired by social choice theory and theories of justice, where choice under uncertainty is adapted to social distributions (Karni, 1998; Mongin, 2001). In particular, an individual is placed under Harsanyi's (1953; 1955) situation of an impartial spectator, where she must decide on a social allocation before learning which identity contingency obtains (i.e., whether she will be the "Self" or the "Other" position). She understands that if there are n individuals in the society, she has a $1/n$ probability of occupying any given position. Although our experiment was not designed to place subjects under the conditions of Harsanyi (1953, 1955), it aimed to test a form of anonymity principle, according to which an individual might be indifferent to swapping her own payoff with that of the other person if what matters to her is solely the level of inequality between the two. Note that such a principle would more intuitively apply to an egalitarian rather than a non-egalitarian, as swapping the allocations chosen by an egalitarian would yield the same outcome as the original allocation.

3 Experimental Design

The sessions were conducted at the Laboratory for Experimental Economics of Strasbourg in France, in the spring of 2024. We recruited a total of 378 subjects from the subject pool of the University of Strasbourg.[4] We employed a between-subject design, with 128 subjects assigned to the risk task, 125 to the time task, and 125 to the social task. Each session lasted approximately one hour. Subjects received a flat participation fee of €12, in addition to a variable incentive payment, ranging from €0 to €24, depending on their choices. All payments were made in cash.

We implemented the PRINCE method (Johnson et al., 2021), in which the payment-relevant choice situation is predetermined. That is, the choice situation to be paid was randomly selected *before* the experiment began. In practice, we prepared 24 opaque and sealed envelopes, each containing a number from 1 to 24, corresponding to the ID of a choice task. At the start of each session, we invited a volunteer to draw a number that determined one's cubicle seat. The subject assigned to that cubicle then randomly selected one of the 24 sealed envelopes. We took this envelope and attached it visibly to the wall for the duration of the entire experiment. At the end of the session, we opened the envelope, and subjects entered the ID of the choice task on their screen to see their actual choice in this task, with the corresponding payment amount. Doing so, the procedure was fully randomized and repeated at the beginning of each session.

[3] A violation of SYM rejects more than just expected utility. In fact, rank-dependent utility also predicts symmetric indifference curves for equiprobable events. The only difference is the presence of a kink at the certainty line.

[4] 89% of our subjects were students in various disciplines (medicine, computer sciences, biology, geography, sociology, political sciences, law, management, psychology, economics, literature, among others). The remainder 11% were non-students. Detailed sample characteristics are available upon request.

For uncertainty resolution across the risk tasks, we flipped a coin at the center of the experimental room and asked subjects to observe and confirm the event ("Heads" or "Tails"). Subjects were then immediately paid according to their allocation in the selected choice task. For the time tasks, the screen displayed the total amount to be collected immediately and the amount available either two or four weeks later. If the subject had chosen in the selected choice task to collect some money at a later point in time (two or four weeks later), she was informed that she could collect it at the given time period in the lab manager's office, which is located in the same building as the Laboratory for Experimental Economics of Strasbourg. We ensured that the later periods did not overlap with any holidays or vacations. To control for present-biased choices driven by external factors, the post-experimental questionnaire asked subjects whether they expected to be available and able to collect their payments at the earlier and the later payment dates. For the social tasks, the screen showed the total amount allocated to oneself and the total amount allocated to another (anonymous) person. The information we provided about the other person was that she was another participant in the larger project, but not in this specific experiment. This anonymity statement was meant to eliminate any noise related to the characteristics of the other individual, as well as other factors that may influence one's distribution, e.g., reciprocity or punishment. Appendix A provides the full set of instructions for risk tasks, as they appeared on subjects' screens.

Our experiment is split into three phases. In the "Choice Tasks" phase (Section 3.1), we elicit subjects' preferences in their respective domain (risk, time, or social). In the "Choice Reconsideration" phase (Section 3.2), we show them some of their previous choices and give them the opportunity to change them. We also collect their reasons for changing or not their choices. In the "Quality Reports" phase (Section 3.3), we ask them to provide us with some qualitative reports about their previous choices, including the strategies they may have had during the choice tasks.

3.1 Choice Tasks

To measure subjects' preferences, we used a series of 24 money allocation tasks in each domain. The choice situations are listed in Table 1.

Table 1: Choice situations

ChoiceID	Tokens	Risk		Time				Social	
		r_H	r_T	Early	Late	r_E	r_L	r_S	r_O
1	10	**1.80**	**0.60**	**0**	**2**	**1.20**	**2.30**	**1.80**	**0.60**
2	10	**1.80**	**0.90**	0	2	1.20	1.20	**1.80**	**0.90**
3	10	2.40	0.60	0	2	0.60	2.40	2.40	0.60
4	10	1.15	2.25	**0**	**2**	**0.90**	**1.80**	1.15	2.25
5	10	**1.20**	**1.20**	0	2	0.60	1.80	**1.20**	**1.20**
6	10	0.60	2.40	0	2	0.80	1.40	0.60	2.40
7	10	2.25	1.15	**0**	**2**	**1.20**	**1.70**	2.25	1.15
8	10	**0.90**	**1.80**	0	2	1.00	1.20	**0.90**	**1.80**
9	10	1.50	1.35	0	2	0.70	0.90	1.50	1.35
10	10	**0.60**	**1.80**	0	2	0.80	0.90	**0.60**	**1.80**
11	10	1.80	0.75	2	4	1.20	2.30	1.80	0.75
12	10	0.75	1.80	2	4	1.20	1.20	0.75	1.80
13	10	2.40	0.70	2	4	0.60	2.40	2.40	0.70
14	10	0.70	2.40	**2**	**4**	**0.90**	**1.80**	0.70	2.40
15	10	1.60	1.60	2	4	0.60	1.80	1.60	1.60
16	10	0.70	2.40	**2**	**4**	**0.80**	**1.40**	0.70	2.40
17	10	**2.40**	**0.70**	**2**	**4**	**1.20**	**1.70**	**2.40**	**0.70**
18	10	1.60	1.35	2	4	1.00	1.20	1.60	1.35
19	10	1.35	1.60	2	4	0.70	0.90	1.35	1.60
20	10	0.70	1.80	2	4	0.80	0.90	0.70	1.80
21	10	1.80	0.90	0	2	0.80	1.40	1.80	0.90
22	10	**1.20**	**1.20**	**0**	**2**	**1.20**	**2.30**	**1.20**	**1.20**
23	10	0.90	1.80	**2**	**4**	**0.80**	**1.40**	0.90	1.80
24	10	**2.40**	**0.70**	2	4	1.20	2.30	**2.40**	**0.70**

Note. r_H and r_T indicate the returns per token for the events “Heads” and “Tails”. “Early” and “Late” indicate the time periods, where "Early" either takes the value of 0 (now) or 2 (two weeks), and where "Late" either takes the value of 2 (two weeks) or 4 (four weeks). r_E and r_L indicate the returns per token for "Early" and "Late". Note that we imposed $r_E \leq r_L$ to compensate for positive time discounting. r_S and r_O indicate the returns per token for the subject (Self) and the anonymous person (Other). Each graph, depicting the possible allocations in each domain, is provided in Appendix B.

The values in bold correspond to the parameters we selected to measure the adherence to RETEST, SYM, STAT and ANON in the “Choice Reconsideration” phase.[5] The calibration was conducted according to the following criteria. First, we needed a sufficiently high number of choice tasks to obtain a precise measure of subjects’ preferences in their respective domains (risk, time, social), but not so many as to compromise subjects’ attention. Second, we ensured that the set of choice tasks used to measure adherence to rationality principles were sufficiently spaced apart to prevent subjects from easily

[5]As Table 1 shows, there could have been other pairs to explore the adherence of the rationality principles in each domain. For example, the pair (2, 21) could test the adherence of RETEST in the risk and social domains. However, our experimental setup required making a choice within the selected parameters, while testing all possibilities would have increased the number of tasks in the “Choice Reconsideration” phase. This could have led to subject fatigue and, consequently, increased noise in responses. Note also that we used the same calibration for the risk and social domains, as SYM and ANON are equivalent.

tracking them.[6] Figure 1 shows the risk interface as it appears on subjects' screens.

Figure 1: Risk | Choice interface before allocation

Note. In each choice situation, subjects had to exhaustively allocate 10 tokens between the two boxes. They could either use the drag-and-drop function of the mouse or the "$+$" and "$-$" buttons. Each box represented the associated contingency related to its respective domain. For risk, we depicted the "Heads" and "Tails" sides of a €1 French coin. For time, the label of each box corresponded to two different points in time: either now (Early) vs. two weeks (Late), or two weeks (Early) vs. four weeks (Late). For social distribution, we used the labels "for myself" (Self) and "for the other" (Other).

Before subjects began the choice tasks, they first completed a tutorial to familiarize themselves with the interface and understand how it worked by moving tokens from one box to another. At the end of the tutorial, they had to answer a comprehension question and could not proceed until they provided the correct answer.[7] An example choice situation, corresponding to ChoiceID No. 1 in Table 1 for risk, is depicted in Figure 2. For illustrative purposes only, the individual decision-maker chooses to allocate four tokens to "Heads" with $r_{\text{H}} = 1.8$, and six tokens to "Tails", with $r_{\text{T}} = 0.6$.

[6] More specifically, we arranged the choice situations in blocks, where the first block contained choice situations that were repeated in the last block. We then presented choice situations in each block in individual random order.

[7] In the tutorial interface, and for all domains, we deliberately used the returns $r_{\omega} = 1.2$ and $r_{\omega'} = 0.8$, which are different from those in the 24 choice tasks listed in Table 1.

Figure 2: Risk | Choice interface after allocation

Note. As instructed to the subjects, each token transforms into the value of the given return when allocated to a contingency (here, "Heads" and "Tails"). This means that in each choice situation, they must first attempt to distribute some tokens between the two boxes to know the returns in the different contingencies, then choose their preferred allocation, and then click the "Continue" button to confirm their choices. The total value of the allocated tokens (in this example, €7.2 for "Heads" and €3.6 for "Tails") appears at the top of each box. To avoid a framing effect, tokens were returned to the initial endowment box when clicking on "−". When tokens were exhaustively allocated between both boxes and the subject pressed the "+" button, it increased the allocation in one box by removing a token from the other. Note that an earlier version of the interface included scaling of the tokens based on their value. We decided to remove this feature in the final version, as we believed it could influence responses toward the higher return.

Our design offers several advantages. First, the interface is presented without a preselected default, thereby limiting potential anchoring effects. Second, the interface is fully adaptable across every contingency and domain, requiring only a change in the labels at the top of each box.[8] Third, the task is extremely simple, requiring subjects to do nothing other than indicating their preferred distribution in each choice task. This contrasts with other tasks that may involve trade-offs between contingencies of different nature, such as outcome vs. probability (Breig and Feldman, 2024), which demand greater cognitive effort. The only cognitive effort required in our experimental setup is focusing on the varying returns across the 24 choice tasks, which subjects were explicitly instructed to pay close attention to during the tutorial phase (Appendix A).

[8]For example, it can also be adapted to ambiguity, where subjects distribute tokens based on their beliefs about the likelihood of two events, e.g., whether it will rain tomorrow or be sunny. A common feature in the preference elicitation literature is the use of binary contingencies. Still, the interface can be adapted to more than two contingencies, with each box corresponding to one.

3.2 Choice Reconsideration

After subjects completed the choice tasks, we instructed them that we would present some of the choices they had just made and ask them questions about them. Subjects knew from the instructions at the beginning of the experiment that they would be asked questions about their choices, but they did not know what the actual questions would be. That is, they did not know that we would present some of their previous choices and ask whether they wanted to modify them. This approach helps maintain response quality by preventing subjects from anticipating the questions, which could lead to response bias during the choice tasks. At the same time, it remains in line with the ethical standards of experimental economics, ensuring that subjects are fully informed from the beginning about the overall content and structure of the experiment. In any case, since the "Choice Reconsideration" phase is the core of our experiment, it is essential not to reveal to subjects what exactly is being measured, as is standard in exploratory experimental studies. In order to mitigate the experimenter demand effect, we emphasized, once again, and in bold, that "there are no right or wrong answers" and that "we are simply interested in how people make choices".

Adherence to the rationality principles we control for, namely RETEST, SYM, STAT and ANON, is revealed by showing subjects their previous choices and providing them with the opportunity to change them. If the subject is shown two of her previous choices that are not aligned with the given principle, and decides to change her choice by conforming to the given principle, it provides us with evidence of the subject's *adherence* to it. On the contrary, if the subject is shown two of her previous choices that are not aligned with the given principle, and decides to maintain her choices, it provides us with evidence of the subject's *rejection* for it.[9]

We draw this approach from Benjamin et al. (2020), recognizing at least one caveat. In particular, one may argue that actual preferences over the rationality principles that are being controlled for are not, strictly speaking, elicited through this method because the subject is never explicitly shown the principles nor asked whether she prefers to conform to them. By contrast, the approach of Nielsen and Rehbeck (2022) is closer to elicit those preferences, where subjects are being presented some rationality principles in a simple format (depicted balls) and asked if they want to delegate their choices to a third party by following those rules of behavior. This is why we say *adherence* instead of *preference* over the rationality principles. Still, we argue in Section 5 that our approach, in our view, has more advantages than drawbacks compared to the other methods used in related experiments. Furthermore, a particularity of our design is that the "Choice Reconsideration" phase is shown to *all* subjects, whereas most of the related literature directs this phase at subjects who chose against the rationality principles being controlled

[9]We emphasize that this identification strategy concerns only adherence to or rejection of a given rationality principle. It does not make any claim about whether the subject made a mistake. Note also that this identification strategy relies on the assumption that the *second* choice—made during the reconsideration phase—reveals subjects' adherence to or rejection of the rationality principle under consideration. From a philosophical perspective, see Mitrouchev and Buonomo (2024), who study which choice should be taken as normatively relevant in intertemporal decision-making, and emphasize that no single answer clearly dominates. In our setting, privileging the second choice is justified by the information structure of the experiment, where subjects are being shown their earlier choices and are given the opportunity to revise them. We interpret this opportunity as a form of self-assessment, which closely relates to the psychological literature of meta-cognition.

for.[10] In addition, our setup allows measuring rationality adherence in degree, with more or less deviation depending on the absolute difference of token allocation, i.e., not as a binary response.

In order to measure subjects' adherence or rejection regarding RETEST, we randomly selected two ChoiceIDs such as $C = C'$ (same contingencies ω, ω', and same returns $r_\omega, r_{\omega'}$) and gave them the opportunity to change their choices. In order to measure subjects' adherence or rejection regarding SYM, STAT, and ANON, we randomly selected two ChoiceIDs that allowed for such a test, and gave subjects the opportunity to change their choices. In particular, for SYM, we randomly selected two ChoiceIDs such that $r^{\text{swap}} = \{r_{\text{T}}, r_{\text{H}}\}$ is the flipped return configuration of $r = \{r_{\text{H}}, r_{\text{T}}\}$. For STAT, we randomly selected two ChoiceIDs with equal returns r_{E} and r_{L}, but where the values of "Early" and "Late" are shifted over time by two weeks. As for ANON, we randomly selected two ChoiceIDs such that $r^{\text{swap}} = \{r_{\text{O}}, r_{\text{S}}\}$ is the flipped return configuration of $r = \{r_{\text{S}}, r_{\text{O}}\}$. The list of the ChoiceID pairs selected to test adherence to all rationality principles, along with their corresponding returns (and dates, for time only), is described in Table 2.

Table 2: ChoiceID pairs selected to test adherence to rationality principles

ChoiceIDs	**Returns**	**Dates**	**Rationality principles**		
			Risk	**Time**	**Social**
(5, 22)	(1.2 – 1.2)	–	RETEST	–	RETEST
(17, 24)	(2.4 – 0.7)	–	RETEST	–	RETEST
(2, 8)	(1.8 <> 0.9)	–	SYM	–	ANON
(1, 10)	(1.8 <> 0.6)	–	SYM	–	ANON
(1, 22)	(1.2 – 2.3)	0 ↔ 2	–	RETEST	–
(16, 23)	(0.8 – 1.4)	2 ↔ 4	–	RETEST	–
(7, 17)	(1.2 – 1.7)	0 ↔ 2, 2 ↔ 4	–	STAT	–
(4, 14)	(0.9 – 1.8)	0 ↔ 2, 2 ↔ 4	–	STAT	–

In total, each subject faced four "Choice Reconsideration" situations, appearing in the order described in Table 3.

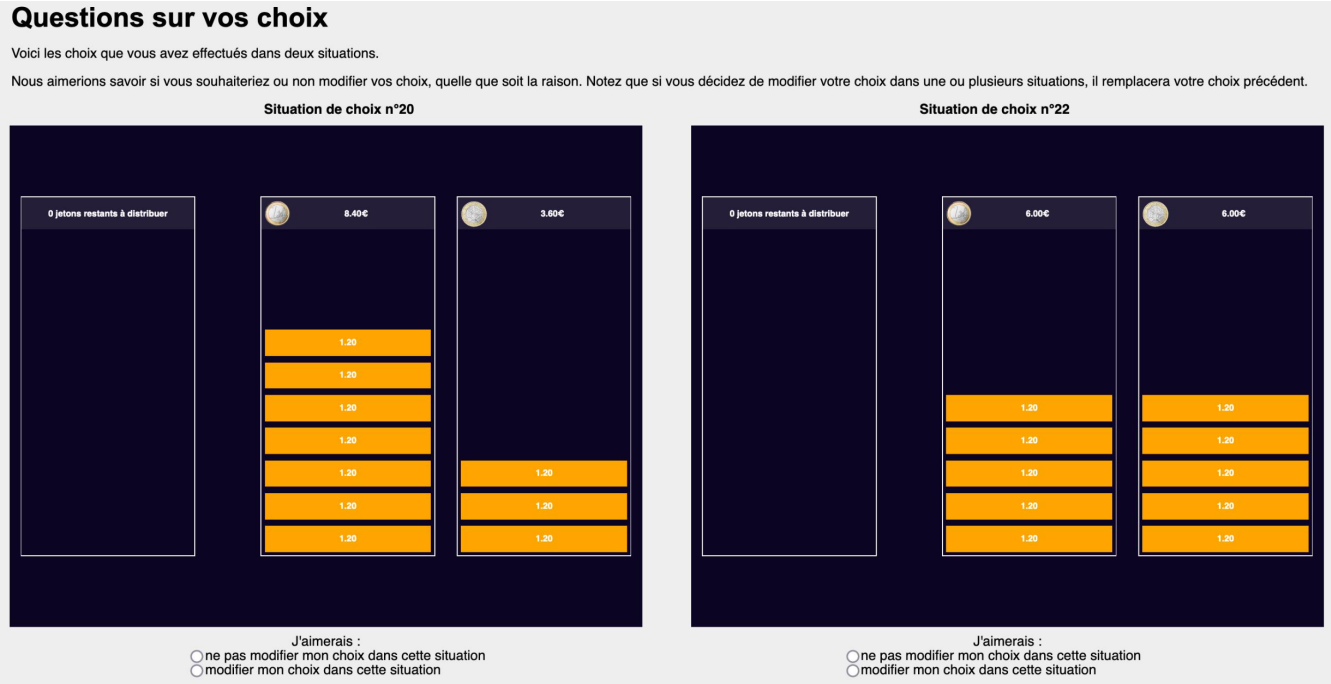
Table 3: Order of appearance of the rationality principles

Domain	**Order of appearance**			
	1	**2**	**3**	**4**
Risk	RETEST	SYM	RETEST	SYM
Time	RETEST	STAT	RETEST	STAT
Social	RETEST	ANON	RETEST	ANON

Figures 3 and 4 show examples of the choice reconsideration interface in the risk domain for RETEST and SYM (respectively), as it appears on subjects' screens. We provide further screenshots of the time and social domains in Appendix C.

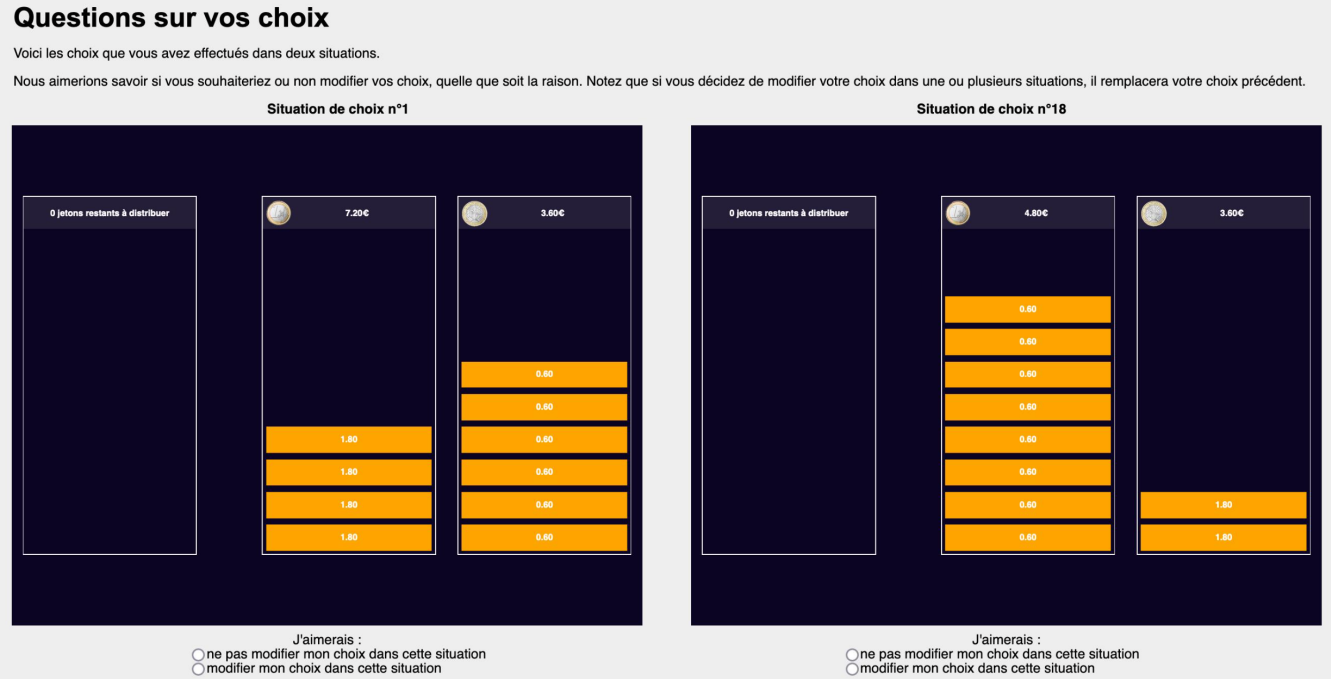
[10] An exception is Herweg et al. (2024).

Figure 3: Risk | Choice reconsideration interface for RETEST



Note. These two choice situations, corresponding to the (5, 22) pair in Table 2, are identical. They both yield returns (1.2 – 1.2). In this example, the violation of RETEST is weak, with an absolute difference of allocation of 2 tokens.

Figure 4: Risk | Choice reconsideration interface for SYM



Note. These two choice situations, corresponding to the (1, 10) pair in Table 2, allow to measure subjects' adherence or rejection of SYM, as their configuration is (1.8 <> 0.9), i.e., swapped returns. In this example, the violation of SYM is moderate, with an absolute allocation difference of 4 tokens. This results in preference reversals, where the majority of tokens are allocated to the flipped event.

As shown in Figures 3 and 4, subjects saw two of their choices side by side and had to undertake a decision. The exact instructions (translated from the original French version) were: "We would like to know whether you would like to modify your choices or not, regardless of the reason. Please note that if you decide to change your choice in one or more situations, it will replace your previous selection". They then saw the "I would like" default sentence, with the following alternatives: "not to modify my choice in this situation" or "modify my choice in this situation". This allows for all possibilities: changing only one choice, both, or none.

After subjects decided whether they wanted to modify their choices, and *before* they actually did so, we asked for their reasons, offering the following multiple-choice alternatives in addition to an open-text field (`others`) where they could freely express their reasoning if they wished (Table 4).

Table 4: Reasons

Variable	**Statement**
`difference`	These two situations are not the same for me
`indifference`	These two situations are equally good to me
`diversification`	I want to make diverse/different choices
`mistake`	I made a mistake in either or both of the choice situations
`mind-change`	I changed my mind
`undecided`	I do not know
`no-reason`	I do not have any particular reason
`others`	Other reason(s)

We borrowed this approach from Benjamin et al. (2020), which is, to our knowledge, the only recent study that explicitly include reasons behind subjects' desire to change or not their choices. For example, we assumed `difference` to possibly confirm rejection of SYM, STAT, and ANON for some subjects, since the two displayed choice situations are not identical in these cases. But we also assumed `indifference` to possibly confirm adherence to SYM, STAT, and ANON for some other subjects, as the two displayed choice situations in these cases can be viewed as similar. The variety of responses we offer, as well as the possibility for subjects to select as many reasons as they wished, accommodates any form of reasoning.

`mistake` is crucial to our investigation, as it represents the mainstream assumption in the related experimental literature and, more generally, the core of behavioral paternalism (Camerer et al., 2003; Thaler and Sunstein, 2003, 2009), as well as behavioral welfare economics (Bernheim and Rangel, 2007, 2009). It indicates that the subject acknowledges having made an error in her chosen allocations and would like to correct it. Theoretically, this suggests that the subject had an exogenous (or *ex-ante*) preference but, due to some noise or cognitive disturbance, failed to satisfy it.[11] `mind-change`

[11]One might argue that a psychological caveat is that "people are never wrong", suggesting that individuals are unlikely to admit having made a mistake, even if they believe they have. One possible counterargument is that they may view this question as an opportunity for self-assessment and improvement rather than as a judgment of their abilities. Specifically, we emphasize in bold and repeat twice throughout the entire set of instructions that "there are no right or wrong answers" and that "we are simply interested

differs from `mistake` in that the "Choice Reconsideration" phase may lead subjects to restructure or update their preferences (typically through learning), without assuming that these preferences preexisted their choices. This aligns with the body of theoretical models of endogenous preferences (Dietrich and List, 2013, 2016; Boissonnet et al., 2023) and relates to the constructed preference hypothesis (Slovic, 1995; Lichtenstein and Slovic, 2006).

Once subjects provided us with their reason(s) for modifying or not modifying their choices, those who chose to modify one or both were shown each interface again sequentially, with the possibility to modify their allocation as they wished. An advantage of this experimental setup is that subjects were never primed toward any specific form of behavior. If a subject decided to modify one or both choices in this phase, we intentionally provided no additional information beyond their original allocation. Indeed, adding information such as pedagogical instructions on the rationality principle being tested, like in MacCrimmon (1968), Moskowitz (1974) and Slovic and Tversky (1974), could risk biasing subjects' responses toward that principle. In the most extreme case, such wording could even be considered a *nudge* itself, defined as any alteration of the choice architecture designed to influence subjects' behavior toward a desired outcome (Thaler and Sunstein, 2009). The only instructions given at this stage were: "You can now modify your choice in the displayed choice situation".

3.3 Quality Reports

Afterward, we asked subjects to report some information about their choices in the previous phases. This third phase can be seen as a personal feedback from the subjects and helps us gain a deeper understanding of their choice processes, particularly their reasoning and the subjective consistency of their behavior. We first measured "quality decision-making", defined as features of the choice architecture that may have hindered or helped subjects in making their choices. In particular, we control for *choice overload*, i.e., if there is any cost associated to the number (24) of choices subjects have to make, the value of the *information* about their previous choices, and the value of the *opportunity* to modify their previous choices, which can be interpreted as subjects' perception of their opportunity cost of changing or not their choices. The description of each variable is listed in Table 5.

Table 5: Quality decision-making

Variable	Question
`choice-overload`	Did making all these choices (in the 24 situations) hinder or help you make your choices?
`information`	Did having some of your choices presented to you hinder or help you make better choices?
`opportunity`	Did giving you the opportunity to modify your choices hinder or help you make better choices?

Note. 1-7 Likert scale: Very hindered, Somewhat hindered, Slightly hindered, Neither hindered nor helped, Slightly helped, Somewhat helped, Very helped.

We then asked subjects to report their strategies—if they had any during the choice

in how people make choices". Additionally, their responses are fully anonymous, eliminating any peer effects.

tasks—and to describe them in an open field if they used strategies other than those we proposed. We proposed a list of six strategies in the risk domain (Table 6), seven in the time domain (Table 7), and ten in the social domain (Table 8).

Table 6: Risk | Strategies

Variable	Statement
variety	I aimed to make diverse or different distributions
consistency	I aimed to make similar or identical distributions in similar or identical situations
randomization	I aimed to distribute tokens in a random manner
maximization	I aimed to distribute tokens to achieve the highest expected gains
high-stakes	I aimed to bet a significant amount (including betting everything) on one of the two events when the potential gains were high
egalitarian	I aimed to make roughly equal distributions

Table 7: Time | Strategies

Variable	Statement
variety	I aimed to make diverse or different distributions
similarity	I aimed to make similar or identical distributions in similar or identical situations
randomization	I aimed to distribute tokens in a random manner
single-period	I aimed to allocate all tokens to one period only
two-period	I aimed to distribute tokens in both periods
high-stakes	I aimed to allocate a significant amount (including all tokens) to one of the two periods when the potential gains were high
egalitarian	I aimed to make roughly equal distributions

Table 8: Social | Strategies

Variable	Statement
variety	I aimed to make diverse or different distributions
similarity	I aimed to make similar or identical distributions in similar or identical situations
randomization	I aimed to distribute tokens in a random manner
selfish	I aimed to maximize my own gains without considering the other
altruist	I aimed for the other to achieve the highest possible gains without worrying about myself
egalitarian	I aimed to make roughly equal distributions between myself and the other
low-regarding	I aimed to ensure that I always left a small amount (even if very small) for the other
high-regarding	I aimed to ensure that I always left a small amount (even if very small) for myself
high-stakes	I focused on my own outcome when the potential gains were particularly high for me
compassion	I focused on the other when the potential gains were particularly low for him/her

Since the strategy reports were multiple-choice forms, subjects could select several strategies reflecting all kinds of approaches they may have applied across the 24 choice tasks—for example, choosing high stakes when returns were unequal, diversifying when returns were the same, and so on.

4 Results

We now present the results for each domain: risk (Section 4.1), time (Section 4.2), and social (Section 4.3). We analyze the proportion and direction of revision, that is, the degree of adherence to the rationality principles under study. We then investigate the reasons subjects provide conditional on their revisions.[12]

4.1 Risk Domain

We first examine consistency toward RETEST before and after revision. Table 9 reports consistency rates toward RETEST as proportions. Note that RETEST-Consistency requires that subjects reveal exactly the same allocation in the two situations forming the pair. RETEST-Consistency is shown conditional on whether subjects revised their choices.

Table 9: Risk | RETEST-Consistency conditional on revision

		RETEST-Consistency					
		Returns (1.2 – 1.2)			**Returns (2.4 – 0.7)**		
		no	yes	Total	no	yes	Total
Revision	no	15.6%	72.7%	88.3%	27.3%	47.7%	75.0%
	yes	10.2%	1.5%	11.7%	23.4%	1.6%	25.0%
Total		25.8%	74.2%	100.0%	50.8%	49.2%	100.0%

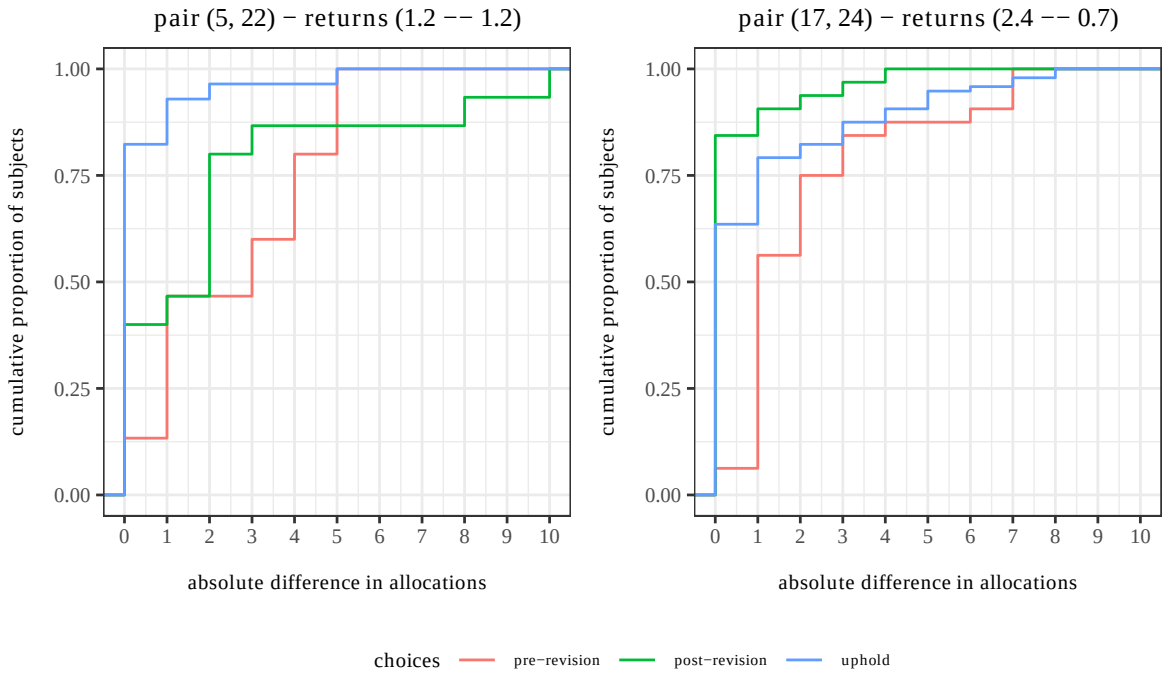
Note. Only an absolute difference of 0 token between the two choice situations forming a pair is considered consistent. Any deviation is classified as inconsistent. This applies to all contingency tables we show in the main text. Consistency tests which allow a difference of one token between the two allocation to be still considered as consistent toward the given principle (`maxDiff=1`) are provided in Appendix E. Here two-sided Fisher exact tests for both returns yield p-values of ≈ 0, strongly rejecting independence between revision and RETEST-Consistency. This confirms that revision is systematically associated with lower RETEST-Consistency, and is coherent with the interpretation that revision reflects awareness and updating of previous inconsistencies rather than random noise.

Unsurprisingly, almost no subject who was consistent with RETEST revised their choices: only around 1.5% did so, for both pairs. Among those who were inconsistent according to RETEST, fewer than half took the opportunity to modify their choices: 39% for returns (1.2 – 1.2) and 46% for returns (2.4 – 0.7).[13] Also, fewer subjects revised their choices when the two returns were perfectly equal (1.2 – 1.2), than when the returns were extremely unequal (2.4 – 0.7). Figure 5 depicts the cumulative distribution of subjects' absolute differences in token allocation between repeated trials.

[12]The aggregated results regarding quality decision-making are provided in Appendix D.

[13]This includes all subjects who indicated that they wanted to revise their choice(s), even if they ultimately did not change it/them. However, the number of subjects revising their choices is too small to draw firm conclusions about those who upheld their choice, despite initially indicating a desire to revise.

Figure 5: Risk | Revision comparison for RETEST



"Uphold", depicted by the blue line, shows the cumulative proportion of subjects who chose not to modify their choices in that pair. For instance, over 80% of subjects made exactly equivalent allocations for ChoiceID n°5 and ChoiceID n°22, and 100% deviated by fewer than 5 tokens between the two choice situations.[14] Comparing the two panels, we see that the RETEST-Consistency rates of the "uphold" group are overall higher when the "Heads" and "Tails" returns were equal (1.2 – 1.2) than when they differed substantially (2.4 – 0.7).

For subjects who revised their choices by updating one or both allocations, we observe an overall improvement in terms of RETEST-Consistency. This is depicted by the moving from "pre-revision" in red to "post-revision" in green. There is also a clear improvement for returns (2.4 – 0.7), where the subjects revised their choices in a way that exceeds the level of RETEST-Consistency of those who upheld their choices. This is depicted by the green line, which always lies above the blue line. The pattern looks somewhat different when the two returns were equal (1.2 – 1.2). Here we find an overall improvement in terms of RETEST-Consistency at lower allocation differences (less or equal than 5 tokens), but the opposite at higher allocation differences (more than 5 tokens). Moreover, subjects who revised their choices do not exceed the consistency of those who upheld them. That is, the green line always lies below the blue line. We now move on to SYM with Table 10, which reports the consistency rates toward SYM conditional on whether subjects' revised their choices.

[14]Note that the allocation interval is bounded below at 0 and above at 10.

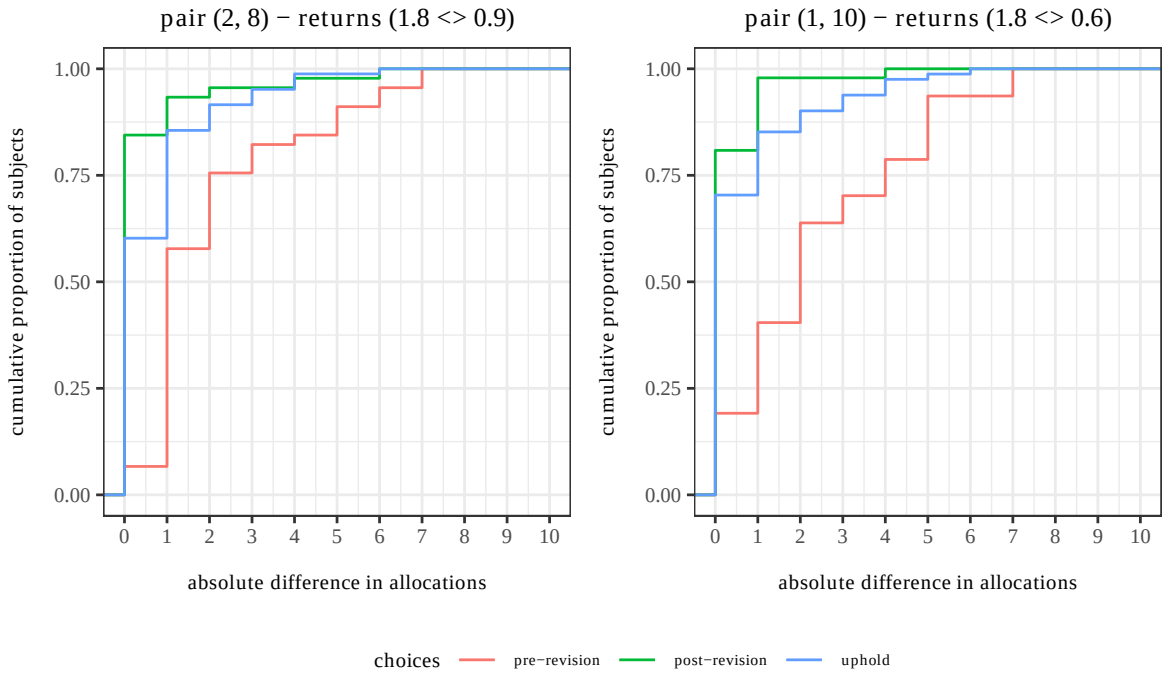
Table 10: Risk | SYM-Consistency conditional on revision

		SYM-Consistency					
		Returns (1.8 <> 0.9)			Returns (1.8 <> 0.6)		
		no	yes	Total	no	yes	Total
Revision	no	25.8%	39.0%	64.8%	18.8%	44.5%	63.3%
	yes	32.8%	2.3%	35.1%	29.7%	7.0%	36.7%
Total		58.6%	41.4%	100.0%	48.4%	51.6%	100.0%

Note. Two-sided Fisher exact tests for both returns yield p-values of ≈ 0.

Only about 41-52% of subjects made choices that are consistent with SYM across the two trials. This is a much lower rate of consistency than in RETEST (but also compared to SYM for `MaxDiff=1`, see Appendix E). Most consistent subjects did not revise (39% vs. 2.3% and 44.5% vs. 7%). Among inconsistent subjects toward SYM, a slight majority of 56% and 61% revised their choices in both pairs. By contrast, only a tiny proportion of the subjects who were consistent with SYM took the opportunity to revise their choice (0.05% and 0.1%). Among all the rationality principles examined in the present study, SYM shows the highest rate of choice revision (35–37%). Figure 6 depicts the cumulative distribution of subjects' absolute differences in token allocation between SYM trials.

Figure 6: Risk | Revision comparison for SYM



As the graphs show, there is a substantial improvement in adherence to SYM after choice revision for both pairs. Among the subjects who revised their choices, the proportion whose initial allocations satisfied SYM (around 10% and 24%, respectively) increased dramatically after revision, reaching approximately 80%. Subjects who revised

their choices even outperformed those who did not in terms of SYM-Consistency. This pattern is illustrated by the red line (pre-revision) lying below the blue line (uphold) and the green line (post-revision) lying above the blue line. Next, we investigate the reasons for which subjects revised or not their allocations conditional on revision. Recall that reasons are being asked to *all* subjects right after they stated their decision to revise none, one, or both choices, and right before they revised their choices if they actually decided to undertake the revision. Table 11 displays the proportions of reasons for RETEST and SYM conditional on revision.[15]

Table 11: Risk | Reasons conditional on revision

Reason	RETEST revision				SYM revision			
	Returns (1.2 – 1.2)		Returns (2.4 – 0.7)		Returns (1.8 <> 0.9)		Returns (1.8 <> 0.6)	
	no	yes	no	yes	no	yes	no	yes
`difference`	5%	13%	12%	12%	14%	13%	11%	23%
`indifference`	80%	7%	53%	0%	55%	7%	65%	2%
`diversification`	10%	47%	25%	28%	29%	27%	15%	17%
`mistake`	0%	7%	0%	16%	0%	20%	0%	26%
`mind-change`	2%	67%	0%	62%	0%	58%	0%	55%
`undecided`	1%	0%	2%	0%	1%	0%	1%	0%
`others`	7%	0%	4%	0%	7%	0%	3%	2%
`no-reason`	0%	0%	0%	0%	0%	0%	0%	0%

Among subjects who chose not to revise ("no"), `indifference` is the most frequently reported reason. This pattern is unsurprising, given that the RETEST condition shows two identical choice situations. A similar pattern emerges under SYM, where 55% and 65% of non-revising subjects report that the choice situations were equally good. By contrast, responses among subjects who revised their choices ("yes") display greater heterogeneity. The `mind-change` reason constitutes the modal and majoritarian response among subjects who revised their choices, for both RETEST and SYM. This may suggest that revision is associated with an internal updating process, potentially reflecting learning or preference construction occurring after subjects have completed the full set of 24 choice tasks and have been reminded of their earlier choices.

4.2 Time Domain

We now move on to the results in the time domain, where we examine consistency toward RETEST before and after revision. Table 12 reports consistency rates toward RETEST as proportions.

[15]The aggregated descriptive statistics of reasons in each domain (risk, time, social) are provided in Appendix F.

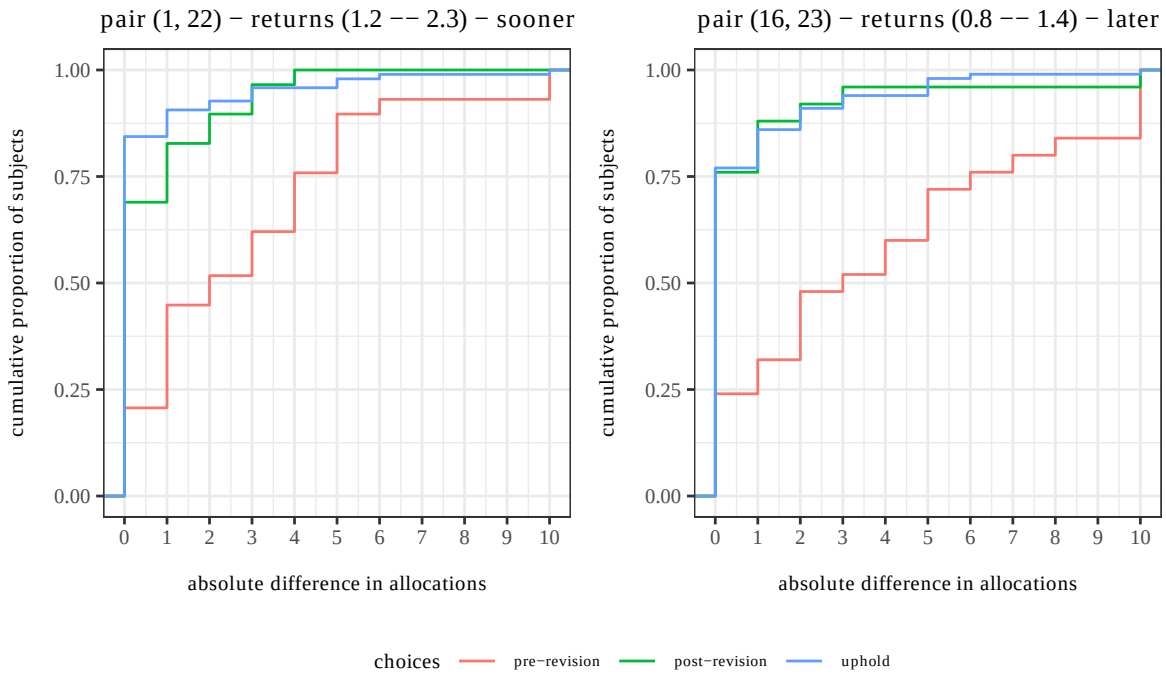
Table 12: Time | RETEST-Consistency conditional on revision

		RETEST-Consistency					
		Returns (1.2 – 2.3)			**Returns (0.8 – 1.4)**		
		no	yes	Total	no	yes	Total
Revision	no	12.0%	64.8%	76.8%	18.4%	61.6%	80.0%
	yes	18.4%	4.8%	23.2%	15.2%	4.8%	20.0%
Total		30.4%	69.6%	100.0%	33.6%	66.4%	100.0%

Note. Two-sided Fisher exact tests for both returns yield p-values of ≈ 0.

There is a highly similar structure than RETEST in risk across both intertemporal choice pairs. Once again, only a small minority of subjects revised their choices when they were already consistent, whereas a substantial fraction revised when inconsistent. In both the *sooner* ($0 \leftrightarrow 2$ weeks) and *later* ($2 \leftrightarrow 4$ weeks) tradeoffs, approximately 20–23% of participants revised their choices. Importantly, the "Choice Reconsideration" phase substantially improved RETEST-Consistency: the proportion of inconsistent choices dropped from a baseline level of roughly 75–80% before revision to about 20–25% afterward. This is depicted in Figure 7, where the shift from the red line (pre-revision) to the green line (post-revision) is substantial.

Figure 7: Time | Revision comparison for RETEST



Most of the subjects who revised their choices reached nearly the same level of RETEST-Consistency as those who upheld their initial choices. Overall, these findings suggest that time preferences are generally stable and well-calibrated when participants

are prompted to reflect on their earlier choices under identical parameters, as captured by the RETEST principle. Consider now adherence to STAT, with Table 13 reporting the consistency rates toward STAT conditional on whether subjects' revised their choices.

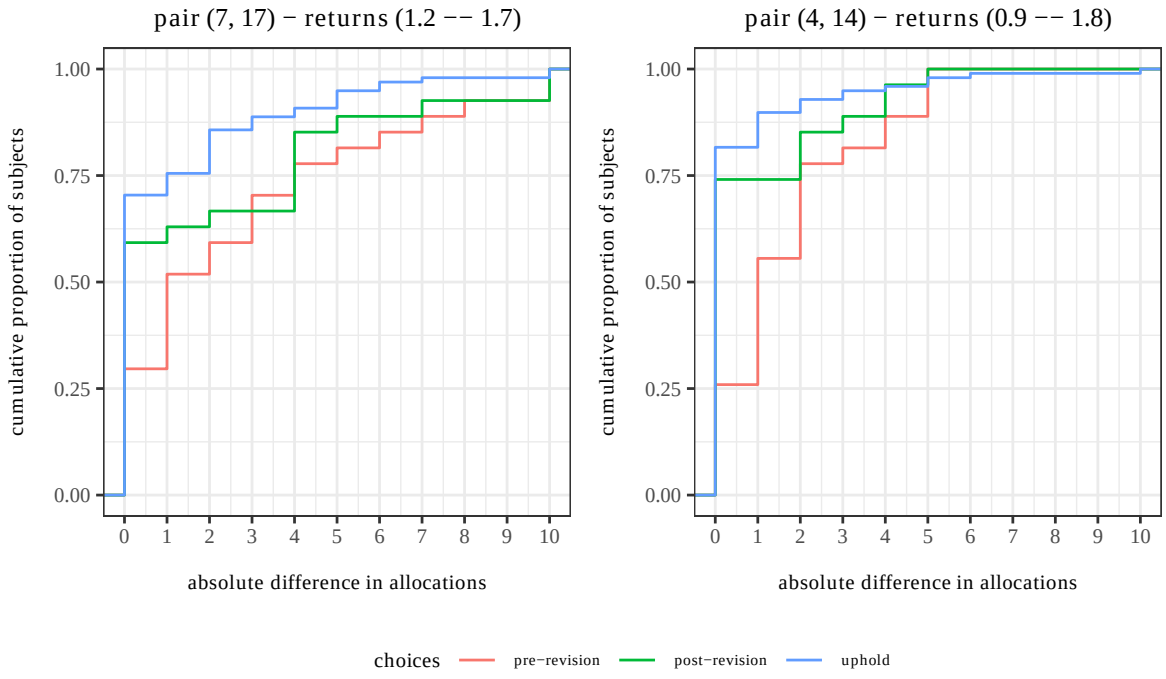
Table 13: Time | STAT-Consistency conditional on revision

		STAT-Consistency					
		Returns (1.2 – 1.7)			**Returns (0.9 – 1.8)**		
		no	yes	Total	no	yes	Total
Revision	no	23.2%	55.2%	78.4%	14.4%	64.0%	78.4%
	yes	15.2%	6.4%	21.6%	16.0%	5.6%	21.6%
Total		38.4%	61.6%	100.0%	30.4%	69.6%	100.0%

Note. Two-sided Fisher exact tests for both returns yield p-values of ≈ 0.

The proportion of revision is quite small in both pairs, where 21.6% of subjects revised their choice(s) vs. 78.4% who did not. Again, only an extreme minority of subjects chose to revise their choices when they were already consistent toward STAT (5-6%), while a small proportion (15-16%) revised their choices when inconsistent. Among the "uphold" group, approximately 70–80% of participants were consistent toward STAT, as indicated by the blue lines in Figure 8.

Figure 8: Time | Revision comparison for STAT



In general, larger differences between the returns (i.e., higher interest rates) were associated with fewer violations of STAT. Subjects who chose to revise their choices ini-

tially exhibited much lower consistency toward RETEST, but the "Choice Reconsideration" phase generally improved their adherence to the principle. Nevertheless, this group did not reach the same level of STAT-Consistency as the "uphold" group. As the two graphs show, the green lines (post-revision) remain below the blue lines (uphold). Table 14 reveals a sharp contrast in stated reasons between subjects who chose not to revise their initial allocation and those who did revise it, across both the RETEST and STAT principles.

Table 14: Time | Reasons conditional on revision

Reason	RETEST revision				STAT revision			
	Returns (1.2 – 2.3)		Returns (0.8 – 1.4)		Returns (1.2 – 1.7)		Returns (0.9 – 1.8)	
	no	yes	no	yes	no	yes	no	yes
`difference`	9%	10%	10%	4%	21%	18%	18%	18%
`indifference`	79%	3%	84%	0%	70%	4%	69%	0%
`diversification`	6%	31%	5%	16%	8%	22%	5%	30%
`mistake`	0%	24%	0%	24%	0%	18%	0%	33%
`mind-change`	1%	48%	0%	60%	1%	52%	0%	44%
`undecided`	0%	0%	0%	0%	0%	0%	2%	0%
`others`	9%	10%	8%	12%	7%	4%	9%	11%
`no-reason`	0%	0%	0%	0%	0%	0%	0%	0%

Conditional on no revision, the overwhelming majority of responses are classified as `indifference` for both RETEST (from 79% to 84%) and STAT (from 69% to 70%). By contrast, almost no subjects who revised their choices reported that the two choice situations were "equally good": the corresponding proportions range from 0% to 3% under RETEST, and from 0% to 4% under STAT. Instead, among subjects who revised their choices, the most frequent reason is again "I changed my mind" (`mind-change`). The `mistake` category is reported by revising subjects with a proportion of 24% under RETEST (for both returns), and from 18% to 33% under STAT. Although this provides evidence that the reconsideration stage allows some individuals to correct *ex-ante* errors, they remain minority responses (never exceeding 33%), corresponding to a very small absolute number of subjects. Finally, the absence of responses in the `undecided` and `no-reason` categories across all conditions indicates that subjects generally articulate a clear rationale for their behavior.

4.3 Social Domain

We repeat the same exercise for the social domain by first examining consistency toward RETEST before and after revision. Recall that the parameters in the social domain are identical with the risk domain (Table 2). The results are presented in Table 15.

Table 15: Social | RETEST-Consistency conditional on revision

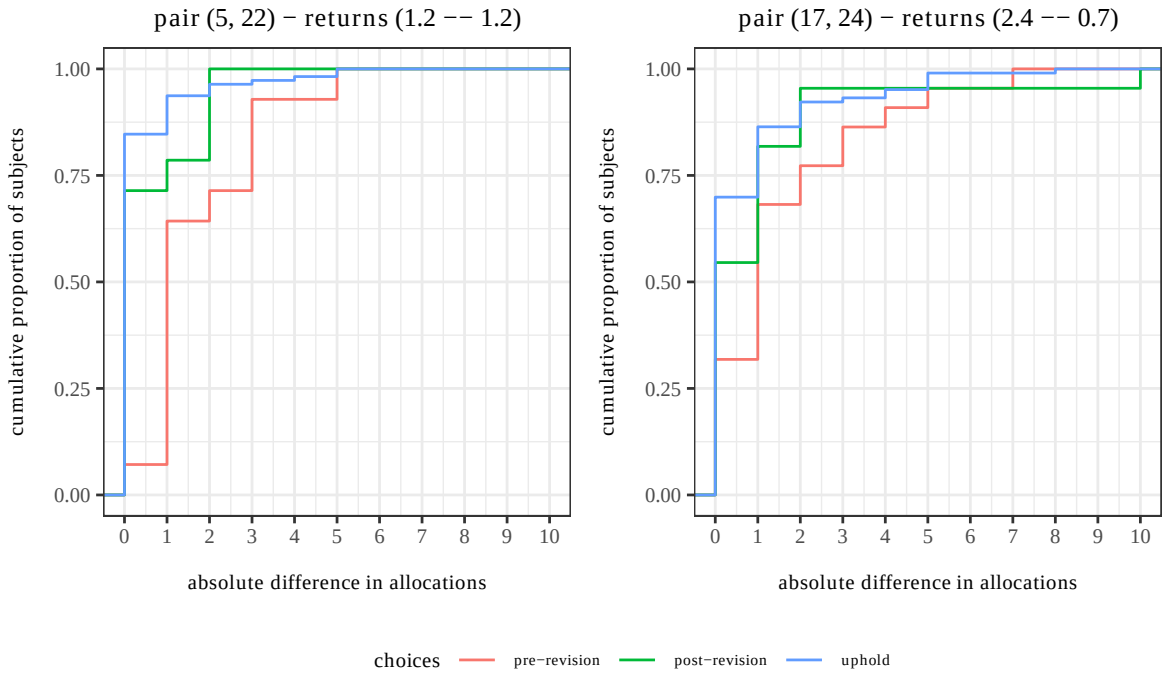
		RETEST-Consistency					
		Returns (1.2 – 1.2)			**Returns (2.4 – 0.7)**		
		no	yes	Total	no	yes	Total
Revision	no	13.6%	75.2%	88.8%	24.8%	57.6%	82.4%
	yes	10.4%	0.8%	11.2%	12.0%	5.6%	17.6%
Total		24.0%	76.0%	100.0%	36.8%	63.2%	100.0%

Note. Two-sided Fisher exact tests for both returns yield p-values of ≈ 0 and 0.001, respectively.

The social domain shows the lowest rates of revision compared to the risk and time domains. Here, only 11–18% of participants revised their choices across the two pairs, with a slightly higher rate of revision when differences in returns were larger. In contrast, 82–89% did not revise, representing a clear majority. This may suggest that people have a stronger sense of their preferences when it comes to social distribution, which, unlike individualistic risk and time decisions, involves ethical values that are generally stable and grounded in principles of justice. Again, only a tiny fraction of subjects chose to revise their choices when they were already consistent (0.8% and 5.6%), whereas revisions were only slightly more common among those who initially violated consistency toward RETEST (10% and 12%).

As the blue line (uphold) in Figure 9 shows, consistency toward RETEST was generally high among those who upheld their initial choices (around 75–80%), while subjects who revised their choices (pre-revision) displayed slightly lower initial consistency toward the principle.

Figure 9: Social | Revision comparison for RETEST



Subjects who revised their choices also slightly improved their consistency toward RETEST, approaching the level of the "uphold" group. This is depicted by the shift from the red line (pre-revision) to the green line (post-revision). Finally, we study adherence to ANON. Table 16 reports the consistency rates toward ANON conditional on whether subjects' revised their choices.

Table 16: Social | ANON-Consistency conditional on revision

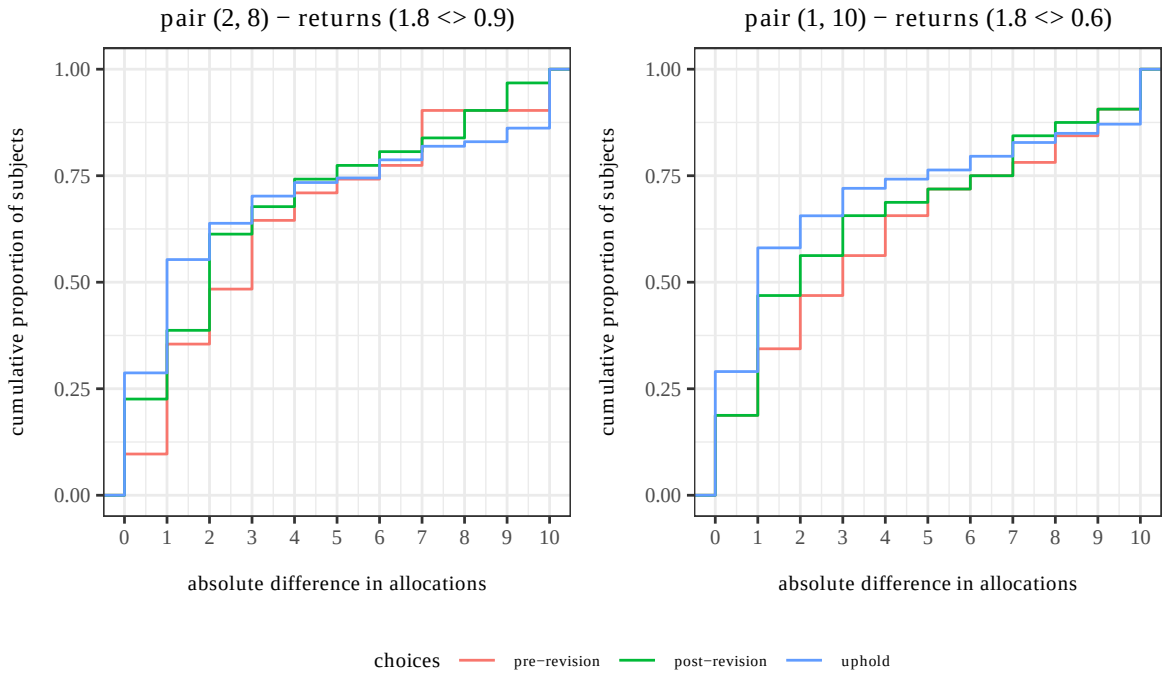
		ANON-Consistency					
		Returns (1.8 <> 0.9)			**Returns (1.8 <> 0.6)**		
		no	yes	Total	no	yes	Total
Revision	no	53.6%	21.6%	75.2%	52.8%	21.6%	74.4%
	yes	22.4%	2.4%	24.8%	20.8%	4.8%	25.6%
Total		76.0%	24.0%	100.0%	73.6%	26.4%	100.0%

Note. Two-sided Fisher exact tests for both returns yield p-values of ≈ 0.

Here again, the proportion of revision remains relatively low for both pairs (24–25%) compared to those who upheld their choices (74–75%), although it is notably higher than for RETEST. The contingency tables further show that choice revision is rare among consistent subjects toward ANON (2.4% and 4.8%), whereas a larger share of inconsistent subjects toward the principle revised their choices (22.4% and 20.8%). There is a very slight, but almost negligible, shift from inconsistency toward greater consistency according to ANON, as depicted by the shift between the red line (pre-revision) and the green

line (post-revision) in Figure 10.

Figure 10: Social | Revision comparison for ANON



Here subjects display the highest proportions of violations (76% and 73.6%) compared to all other principles across the risk, time and social domains. This suggests that ANON is less intuitive or less readily endorsed. Table 17 reports the reasons conditional on revision for the social domain.

Table 17: Social | Reasons conditional on revision

Reason	RETEST revision				ANON revision			
	Returns (1.2 – 1.2)		Returns (2.4 – 0.7)		Returns (1.8 <> 0.9)		Returns (1.8 <> 0.6)	
	no	yes	no	yes	no	yes	no	yes
`difference`	4%	0%	5%	14%	26%	19%	27%	12%
`indifference`	82%	14%	71%	4%	49%	0%	51%	0%
`diversification`	5%	36%	9%	27%	13%	39%	11%	25%
`mistake`	0%	7%	0%	18%	0%	13%	0%	25%
`mind-change`	0%	57%	0%	50%	0%	65%	1%	50%
`undecided`	1%	0%	0%	0%	1%	0%	1%	3%
`others`	5%	21%	12%	0%	9%	7%	10%	6%
`no-reason`	0%	0%	0%	0%	0%	0%	0%	0%

Among subjects who chose not to revise, `indifference` is the most frequently reported reason, particularly in the RETEST condition, with 71-82% of subjects. Again, this pattern is unsurprising, given that the RETEST principle involves identical choice situations. A comparable, though less pronounced, pattern is observed in the ANON condition, where roughly half of the non-revising subjects (49% and 51%) report `indifference`.

Responses among subjects who revised their choices are more heterogeneous. In the ANON condition, the `mind-change` reason again clearly dominates, accounting for at least 50% of revised choices, which again may point to an internal updating process whereby subjects reassess their preferences after being reminded of their earlier choices. The `mistake` category, while never modal, is non-negligible among revising subjects, especially under asymmetric return structures, suggesting occasional recognition of prior errors. Reasons such as `undecided` and `no-reason` are again reported only marginally or not at all across conditions. This indicates yet again that most subjects articulate a clear rationale for their behavior.

4.4 Summary

Result 1. ***Overall, subjects are highly consistent with almost all rationality principles.*** This contrasts with most of the related literature (Section 5), reporting, in general, low consistency rates for the rationality principles being examined across their tasks (mainly, under risk). We suspect the experimental design and preference elicitation method play a crucial role here. In particular, we believe our approach avoids creating situations in which subjects are inadvertently "trapped" into inconsistent choices, as we purposefully aimed to make the choice itself simple. Recall that the only parameter that subjects are expected to pay close attention to is the value of the returns, which constantly varies from one situation to another (except for pairs designed to measure adherence to RETEST). This allows us to rule out, as far as possible, the hypothesis that task difficulty itself reduces consistency toward the controlled rationality principles.

Result 2. ***Subjects tend to maintain their choices.*** Across all domains, we observe a relatively low rate of revision, with SYM showing the highest rates (35–37%). Also, among the subjects who upheld their initial choices, a substantial proportion were consistent with the rationality principles under study. We interpret these results as being driven by our design, in which preferences are first elicited, and then subjects are shown their initial choices and asked whether they want to change their choice(s). We believe this approach to be minimally intrusive, as subjects are not required to commit to a decision rule, and therefore are less exposed to potential biases (such as experimenter demand effect).

Result 3. ***Inconsistent subjects are not particularly prone to revise their choices.*** Across all domains, inconsistent responses conditional on revised and persistent choices occur at roughly equal rates. This pattern appears to depend less on the rationality principles themselves and more on the relative rates of return between the pairs under study. Our findings show that when the rates are more unequal, inconsistent participants are more likely to revise than to uphold their choices, within the same rationality principle being controlled for. Still, compared to related findings (e.g., Nielsen and Rehbeck, 2022), subjects who were inconsistent toward the rationality principles displayed a relatively low rate of choice revision. At first glance, this may provide limited support for the value of offering participants the opportunity to reconsider their choices. However, the next result contrasts this intuition.

Result 4. ***Inconsistent subjects who revise their choices align more with the rationality principles.*** Among those who were inconsistent *and* revised their choices, we

observe greater adherence to nearly all rationality principles under study. Table 18 shows the aggregated improvement for each domain.

Table 18: Consistency improvement per domain

	Risk			
	RETEST-Consistency		**SYM-Consistency**	
Assessment	**(1.2 – 1.2)**	**(2.4 – 0.7)**	**(1.8 <> 0.9)**	**(1.8 <> 0.6)**
Before revision	74.2%	49.2%	41.4%	51.6%
After revision	77.3%	68.8%	68.8%	74.2%
Consistency improvement	+3.1 pp.	+19.5 pp.	+27.3 pp.	+22.7 pp.
	Time			
	RETEST-Consistency		**STAT-Consistency**	
Assessment	**(1.2 – 2.3)**	**(0.8 – 1.4)**	**(1.2 – 1.7)**	**(0.9 – 1.8)**
Before revision	69.6%	66.4%	61.6%	69.6%
After revision	80.8%	76.8%	68.0%	80.0%
Consistency improvement	+11.2 pp.	+10.4 pp.	+6.4 pp.	+10.4 pp.
	Social			
	RETEST-Consistency		**ANON-Consistency**	
Assessment	**(1.2 – 1.2)**	**(2.4 – 0.7)**	**(1.8 <> 0.9)**	**(1.8 <> 0.6)**
Before revision	76.0%	63.2%	24.0%	26.4%
After revision	83.2%	67.2%	27.2%	26.4%
Consistency improvement	+7.2 pp.	+4.0 pp.	+3.2 pp.	0.0 pp.

Note. Improvement tests which allow a difference of one token between the two allocation to be still considered as consistent toward the given principle (`maxDiff=1`) are provided in Appendix G.

These results may suggest that inconsistent subjects who revised their choices, although a small proportion, made mistakes by failing to behave according to the rationality principles. Yet again, the next main result provides empirical evidence against this common interpretation.

Result 5. ***Choice improvements toward the rationality principles are not driven by mistakes.*** The observed improvements across all rationality principles we control for—particularly RETEST in all domains, as well as SYM—are *not* due to explicitly reported mistakes. This suggests that subjects' revisions reflect factors other than the correction of mistakes, contrary to what is often assumed in both the theoretical (Camerer et al., 2003; Thaler and Sunstein, 2003, 2009; Bernheim and Rangel, 2007, 2009) and experimental literature (Benjamin et al., 2020; Nielsen and Rehbeck, 2022; Breig and Feldman, 2024; Herweg et al., 2024). By contrast, the modal response in nearly all of our consistency tests is the `mind-change` reason, formulated as "I changed my mind". This is fundamentally different from the act of making a mistake. By definition, a mistake implies a deviation from a pre-existing intended choice. However, a subject may not necessarily have a fixed intention, or she may simply change her preference during the course of the experiment, as the choice tasks unfold. In Appendix H, we provide supporting evidence that subjects

did have a clear rationale in mind and aimed to achieve a certain form of coherence according to their own personal rules.

5 Literature Review

We now contrast our approach with related literature. Several experiments have examined people's adherence to rationality principles, mostly (though not exclusively) in the risk domain. Each study introduces specific methods for assessing such adherence, each with its own advantages and drawbacks that we discuss here. To our knowledge, MacCrimmon (1968) was the first experimental study in this area of research. The subjects were business executives who were allowed to reconsider their responses after being exposed to arguments for and against some principles of rationality. Overall, MacCrimmon found that a majority of subjects wanted to conform to the rationality principles after receiving *ex-post* feedback from the experimentalist. The occurrence of subjects self-reporting that they had made a mistake was observed several times, although the exact number of occurrences is not made explicit.

Moskowitz (1974) examined adherence to the principle of Independence of Irrelevant Alternatives (henceforth IIA), which was, at the time, the most empirically contested axiom following Allais (1953).[16] His experiment involved presenting subjects with Allais' type problems, framing the tasks in three different formats: words, trees, and matrices. Participants initially made choices, then viewed other students' responses to assess logical coherence, and, in one treatment group, were allowed to discuss their choices. After discussions, participants were more likely to align with rational principles, whereas those without discussion opportunities showed little change in their responses. An important strength of this experiment lies in the acknowledgment that *framing*—that is, the presentation of choice problems in alternative formats such as verbal descriptions, decision trees, or matrices—can lead to different responses, and potentially even shape individuals' preferences (see in particular Tversky and Kahneman, 1981).

A limitation of this experimental setup, however, concerns the possibility of *social conformity bias* arising from deliberation. Group discussions may have exerted pressure on participants to modify their choices in accordance with the majority view rather than expressing their authentic preferences. This echoes the conformity experiments of Asch (1951), where individuals, in a group setting, often conformed to incorrect majority statements even when making objective judgments. The experiment of Slovic and Tversky (1974) about the adherence to IIA yielded contrasting results. In their study, students were given choices in Allais' and Ellsberg's problems without social or corrective feedback but with explanation of the principles from the experimentalists. This was in response to some concerns regarding the results of MacCrimmon (1968), where, in their view, subjects may have been affected by the experimenter demand effect (Zizzo, 2010) and the peculiar characteristics of the sample (namely, business executives). According to their results, most participants maintained their choices, even when they conflicted with the principle.

[16] See in particular Mongin (2019) for a historical analysis of the early experiments aimed at measuring the normative appeal of IIA.

In a yet another related experiment, MacCrimmon and Larsson (1979) aimed to study the empirical violations of expected utility theory in decision-making experiments, where they allowed for an acceptance test of some rationality principles under risk, including IIA. Subjects were asked to rate, on an 11-point scale ranging from “strongly disagree” (0) to “strongly agree” (10), the appeal of each principle as a decision norm. This approach serves as another way to measure people’s adherence of the principles and study the discrepancy between the adhered principles and actual behavior.

Until Nielsen and Rehbeck (2022), there was, however, no experimental study aimed at incentivizing people’s preferences over rationality principles. The originality of their approach lies in the elegant and simple design of presenting some rationality principles to subjects using differently colored balls. Subjects are then asked whether they would be willing to delegate their choice to the experimentalist if such an option had been available, in which case their decision in the given task would be selected for payment. One drawback of their design, however, is that the choice reconsideration phase is highlighted in red, which may create a psychological effect suggesting that something went wrong. This could influence subjects toward more consistent decisions. There is also the issue of wording: using phrases such as “your choices were inconsistent with this rule” instead of a more neutral formulation like “you picked rule x and you did y” emphasizes the challenge of crafting instructions that do not push subjects toward one behavior or another. This difficulty, noted by Slovic and Tversky (1974), is shared across all studies that involve subjects reconsidering their choices.

A crucial point, as previously noted, is that we do not make any assumptions about mistaken choices based on observed behavior. The particular focus of Nielsen and Rehbeck (2022) is on *mistakes*—as made explicit by the title of their paper, "When Choices Are Mistakes". Specifically, the authors assume that when participants adjust their axiom and lottery selections to achieve consistency, this adjustment signals that their earlier choices were mistaken. They argue that this assumption is reasonable because the revision process gives participants additional information about how the axiom relates to their prior decisions. Accordingly, they treat the choices made in the "Choice Reconsideration" phase as a more accurate reflection of individuals’ authentic or “true” preferences.

Two underlying assumptions support this approach: (i) that people have pre-defined and fixed preferences prior to making a choice, and (ii) that any new information clarifies the situation and may either help subjects align with these preferences (if they exist) or form them (if they were not predefined). As previously mentioned, we only adopt assumption (ii). This is because, in our view, the reasons and the subjective definition of what it means to be consistent are internal to the subjects. An even more radical assumption is made by Breig and Feldman (2024) in a related experiment, who, in their words, “argue that a choice that is modified, absent any informational change, is revealed to have been a mistake”. In their experimental setup, subjects are given 25 relatively complex choice tasks, each involving a trade-off between an outcome and its probability, where each task is randomly repeated once throughout the study to test for consistency. A limitation of this design is that the measurement of inconsistency relies on the assumptions of the experimentalist, without any explicit "Choice Reconsideration" phase such as in Nielsen and Rehbeck (2022) and related experiments.

Using a design inspired by Nielsen and Rehbeck (2022), the experiment of Herweg et al. (2024) required respondents to revise all their decisions (i.e. not only the consistent ones), with all lottery choices related to the same axiom being revised simultaneously. Doing so, they aim to distinguish violations of canonical choice axioms under risk as either mistakes or genuine expressions of preferences. Although appealing, we believe this binary interpretation—categorizing behavior as either a mistake or a manifestation of true preferences—rules out several possibilities. For instance, it does not account for simple changes of mind—as supported by our results—meta-preferences (e.g., diversification in risk, where all the risk tasks can be treated as a risk portfolio), or other factors that cannot be anticipated *ex-ante* by the experimentalist.

Moreover, just like in Nielsen and Rehbeck (2022) and Breig and Feldman (2024), the experimental design of Herweg et al. (2024) implicitly assumes *ex-ante* preferences, treating them as fixed and exogenous—that is, independent of the choice task and process. We argue this view to be problematic for at least two reasons. First, it is unrealistic in the artificial context of a lab experiment, where the choice tasks are often detached from real-world alternatives about which subjects may feel confident enough to state genuine preferences. Second, it is empirically unfounded, since testing such an assumption would require at least that (i) a subject fail to adhere to a rationality principle, (ii) revise her choice, (iii) do so in the direction of greater consistency toward the given rationality principle, *while also* (iv) acknowledging that the original choice was a mistake. Our results provide some empirical evidence for (i), (ii), and (iii), but not (iv).[17]

Rather than "choice revision", we use the wording "choice reconsideration", which captures the possibility of reconsideration without implying pre-commitment to correcting errors or revealing "true" preferences. Practically, we extend the binary interpretation of behavior in Herweg et al. (2024) by allowing subjects to construct multiple narratives of their decision-making. These narratives are collected through our "reasons" and "strategies" questions, as well as through an open-text field where subjects can describe a "coherent story" of their own if they have one (see Appendix H). We inspired this experimental setup from Benjamin et al. (2020), where a distinctive feature of their study is the use of a simple algorithmic procedure to collect participants' feedback, allowing respondents to describe the reasoning behind their choices.

Specifically, participants in the experiment of Benjamin et al. (2020) were first asked: "Do you think the two situations are different enough that it makes sense to make different choices, or should they be the same?" Depending on their response, they were then presented with multiple options to explain their reasoning. These options included "I made a mistake", but also allowed for more nuanced responses such as "answering all of these questions made me change what I want", "I don't know which options I prefer",

[17]We speculate that the commitment to this "mistake" assumption is in fact mainly due to the standard view in microeconomic theory, which states that individual agents hold fixed and stable preferences over all states of the world (Mas-Colell et al., 1995). While theoretically appealing, experimental economics—drawing on historical studies that challenge the normative appeal of some rational choice principles (Allais, 1953) and demonstrate that people are not expected-utility maximizers (Kahneman and Tversky, 1979)—is ready to move beyond this "fixed preference" hypothesis. See in particular Cubitt et al. (2001) and Infante et al. (2016) for a critical and methodological assessment of what they call the "discovered" and "true" preference hypotheses, respectively. For a brief history of how "normativity" evolved in behavioral economics, see Mitrouchev and Dold (2025).

and even “I chose how I thought the experimenters wanted me to choose”. Related to the latter point, we aimed to avoid framing the opportunity to change a choice as an invitation to “make right what was wrong”, which could prime participants and introduce experimenter demand effects. To reduce such bias, the opportunity to reevaluate was presented to *all* participants, regardless of whether their initial choices were consistent with the rationality principles under study. Moreover, the first option offered to the subjects was always “I do not want to change my choice in this situation”, to avoid suggesting that revision is expected.[18]

All the studies cited above focus on choice under risk. To our knowledge, Andersson et al. (2023) is the only experiment on choice reconsideration in the social domain. The authors elicit preferences for redistribution principles and measure social distribution preferences in a general population sample. Half of their sample is randomly assigned to, in our words, a “Choice Reconsideration” phase in cases where participants’ preferred principles that conflict with their choices. This allows the authors to study how individuals resolve personal conflicts about social redistribution. Their findings indicate that the majority of participants abandon their initial adherence to the selected principle.

Crosetto and Gaudeul (2024) examine the decoy effect in choice under certainty. Their experimental design presents participants with three alternatives under a 20-second timer. The incentivized mechanism requires selecting an option quickly, after which participants can use the remaining time to change their choice if they see a preferable option. While this study is related to the literature on choice reconsideration because it is about letting subjects the possibility to change their choices, it is less directly aligned with it as its primary aim is not to measure adherence to a given rationality principle.

Beyond experimental studies, there is also theoretical and philosophical justification for using “confirmed choice” as a proxy for welfare (Ferreira, 2023). Distinguishing between context-independent choices and reason-based choices, Ferreira proposes an alternative criterion based on choices that individuals *confirm* after reflecting on their behavior. We consider this approach to be already implemented in practice through the various methods designed to measure adherence to rationality principles discussed in the present Section, but the question remains how this confirmation should be implemented: through communication with the experimenter (e.g., MacCrimmon, 1968; Slovic and Tversky, 1974), group deliberation (e.g., Moskowitz, 1974), or individual evaluation (e.g., Nielsen and Rehbeck, 2022). Table 19 summarizes the contributions of related experiments in comparison with our own.

[18] Still, as Thaler and Sunstein (2003) rightly emphasize in their seminal proposition of libertarian paternalism, *every* choice architecture entails framing, such that no choice situation can be, strictly speaking, "neutral". In this regard, we acknowledge that complete neutrality in a choice situation is unlikely to exist. However, this does not imply that neutrality is meaningless in *comparative* terms. That is, one choice situation may be more neutral than another. Framing also raises more fundamental concerns, such as the existence of free will (Gilboa, 2009)

Table 19: Summary of experiments on choice reconsideration

Experiment	Sample	Domain(s)	Choices	Preference elicitation	Multiple framing	Rationality adherence	Mistake assumption
Non-incentivized							
MacCrimmon (1968)	38 business executives	Risk	Hypothetical	Binary choice (Allais-type)	No	Observed behavior & interviews	No
Moskowitz (1974)	134 students	Risk	Hypothetical	Binary choice (Allais-type)	Yes	Observed behavior & group discussions	No
Slovic and Tversky (1974)	49 students	Risk	Hypothetical	Binary choice (Allais & Ellsberg-type)	No	Likert-scale approval	No
MacCrimmon and Larsson (1979)	19 students	Risk	Hypothetical	Binary choice (Allais-type)	No	Likert-scale approval	No
Incentivized							
Nielsen and Rehbeck (2022)	110 students	Risk	Real	Binary choice (lottery pairs)	No	Stated preference	Yes
Breig and Feldman (2024)	181 students	Risk	Real	Allocation (trade-off method)	Yes	Not measured	Yes
Herweg et al. (2024)	349 students	Risk	Real	Binary choice (lottery pairs)	Yes	Stated preference	Yes
Benjamin et al. (2020)	237 students	Risk	Hypothetical	Binary choice (lottery pairs)	Yes	Observed behavior & qualitative reports	Yes
Andersson et al. (2023)	2 295 British citizens	Social	Hypothetical	Binary choice (redistribution)	No	Stated preference	Yes
Crosetto and Gaudeul (2024)	111 mostly students	Certainty	Real	Ternary choice (deterministic)	No	Not measured	N/A
This experiment	378 mostly students	Risk & Time & Social	Real	Allocation (budget method)	No	Observed behavior & qualitative reports	No

6 Discussion

We have arrived at a classic question that has preoccupied a vast amount of discussions in decision theory: *what is rationality?* In standard models of decision-making, rationality is rooted in axiomatic foundations that reflect an individual's values regarding what one *should do*, making it appear inherently *normative*. Beyond the intuitive appeal of the axioms, that is most certainly grounded in the normative status of logic (Steinberger, 2022), there seems to be little, however, that makes any specific set of axioms an obvious or universal choice for defining rationality. This raises the question: how should the criteria for defining rationality be determined in the first place?[19] Given the plurality of values, attempting to characterize rationality from an *a priori* perspective—much like philosophers or theorists typically do—seems destined to be an endless endeavor. In addressing normative rationality empirically, our goal is not to supply a theoretical answer—one that might simply perpetuate this long-running debate.[20] Instead, building on the body of research in Section 5 that has made significant efforts to provide empirical insights and bring the concept of (normative) rationality closer to practical understanding, our approach takes a step further by seeking to identify some key principles of rational behavior, ultimately helping us to "uncover" rationality as a subjective concept.

Consider three major factors in cognitive processes: *information*, *computation* and *time*. In our lab setting, participants had never encountered the specific choice tasks before entering the experimental room, and had no prior training to prepare them. Although the task was designed to be extremely simple, it still required participants to exert minimal computational effort through the provided information—mainly the varying returns—in order to determine the allocation they preferred in each of the 24 choice situations. This information and computation, combined with the time subjects took to make their decisions (which was unconstrained by us but implicitly self-imposed by the duration of the sessions) led to some heuristics, such as the ones self-reported by some subjects in Appendix H. Our raw data (available upon request) shows that (i) some subjects developed several rules of their own to guide their decision-making and (ii) *throughout* the choice tasks. That is to say, the more subjects advanced in the choice tasks, the clearer representation they had about their allocation rules. This finding aligns closely with Simon's (1955) approach of rational choice, as well as the principles of the fast-and-frugal heuristics program, which emphasizes the adaptive use of simple decision-making strategies in a given choice context (Gigerenzer and Selten, 2001; Todd and Gigerenzer, 2012). In this approach, rationality is not about satisfying some principles/axioms of

[19] The "problem of the criterion" is ancient in epistemology. See Chisholm (1973) for an overview.

[20] This question posed, for example, a significant challenge for Savage in justifying why he believed he made a mistake when confronted with Allais' (1953) hypothetical experiment after he violated himself the IIA axiom. It was evident that Savage believed he made a mistake because he had faith in his own theory, which is expected utility theory. But *why*, specifically? Was it because the axioms of expected utility theory are logical, and, given that logic prescribes how we ought to think, we are compelled to accept the axioms? Alternatively, was it because Savage (1954) had invested considerable time and effort in developing expected utility, and thus felt compelled to uphold the theory as a normative rather than a descriptive one? Or was it perhaps a psychological reason—could it be that Savage, out of ego, could not admit that his own theory was flawed, and therefore felt the need to provide another interpretation for it (*ex-post rationalization* or *cognitive dissonance*)? For a historical analysis of Allais' paradox, which followed the experiments of MacCrimmon (1968), Moskowitz (1974), Slovic and Tversky (1974) and MacCrimmon and Larsson (1979), see Mongin (2019).

rational choice, nor about optimization, but more specifically, about optimization *given the choice context*.

In this matter, our findings support the following definition of rationality. *Rationality is about following a path of behavior that is aligned with personal motives given the available constraints in a specific context.* By "motives", we remain deliberately vague, as the factors that drive individual motivation are private, and therefore often inaccessible to the experimentalist (these can even be unconscious to the subjects themselves). By "constraints", we refer to factors—in particular *information treatment*, *computational ability* and *time management*—which are subject to individuals' personal capacities. In particular, subjects' scored differently in the Cognitive Reflection Test (CRT), and the time they took to complete the choice tasks varied considerably.[21] By "choice context" (often also refereed to as "choice architecture"), we mean *everything* related to the situation in which subjects have to make choices. The contextual factors can be external to the subjects, typically, the framing of the choice tasks, including the preference elicitation method that is being used. But they can also be internal, e.g., one's energy (fatigue) or mood of the day.[22]

Yet to meaningfully distinguish rational from irrational behavior, some fundamental criteria seem required. We propose two. First, as long as individuals can provide a narrative account of their choices, they may be considered rational. That is, they are able to organize their motives and constraints into a coherent story. A first potential criterion for rationality is therefore *narrativity*, defined as the meaning individuals attribute to the psychological relations among memories, desires, and preferences—see in particular Mitrouchev and Buonomo (2024). Second, another possible criterion is the notion of *regret* as an emotion or feeling, which can serve as a signal of what was truly desired.[23] In our experimental setting, a subject may leave the experimental room and subsequently reflect on a choice, for instance, adopting an "all-in" strategy for the return 2.4 in the return configuration (0.7 – 2.4), and later regret having taken such a risk after observing that the coin flip yielded the low return of 0.7. Within a given time window—that is necessarily chosen at the discretion of the decision-maker or the experimentalist—rationality can therefore be evaluated through (i) the possibility of choice revision, (ii) the opportunity to construct a coherent narrative based on that revision, and (iii) the *ex-post* affective response once uncertainty is resolved, regarding probabilities (in the risk domain) and outcomes (in all domains).

As previously emphasized, our results are closely aligned with the constructed preference hypothesis (Slovic, 1995; Lichtenstein and Slovic, 2006), according to which preferences are shaped and refined through the act of choice itself rather than existing *ex-ante*. This means that preferences highly depend on the choice context, where subjects may engage in *ex-post rationalization* following choice revision. They are likely to infer, or construct, a narrative about their preferences based on their previous actions. Based on

[21] We found, however, no robust association between CRT scores or time responses, on the one hand, and consistency with the rationality principles under study, on the other hand. Full results are available upon request.

[22] See Lecouteux and Mitrouchev (2024) for a broad characterization of context-dependent preferences.

[23] See in particular regret theory (Loomes and Sugden, 1982). Note that such a desire need not have pre-existed the choice, but may instead be formed specifically after the experience of regret. This accommodates the possibility that a genuine preference or desire can be produced *ex-post*.

our results, more realistic to us is the view that subjects *learn*, *discover*, and *adapt* their preferences in response to the task they face.[24]

Note that our study does not support the idea that individuals do not exhibit biases, since a bias is conceptually distinct from a mistake. A *cognitive bias* is a systematic pattern of deviation from a normative standard of judgment, potentially leading to perceptual distortion, inaccurate judgment, or illogical inference. Identifying a bias therefore requires an objective benchmark. A *mistake*, by contrast, is an act or judgment that is misguided or wrong *from the subject's own perspective*. Thus, while the benchmark for identifying a bias is *objective*, the benchmark for identifying a mistake is inherently *subjective*: it depends on whether the individual herself evaluates the deviation as erroneous. In this sense, just as Gigerenzer (2018) refer to the 'bias bias' as the tendency to systematically spot biases (when there might in fact be none), there may also be a bias in assuming that revising choices toward greater consistency with rationality principles exhibit original mistakes.

Based on existing models of decision-making, several sources of behavior can, more generally, explain how individuals' choices fail to conform to certain rationality principles. First, stochastic choice models that accommodate errors through a Fechnerian error term—such as random utility models or the Luce (1988) model—naturally align with the interpretation of mistakes. If an observed inconsistency is generated by a mistake, the individual should recognize it and be willing to revise her choice. Second, individuals may have imprecise or random preferences. A vast literature explores this source of apparent inconsistency, including work on random preferences (Gul and Pesendorfer, 2006; Cerreia-Vioglio et al., 2019), incomplete preferences (Hill, 2016), and empirical studies of preference imprecision (Butler and Loomes, 1988). In this case, individuals hold vague or fuzzy preferences over available options. The closer two alternatives are to indifference, the more likely it is that such imprecision arises. Under this explanation, subjects should not be willing to revise their choices. Instead, they should report indifference. Distinguishing this case from deliberate stochasticity is challenging. Under preference imprecision, subjects essentially do not know which option they prefer. An open question is how this maps into incentive responses, which may differ from those associated with indifference. Third, individuals may deliberately randomize, as in the literature on preferences for randomization (Agranov and Ortoleva, 2022). If deliberate randomization drives inconsistency, individuals may be willing to delegate their choice to a randomization device, potentially incurring a small cost to do so. Once the outcome is determined randomly, the subject should not wish to revert it, as this behavior would provide evidence of a preference for randomization—a form of diversification. If delegation is costless, the individual should be indifferent.

[24]From a methodological perspective, the main point of Slovic (1995) was in fact to underline a radical issue concerning the very nature of preferences, which is widely shared among preference-elicitation specialists: "If different elicitation procedures produce different orderings of options, how can preferences be defined and in what sense do they exist?" (p. 364).

We conclude by outlining some limitations and identifying directions for future research. One important limitation concerns potential biases inherent to laboratory experiments involving choice reconsideration. In particular, *ex-post rationalization* and *cognitive dissonance* may play a significant role, as subjects may provide a new justification for their choices after seeing the "Choice Reconsideration" phase, or adjust their responses to maintain internal consistency, especially when some of their initial motives conflict. More broadly, our experiment is subject to standard concerns related to experimenter demand effects, which are often regarded as a central challenge in this area of research. While our design seeks to mitigate such effects, they cannot be fully ruled out. Another important limitation is the scope of our analysis. We study rationality exclusively in the context of monetary decision-making. Accordingly, *our results should not be interpreted as general statements about rationality across all domains of life*, as is sometimes implicitly assumed in more abstract theoretical or philosophical accounts.

Several directions for future research naturally emerge from our study. First, our findings regarding the ANON principle differ markedly from those obtained for other rationality principles, suggesting that ANON may be less intuitive or less readily endorsed by subjects. Future studies could adapt our choice-allocation design to more closely examine adherence to the ANON principle in experimental settings where risk and social considerations are intertwined—see in particular Gaertner and Schokkaert (2012), Huang et al. (2019), and Wolf and Dron (2020). Second, further research could build on subjects' behavior (provided in Appendix I) by incorporating revealed-preference tests such as GARP (Generalized Axiom of Revealed Preference). This would allow us to assess consistency with other rationality principle—like Monotonicity and Transitivity—that were not explicitly addressed during our "Choice Reconsideration" phase. This approach has recently been adopted by Breig and Feldman (2024). Finally, our experimental design included additional observable measures, such as preference-survey modules, locus of control, and subjective well-being. Future work could investigate how these individual characteristics relate to conformity with and adherence to the rationality principles, following Choi et al. (2014).

References

Agranov, M. and P. Ortoleva (2022). Revealed preferences for randomization: an overview. *AEA Papers and Proceedings 112*, 426–30.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque : critique des postulats et axiomes de l'école américaine. *Econometrica 21*(4), 503–546.

Andersson, O., M. Lambrecht, and T. Miettinen (2023). Personal and societal conflict of distributive principles and preferences. Technical Report 14, Helsinki Graduate School of Economics.

Arneson, R. J. (1990). Liberalism, distributive subjectivism, and equal opportunity for welfare. *Philosophy and Public Affairs 19*(2), 158–194.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In *Groups, leadership, and men*, pp. 177–190. Carnegie Press.

Benjamin, D. J., M. A. Fontana, and M. S. Kimball (2020). Reconsidering risk aversion. Working Paper 28007, National Bureau of Economic Research.

Bernheim, B. D. (2016). The good, the bad, and the ugly: a unified approach to behavioral welfare economics. *Journal of Benefit-Cost Analysis 7*(1), 12–68.

Bernheim, B. D. and A. Rangel (2007). Toward choice-theoretic foundations for behavioral welfare economics. *American Economic Review 97*(2), 464–470.

Bernheim, B. D. and A. Rangel (2009). Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *The Quarterly Journal of Economics 124*(1), 51–104.

Boissonnet, N., A. Ghersengorin, and S. Gleyze (2023). Revealed deliberate preference change. *Games and Economic Behavior 142*, 357–367.

Breig, Z. and P. Feldman (2024). Revealing risky mistakes through revisions. *Journal of Risk and Uncertainty 68*, 227–254.

Butler, D. and G. Loomes (1988). Decision difficulty and imprecise preferences. *Acta Psychologica 68*(1), 183–196.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue, and M. Rabin (2003). Regulation for conservatives: behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review 151*(3), 1211–1254.

Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2019). Deliberately stochastic. *American Economic Review 109*(7), 2425–45.

Chisholm, R. M. (1973). *The Problem of the Criterion*. Milwaukee, WI: Marquette University Press.

Choi, S., S. Kariv, W. Müller, and D. Silverman (2014). Who is (more) rational? *American Economic Review 104*(6), 1518–50.

Crosetto, P. and A. Gaudeul (2024). Fast then slow: choice revisions drive a decline in the attraction effect. *Management Science 70*(6), 3711–3733.

Cubitt, R. P., C. Starmer, and R. Sugden (2001). Discovered preferences and the experimental evidence of violations of expected utility theory. *Journal of Economic Methodology 8*(3), 385–414.

Dalton, P. S. and S. Ghosal (2012). Decisions with endogenous frames. *Social Choice and Welfare 38*, 585–600.

Dietrich, F. and C. List (2013). Where do preferences come from? *International Journal of Game Theory 42*(3), 613–637.

Dietrich, F. and C. List (2016). Reason-based choice and context-dependence: An explanatory framework. *Economics and Philosophy 32*(2), 175–229.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Ferreira, J. V. (2023). Which choices merit deference? A comparison of three behavioural proxies of subjective welfare. *Economics and Philosophy 39*(1), 124–151.

Fine, B. (1995). On the relationship between true preference and actual choice. *Social Choice and Welfare 12*, 353–361.

Gaertner, W. and E. Schokkaert (2012). *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge: Cambridge University Press.

Gauthier, D. (1986). *Morals by Agreement*. Clarendon Press.

Gigerenzer, G. (2018). The bias bias in behavioral economics. *Review of Behavioral Economics 5*(3-4), 303–336.

Gigerenzer, G. and R. Selten (2001). *Bounded Rationality: The Adaptive Toolbox*. MIT Press.

Gilboa, I. (2009). *Theory of Decision under Uncertainty*. Econometric Society Monographs. Cambridge University Press.

Gilboa, I. (2010). *Rational Choice*. MIT Press.

Goodin, R. (1992). Laundering Preferences. In J. Elster and A. Hylland (Eds.), *Foundations of Social Choice Theory* (reprinted ed.)., pp. 75–101. Cambridge University Press.

Gul, F. and W. Pesendorfer (2006). Random expected utility. *Econometrica 74*(1), 121–146.

Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy 61*(5), 434–435.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy 63*(4), 309–321.

Harsanyi, J. C. (1977). Rule utilitarianism and decision theory. *Erkenntnis 11*(1), 25–53.

Herweg, F., S. Hippel, D. Müller, and F. Römeis (2024). Axiom preferences and choice mistakes under risk. Working paper, CESifo.

Hill, B. (2016). Incomplete preferences and confidence. *Journal of Mathematical Economics 65*, 83–103.

Huang, K., J. D. Greene, and M. Bazerman (2019). Veil-of-ignorance reasoning favors the greater good. *Proceedings of the National Academy of Sciences 116*(48), 23989–23995.

Infante, G., G. Lecouteux, and R. Sugden (2016). Preference purification and the inner rational agent: a critique of the conventional wisdom of behavioural welfare economics. *Journal of Economic Methodology 23*(1), 1–25.

Johnson, C., A. Baillon, H. Bleichrodt, Z. Li, D. van Dolder, and P. P. Wakker (2021). Prince: an improved method for measuring incentivized preferences. *Journal of Risk and Uncertainty 62*, 1–28.

Kahneman, D. and A. Tversky (1979). Prospect theory: an analysis of decision under risk. *Econometrica 47*(2), 263–291.

Karni, E. (1998). Impartiality: definition and representation. *Econometrica 66*(6), 1405–1415.

Koopmans, T. (1960). Stationary ordinal utility and impatience. *Econometrica 28*(2), 287–309.

Köszegi, B. and M. Rabin (2007). Mistakes in choice-based welfare analysis. *American Economic Review 97*(2), 477–481.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Lecouteux, G. and I. Mitrouchev (2024). The view from manywhere: normative economics with context-dependent preferences. *Economics and Philosophy 40*(2), 374–396.

Lichtenstein, S. and P. Slovic (2006). *The Construction of Preference*. Cambridge University Press.

Loomes, G. and R. Sugden (1982). Regret theory: an alternative theory of rational choice under uncertainty. *The Economic Journal 92*(368), 805–824.

Luce, R. D. (1988). Rank-dependent, subjective expected-utility representations. *Journal of Risk and Uncertainty 1*(3), 305–332.

MacCrimmon, K. (1968). Descriptive and normative implications of the decision-theory postulates. In K. Borch and J. Mossin (Eds.), *Risk and Uncertainty: Proceedings of a Conference held by the International Economic Association*, pp. 3–32. Palgrave Macmillan UK.

MacCrimmon, K. and S. Larsson (1979). Utility theory: Axioms versus 'paradoxes'. In M. Allais and O. Hagen (Eds.), *Expected Utility Hypotheses and the Allais Paradox: Contemporary Discussions of the Decisions under Uncertainty with Allais' Rejoinder*, pp. 333–409. Dordrecht: Springer Netherlands.

Manzini, P. and M. Mariotti (2014). Welfare economics and bounded rationality: the case for model-based approaches. *Journal of Economic Methodology 21*(4), 343–360.

Mas-Colell, A., M. Whinston, and J. Green (1995). *Microeconomic Theory*. Oxford University Press.

Mitrouchev, I. and V. Buonomo (2024). Identity, ethics and behavioural welfare economics. *Economics and Philosophy 40*(2), 310–336.

Mitrouchev, I. and M. Dold (2025). From heuristics and biases to agency. *Mind & Society 24*, 489–508.

Mongin, P. (2001). The impartial observer theorem of social ethics. *Economics and Philosophy 17*(2), 147–179.

Mongin, P. (2019). The Allais paradox: what it became, what it really was, what it now suggests to us. *Economics and Philosophy 35*(3), 423–459.

Moskowitz, H. (1974). Effects of problem representation and feedback on rational behavior in Allais and Morlat-type problems. *Decision Sciences 5*(2), 225–242.

Nielsen, K. and J. Rehbeck (2022). When choices are mistakes. *American Economic Review 112*(7), 2237–68.

Railton, P. (2003). *Facts, Values, and Norms: Essays toward a Morality of Consequence*. Cambridge University Press.

Rubinstein, A. and Y. Salant (2012). Eliciting welfare preferences from behavioural data sets. *The Review of Economic Studies 79*(1), 375–387.

Salant, Y. and A. Rubinstein (2008). (A, f): choice with frames. *The Review of Economic Studies 75*(4), 1287–1296.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley and Sons.

Simon, H. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics 69*(1), 99–118.

Slovic, P. (1995). The construction of preference. *American Psychologist 50*(5), 364–371.

Slovic, P. and A. Tversky (1974). Who accepts Savage's axiom? *Behavioral Science 19*(6), 368–373.

Steinberger, F. (2022). The Normative Status of Logic. In E. N. Zalta and U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022 ed.). Metaphysics Research Lab, Stanford University.

Sugden, R. (2022). Debiasing or regularisation? Two interpretations of the concept of "true preference" in behavioural economics. *Theory and Decision 92*(3-4), 765–784.

Thaler, R. H. and C. R. Sunstein (2003). Libertarian paternalism. *American Economic Review 93*(2), 175–179.

Thaler, R. H. and C. R. Sunstein (2009). *Nudge: Improving Decisions about Health, Wealth, and Happiness* (revised and expanded ed. ed.). Penguin Books.

Todd, P. M. and G. Gigerenzer (2012). *Ecological Rationality: Intelligence in the World*. Oxford University Press.

Tversky, A. and D. Kahneman (1981). The framing of decisions and the psychology of choice. *Science 211*(4481), 453–458.

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Wolf, S. and C. Dron (2020). The effect of an experimental veil of ignorance on intergenerational resource sharing: empirical evidence from a sequential multi-person dictator game. *Ecological Economics 175*, 106662.

Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics 13*(1), 75–98.

A Full Set of Instructions

Welcome to the Study

Before starting the study, please read the following instructions carefully.

Procedure

In this study, you will be asked to make choices regarding monetary distributions and answer some questions about your decisions. You will then complete a questionnaire on everyday life topics and take a brief cognitive test. The entire study lasts approximately one hour. It is structured as follows:

1. Tutorial and comprehension question
2. Your choices (24 situations)
3. Questions about your choices
4. Questionnaire on everyday life topics
5. Cognitive test
6. Feedback form
7. Your compensation

Compensation

Upon completing the study, you will receive a participation fee of €12. Additionally, you may earn up to €24 extra depending on the choices you make. In a few moments, we will randomly select a participant who will then draw a sealed envelope containing the number of a specific situation. For transparency, this envelope will be placed on the wall at the beginning of the study and will only be opened at the end so that everyone can see the number it contains. For example, if the envelope contains No. 17, your choice in situation No. 17 will be selected for your compensation. If the envelope contains No. 6, then your choice in situation No. 6 will be used for your compensation—and so on. Since we must wait for all participants to complete the study before opening the envelope, there is no advantage in rushing through the study. **Take your time**. If you finish before others, please wait patiently. Please note that all choices you make across the 24 situations are independent of each other. Since any number (from 1 to 24) could be selected, **it is in your best interest to carefully consider each decision.** Please note that **there are no right or wrong choices.** We are simply interested in how people make decisions. In this sense, each response is purely subjective—that is, unique to each individual. Therefore, you should make your choices based on what you believe is best for you.

Help

If you encounter any issues during the study or have any questions, we are here to assist you at any time. If this is the case, please raise your hand, and we will come to you.

Now, please wait a few moments

We will proceed with the random draw...
Enter here the activation key: [ ]

Instructions

In the following situations, you will make choices that are independent of one another. In each of the 24 situations, you will be asked to distribute 10 tokens between two different options. As a reminder, the envelope placed on the wall will be opened once all participants have completed the study. The number inside the envelope will determine your compensation. Since any number between 1 and 24 could be selected, it is in your best interest to carefully consider your choices in each situation. To familiarize yourself with the user interface, you will first go through a tutorial. Please click "Continue" to proceed.

How does it work?

In each of the 24 choice situations, you will start with 10 blue tokens, which are available in the left box. Your task is to distribute all 10 tokens between two possible events. These two events are represented by the boxes on the right. The situations you will encounter describe a coin flip. The coin can land on heads (represented by the "heads" side of the coin) or tails (represented by the "tails" side of the coin). You can think of these situations as a bet. For example, in a coin flip, if you prefer to bet 3 tokens on heads and 7 tokens on tails, place 3 tokens in the "heads" box and 7 tokens in the "tails" box. If you prefer to bet everything on heads and nothing on tails, place all the tokens in the "heads" box. Conversely, if you want to bet everything on tails and nothing on heads, place all the tokens in the "tails" box—and so on.

Your turn!

Now, familiarize yourself with the interface by moving the tokens. You can move them by dragging them from one box to another or by clicking the “+” and “−” buttons.

Tokens turn into money

As you have just seen, the tokens convert into a different monetary value for each event. The 24 choice situations you will encounter simply represent different ways the tokens are converted into money.

Feel free to experiment before confirming your choices

When you start a new choice situation, first explore the token values in both boxes by moving the tokens. At any time, you can see the total monetary value of the tokens at

the top of each box. This amount represents your potential compensation if this choice situation is selected by the envelope. We will then flip a coin, which will determine whether you receive the amount you allocated to heads or the amount you allocated to tails.

Remember, we are here to help

If anything is unclear, please raise your hand. On the next page, you will be asked a comprehension question to ensure that you have understood the instructions.

Comprehension question

Please answer the comprehension question below. Suppose that the choice situation below is determined by the envelope, and you have allocated 4 tokens to "heads" and 6 tokens to "tails." What would be your earnings for this part of the study if the coin lands on heads? (Please do not include the €12 participation fee in your answer).

Your response: €

Your response: € 4.80

Here we go!

You have correctly answered the comprehension question. You will now begin the study with your actual choices. Please click "Continue."

Here we show the 24 choice tasks sequentially on separated screens, as described in Table 1

Thank you for indicating your choices

In the next step, we will present some of the choices you just made and ask you questions about them. Once again, please note that **there are no right or wrong answers.** We are simply interested in how people make choices. In this sense, each answer is purely subjective—that is, specific to each individual. You should answer as you feel is best for you.

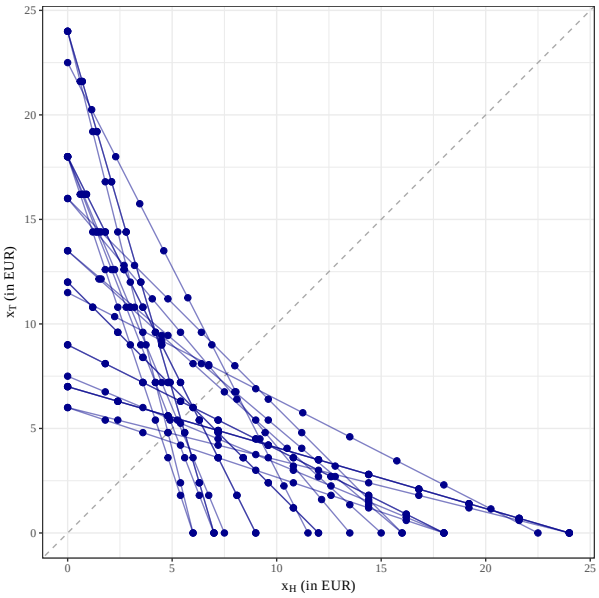
Questions about your choices

Here are the choices you made in two situations. We would like to know whether or not you would like to modify your choices, for any reason. Please note that if you decide to change your choice in one or more situations, it will replace your previous choice.

Here we show the "Choice Reconsideration" situations as described in Table 2 and Table 3

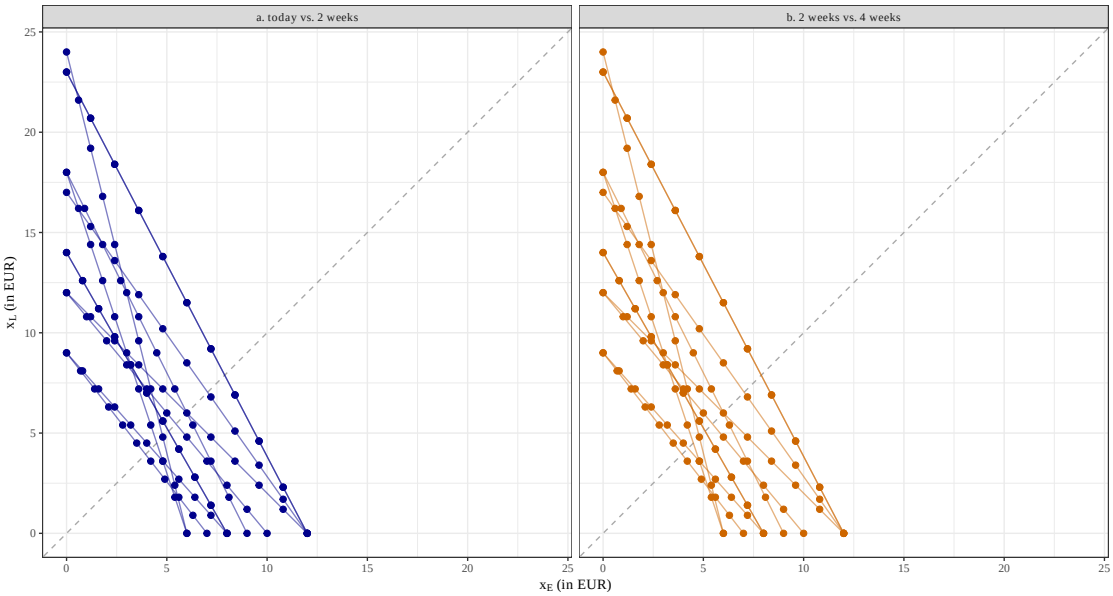
B Payoff Spaces

Figure 11: Risk | Budget lines



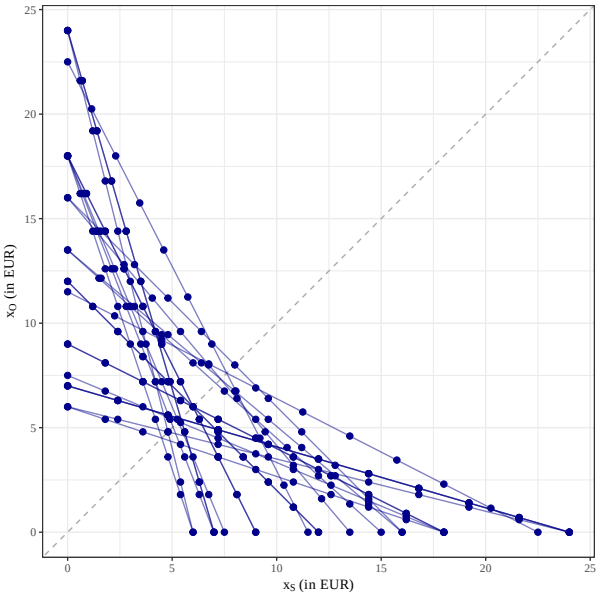
Note. Each budget line represents a choice situation in the risk domain. The dots indicate the feasible outcomes across the two states for given allocations. Let x_{H} denote the outcome if "Heads" is realized and x_{T} the outcome if "Tails" is realized. Outcomes in EUR (€) are computed as the number of tokens allocated to a state multiplied by that state's return—see Table 1 for the exact values. The dashed line corresponds to the certainty line, i.e., the same payoff regardless of which state obtains.

Figure 12: Time | Budget lines



Note. Each budget line represents a choice situation in the time domain. The dots indicate the feasible outcomes between a sooner and a later date for given allocations. Let x_{E} denote the outcome at the earlier point in time and x_{L} the outcome at the later point in time. Outcomes in EUR (€) are computed as the number of tokens allocated to a point in time multiplied by the actual return—see Table 1 for the exact values. The dashed line corresponds to the zero-discounting line (payoffs sooner and later have equal value). Panel a. shows allocations between today and two weeks from today, while panel b. shows allocations between two weeks and four weeks from today. Note that the budget lines have a slope steeper than -1, reflecting compensation for impatience: one EUR (€) received sooner is valued more than one EUR (€) received later.

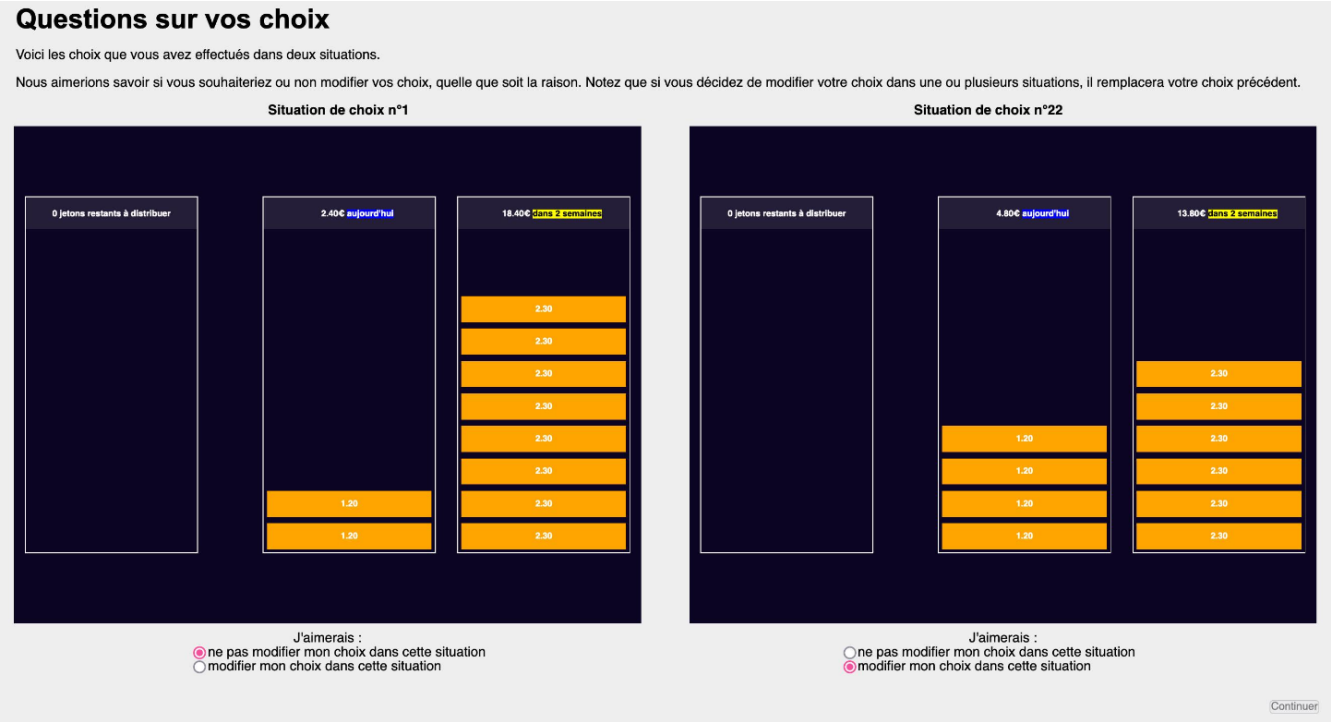
Figure 13: Social | Budget lines



Note. Each budget line represents a choice situation in the social domain. The dots indicate the feasible outcomes between the two persons for given allocations. Let x_{S} denote the outcome for the subject and x_{O} the outcome for the other person. Outcomes in EUR (€) are computed as the number of tokens allocated to a person multiplied by the actual return—see Table 1 for the exact values. The dashed line corresponds to the equality line, i.e., both persons receive the same payoff.

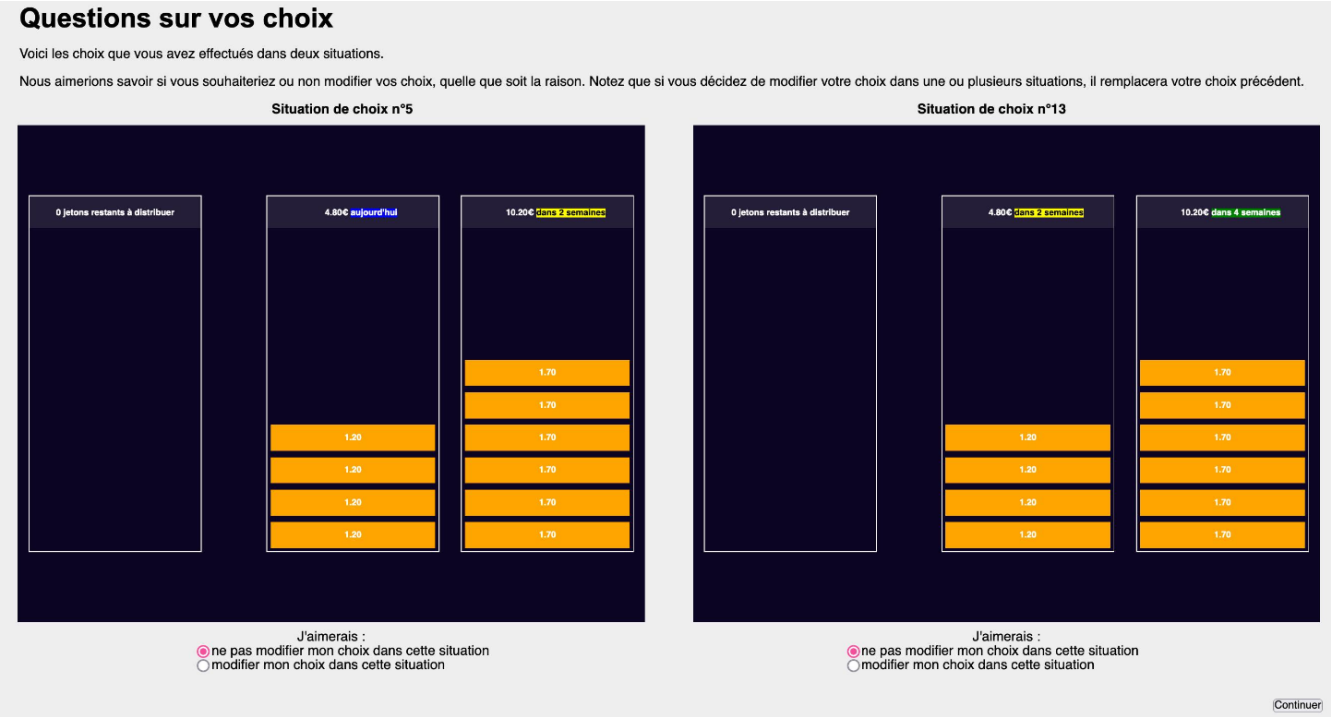
C Additional Screenshots

Figure 14: Time | Choice evaluation for RETEST



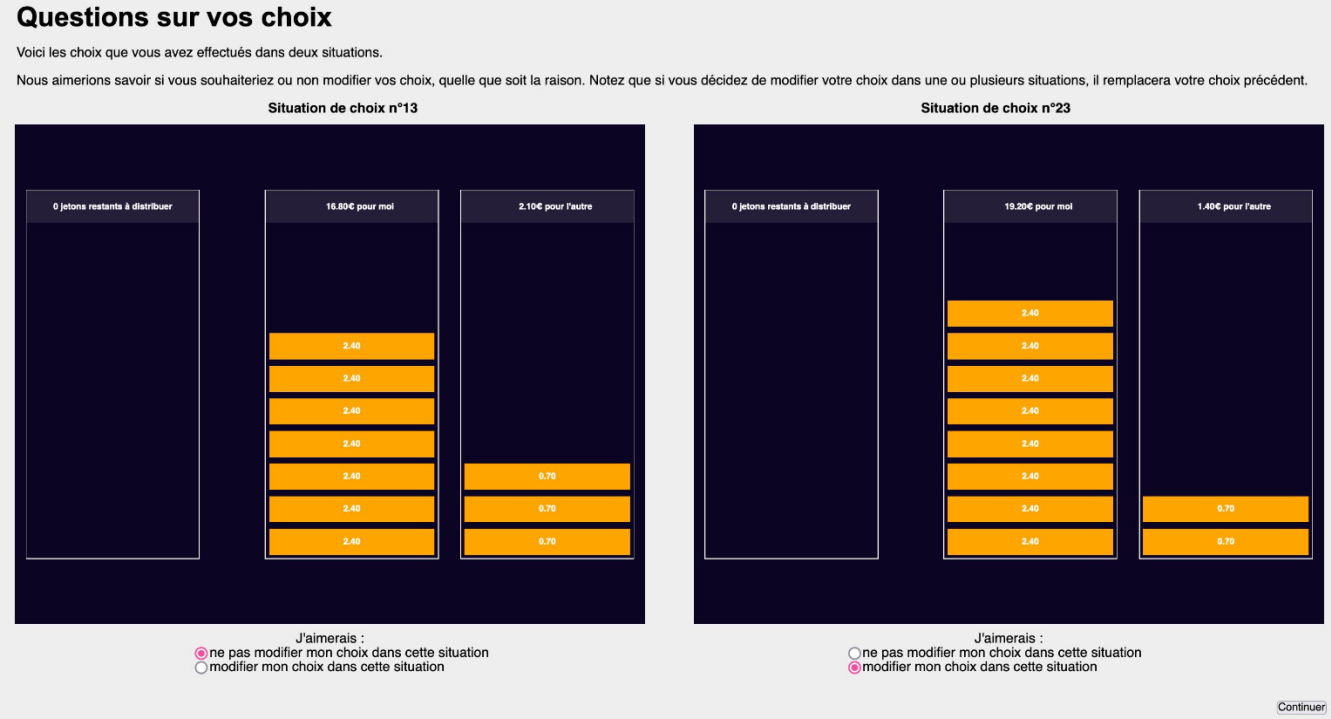
Note. These two choice situations, corresponding to the (1, 22) pair in Table 2, are identical. They both yield returns (1.2 – 2.3) for the same states of the world "Early $= 0$" and "Late $= 2$". In this example, the violation of RETEST is weak, with an absolute allocation difference of 2 tokens.

Figure 15: Time | Choice evaluation for STAT



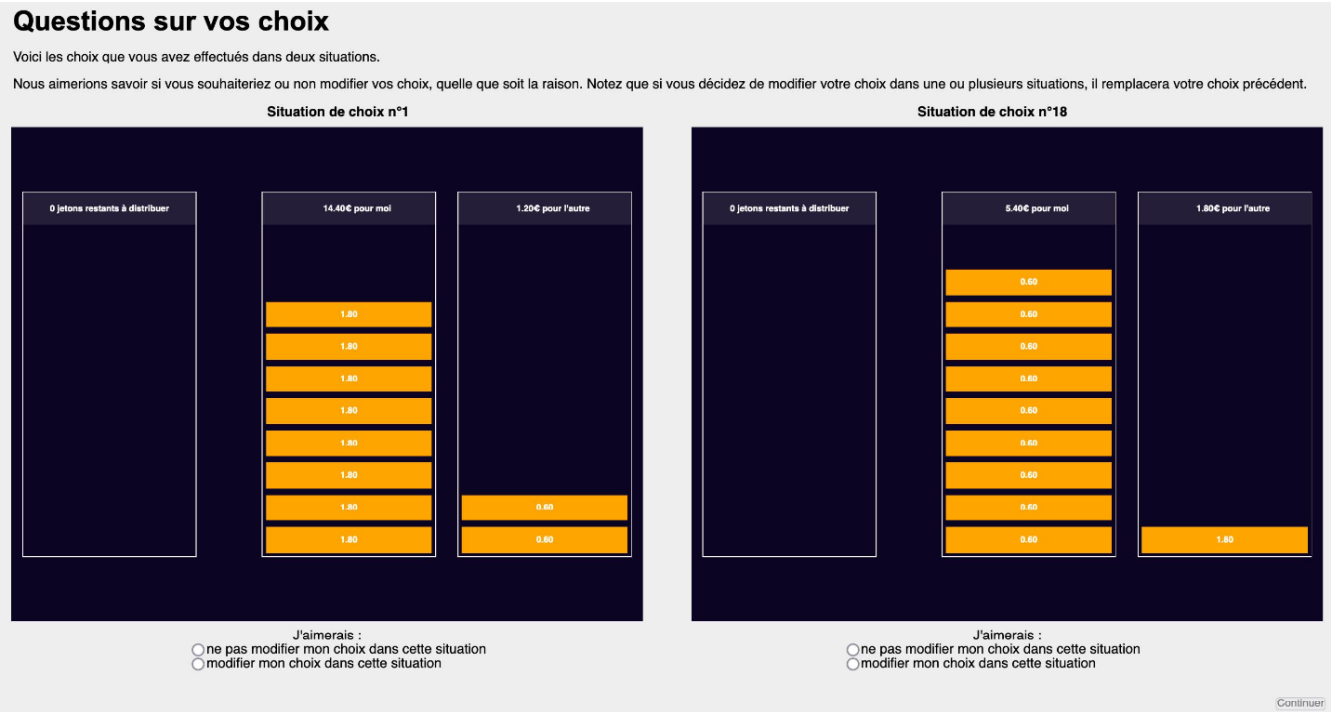
Note. These two choice situations, corresponding to the (5, 13) pair in Table 2, allow us to measure subjects' adherence to or rejection of STAT, as these situations have an identical return configuration that is shifted over time by two weeks (0 vs. 2 weeks and 2 vs. 4 weeks). We used different background colors in the labels to make the different points in time more salient: blue for "today" (0), yellow "in two weeks" (2) and green for "in four weeks (4). This was clearly mentioned in the tutorial instructions. In this example, the individual decision-maker is perfectly consistent with STAT and chooses not to modify her choices, revealing adherence to STAT.

Figure 16: Social | Choice evaluation for RETEST



Note. These two choice situations, corresponding to the (17, 24) pair in Table 2, are identical. They both yield returns (2.4 – 0.7). In this example, the violation of RETEST is weak, with an absolute allocation difference of only 1 token.

Figure 17: Social | Choice evaluation for ANON



Note. These two choice situations, corresponding to the (1, 18) pair in Table 2, allow us to measure subjects' adherence/rejection regarding ANON, as their configuration is (1.8 <> 0.6), i.e., swapped returns. In this example, the violation of ANON is very high, with an absolute allocation difference of 7 tokens.

D Quality Decision-Making: Aggregated Responses

We here provide aggregated descriptive statistics for what we refer to as the "quality decision-making" features. Quality decision-making entails aspects of the choice architecture that may have hindered or helped subjects in making their choices. For the `choice-overload` variable, this relates to whether the number of choice tasks in the experiment (24) had a negative, neutral, or positive impact on their choices. For the `information` and `opportunity` variables, this relates to how helpful the "Choice Reconsideration" phase was for subjects—namely, whether they appreciated being informed of their previous choices and whether they valued having the opportunity to change those choices if they wished. We report descriptive statistics separately for the risk (Table 20), time (Table 21), and social (Table 22) domains.

Table 20: Risk | Proportions of responses across quality decision-making

Response Category	`choice-overload`	`information`	`opportunity`	**Average**
Very helped	2.3%	3.9%	14.8%	7.0%
Rather helped	17.2%	21.1%	18.0%	18.8%
Slightly helped	15.6%	23.4%	17.2%	18.7%
Neither helped nor hindered	50.0%	42.2%	32.8%	41.7%
Slightly hindered	10.9%	6.2%	12.5%	9.9%
Rather hindered	3.9%	3.1%	3.9%	3.6%
Very hindered	0.0%	0.0%	0.8%	0.3%

Overall, most participants perceived the "Choice Reconsideration" phase as either neutral or somewhat helpful. For the `choice-overload` question, 50% of the participants reported feeling neither helped nor hindered, while 35.1% felt helped to varying degrees. Only a small minority reported being hindered (14.8%). Perceptions were slightly more positive for `information`, where almost half of participants (48.4%) indicated being helped, and only about 9% felt hindered. The most positive responses were observed for `opportunity`, where the opportunity to revise one's choices was seen as helpful by 50%, and only about 17% felt hindered. These results suggest that there was no choice overload effect across the 24 choice tasks, and that subjects valued the information provided by their previous choices (`information`) as well as the opportunity to modify them (`opportunity`).

Table 21: Time | Proportions of responses across quality decision-making

Response Category	`choice-overload`	`information`	`opportunity`	**Average**
Very helped	4.8%	4.8%	6.4%	5.3%
Rather helped	10.4%	7.2%	16.0%	11.2%
Slightly helped	9.6%	16.8%	23.2%	16.5%
Neither helped nor hindered	59.2%	52.0%	39.2%	50.1%
Slightly hindered	13.6%	12.8%	12.0%	12.8%
Rather hindered	2.4%	6.4%	3.2%	4.0%
Very hindered	0.0%	0.0%	0.0%	0.0%

For `choice-overload`, most subjects reported neutral or slightly positive experiences: 59.2% felt neither helped nor hindered by the amount (24) of choice tasks, while 24.8%

reported some degree of helpfulness, and only 16% felt hindered to any extent. Similar patterns emerge for presenting subjects `information` about their previous choices, where nearly half of participants (52.0%) felt neutral, 28.8% felt helped, and 19.2% felt hindered. The most positive responses appear again for subjects' `opportunity` to revise their choices, where 45.6% of participants reported some form of helpfulness, and only 15.2% felt hindered. Overall, these results suggest again that there was no effect of choice overload, and that participants generally appreciated the "Choice Reconsideration" phase, in particular through the opportunity to modify their choices.

Table 22: Social | Proportions of responses across quality decision-making

Response Category	`choice-overload`	`information`	`opportunity`	**Average**
Very helped	2.4%	3.2%	9.6%	5.1%
Rather helped	11.2%	12.0%	16.8%	13.3%
Slightly helped	11.2%	12.8%	17.6%	13.9%
Neither helped nor hindered	46.4%	47.2%	36.0%	43.2%
Slightly hindered	8.8%	7.2%	8.0%	8.0%
Rather hindered	16.8%	14.4%	8.8%	13.3%
Very hindered	3.2%	3.2%	3.2%	3.2%

Regarding `choice-overload`, participants tended to report neutral or mildly positive experiences: 46.4% indicated that the number of social allocation tasks (24) neither helped nor hindered them, 24.8% experienced some degree of support, while 28.8% felt somewhat hindered. For `information`, responses were comparable, with nearly half of the participants (47.2%) remaining neutral, 28.0% feeling helped, and 24.8% reporting that it made the choice process more difficult. The most favorable perceptions appear again under `opportunity`, where 44.0% of respondents felt that being able to revise their previous allocations was helpful, and only 19.2% experienced it as a hindrance. Taken together, these findings suggest that in the social domain as well, the "Choice Reconsideration" phase was generally well received. Participants did not experience choice overload, and they tended to value the opportunity to reflect on and, when desired, adjust their prior distributional decisions.

E Consistency Tests for maxDiff=1

Table 23: Risk | RETEST-Consistency conditional on revision (`maxDiff=1`)

		RETEST-Consistency					
		Returns (1.2 – 1.2)			Returns (2.4 – 0.7)		
		no	yes	Total	no	yes	Total
Revision	no	6.6%	82.0%	88.3%	15.6%	59.4%	75.0%
	yes	6.3%	5.5%	11.7%	10.9%	14.1%	25.0%
Total		12.5%	87.5%	100.0%	26.6%	73.4%	100.0%

Table 24: Risk | SYM-Consistency conditional on revision (`maxDiff=1`)

		SYM-Consistency					
		Returns (1.8 <> 0.9)			Returns (1.8 <> 0.6)		
		no	yes	Total	no	yes	Total
Revision	no	9.4%	55.5%	64.9%	9.4%	53.9%	63.3%
	yes	14.8%	20.3%	35.2%	21.9%	14.8%	36.7%
Total		24.2%	75.8%	100.0%	31.3%	68.8%	100.0%

Table 25: Time | RETEST-Consistency conditional on revision (`maxDiff=1`)

		RETEST-Consistency					
		Returns (1.2 – 2.3)			Returns (0.8 – 1.4)		
		no	yes	Total	no	yes	Total
Revision	no	7.2%	69.6%	76.8%	11.2%	68.8%	80.0%
	yes	12.8%	10.4%	23.2%	13.6%	6.4%	20.0%
Total		20.0%	80.0%	100.0%	24.8%	75.2%	100.0%

Table 26: Time | STAT-Consistency conditional on revision (`maxDiff=1`)

		STAT-Consistency					
		Returns (1.2 – 1.7)			Returns (0.9 – 1.8)		
		no	yes	Total	no	yes	Total
Revision	no	19.2%	59.2%	78.4%	8.0%	70.4%	78.4%
	yes	10.4%	11.2%	21.6%	9.6%	12.0%	21.6%
Total		29.6%	70.4%	100.0%	17.6%	82.4%	100.0%

Table 27: Social | RETEST-Consistency conditional on revision (`maxDiff=1`)

		RETEST-Consistency					
		Returns (1.2 – 1.2)			**Returns (2.4 – 0.7)**		
		no	yes	Total	no	yes	Total
Revision	no	5.6%	83.2%	88.8%	11.2%	71.2%	82.4%
	yes	4.0%	7.2%	11.2%	5.6%	12.0%	17.6%
Total		9.6%	90.4%	100.0%	16.8%	83.2%	100.0%

Table 28: Social | ANON-Consistency conditional on revision (`maxDiff=1`)

		ANON-Consistency					
		Returns (1.8 <> 0.9)			**Returns (1.8 <> 0.6)**		
		no	yes	Total	no	yes	Total
Revision	no	33.6%	41.6%	75.2%	31.2%	43.2%	74.4%
	yes	16.0%	8.8%	24.8%	16.8%	8.8%	25.6%
Total		49.6%	50.4%	100.0%	48.0%	52.0%	100.0%

F Reasons: Aggregated Responses

We here provide descriptive statistics for subjects' reasons for changing or not their choices, unconditional on their behavior. Note that the "reasons" questions were multiple-choice, allowing subjects to select more than one option. Consequently, the sum of reported proportions in the following tables may exceed 100%. We report descriptive statistics separately for the risk (Table 29 and 30), time (Table 31 and 32), and social (Table 33 and 34) domains.

Table 29: Risk | Proportions of reasons for RETEST

Reason	**Returns (1.2 – 1.2)**		**Returns (2.4 – 0.7)**	
	Frequency	**Proportion**	**Frequency**	**Proportion**
`indifference`	91	71.1%	51	39.8%
`diversification`	18	14.1%	33	25.8%
`mind-change`	12	9.4%	20	15.6%
`difference`	8	6.3%	15	11.7%
`no-reason`	6	4.7%	12	9.4%
`mistake`	1	0.8%	5	3.9%
`others`	8	6.3%	4	3.1%
`undecided`	1	0.8%	2	1.6%

Table 30: Risk | Proportions of reasons for SYM

Reason	**Returns (1.8 <> 0.9)**		**Returns (1.8 <> 0.6)**	
	Frequency	**Proportion**	**Frequency**	**Proportion**
`indifference`	49	38.28%	54	42.19%
`diversification`	36	28.12%	20	15.62%
`mind-change`	26	20.31%	26	20.31%
`difference`	18	14.06%	20	15.62%
`mistake`	9	7.03%	7	9.38%
`no-reason`	7	5.47%	9	10.16%
`others`	6	4.69%	6	2.34%
`undecided`	1	0.78%	1	0.78%

Subjects' most common reason in the risk domain (for both RETEST and SYM) is `indifference`. The drop in the proportion of respondents citing `indifference` from 71.1% to 39.8% may be explained by differences in the return structures (1.2 – 1.2 vs. 2.4 – 0.7), where subjects are more likely to diversify their choices. `diversification` is the second most common reason reported and can be interpreted as follows. Knowing the "1/24" uncertainty-resolution rule, some subjects may treat their decisions as part of a portfolio or be willing to gamble. They may insure certain amounts in choice pairs with similar returns and go "all-in" in choice pairs with very unequal returns, such as (2.4 – 0.7). The `mind-change` reason appears moderately frequent, mostly expressed by subjects who reconsidered their choices (see Table 11 in the main text). Very few

subjects reported having made a mistake in SYM (7.03% and 9.38%). Interestingly, a small but notable proportion of subjects (6.3% and 11.7%) considered the two choice situations in RETEST as different (provided by `difference`), even though, by definition, these situations are identical in all parameters.[25]

Table 31: Time | Proportions of reasons for RETEST

Reason	**Returns (1.2 – 2.3)**		**Returns (0.8 – 1.4)**	
	Frequency	**Proportion**	**Frequency**	**Proportion**
`indifference`	77	60.6%	84	66.1%
`mind-change`	15	11.8%	15	11.8%
`diversification`	15	11.8%	9	7.1%
`difference`	12	9.5%	11	8.7%
`others`	12	9.5%	11	8.7%
`mistake`	7	5.5%	6	4.7%
`no-reason`	6	4.7%	10	7.9%
`undecided`	0	0.0%	0	0.0%

Table 32: Time | Proportions of reasons for STAT

Reason	**Returns (1.2 – 1.7)**		**Returns (0.9 – 1.8)**	
	Frequency	**Proportion**	**Frequency**	**Proportion**
`indifference`	70	49.0%	68	46.9%
`difference`	26	18.2%	23	15.9%
`mind-change`	15	10.5%	12	8.3%
`diversification`	14	9.8%	13	9.0%
`others`	8	5.6%	12	8.3%
`no-reason`	5	3.5%	12	8.3%
`mistake`	5	3.5%	9	6.2%
`undecided`	0	0.0%	2	1.4%

`indifference` is again the most common response, which is not surprising given the high proportion of non-revision. For both RETEST and STAT, a notable proportion of subjects (around 10%) declared that they changed their mind. Very few subjects reported having made a mistake (less than 6% across all rationality principles). Again, a small but non-negligible proportion of subjects (9.45% and 8.66%) considered the two choice situations in the RETEST condition as different.

[25]This supports our previous claim that, from an experimental perspective, one cannot assume anything about subjects' motivations, as they may treat two objectively identical choice situations as different for purely subjective or personal reasons.

Table 33: Social | Proportions of reasons for RETEST

Reason	Returns (1.2 – 1.2)		Returns (2.4 – 0.7)	
	Frequency	Proportion	Frequency	Proportion
indifference	93	66.9%	74	58.7%
no-reason	12	8.6%	12	9.5%
diversification	11	7.9%	15	11.9%
others	9	6.5%	12	9.5%
mind-change	8	5.8%	11	8.7%
difference	4	2.9%	8	6.4%
undecided	1	0.7%	0	0.0%
mistake	1	0.7%	4	3.2%

Table 34: Social | Proportions of reasons for ANON

Reason	Returns (1.8 <> 0.9)		Returns (1.8 <> 0.6)	
	Frequency	Proportion	Frequency	Proportion
indifference	46	36.5%	47	37.3%
difference	30	23.8%	29	23.0%
diversification	24	19.1%	18	14.3%
mind-change	20	15.9%	17	13.5%
no-reason	16	12.7%	12	9.5%
others	10	7.9%	11	8.7%
mistake	4	3.2%	8	6.4%
undecided	1	0.8%	2	1.6%

The indifference reason is again the most prevalent, although it is much less frequent for ANON (36.51% and 37.3% of respondents) than for RETEST (66.91% and 58.73%). This contrast may indicate the low normative relevance of the ANON principle, coupled with the very low proportion of subjects (a bit more than 25%) who were consistent according to ANON and did not revise their choices. An additional indicator of the low normative relevance of ANON may be the relatively high proportions of subjects who considered the two choice situations as different, where the difference reason was selected by approximately 23% of respondents. The other less frequent reasons were diversification (19.05% and 14.29%), mind-change (15.87% and 13.49%), and no-reason (12.7% and 9.52%). Once again, the proportion of reported mistakes remains very low in the social domain (below 7%).

G Consistency Improvement for maxDiff=1

Table 35: Consistency improvement for maxDiff=1

	Risk			
	RETEST-Consistency		SYM-Consistency	
Assessment	**(1.2 – 1.2)**	**(2.4 – 0.7)**	**(1.8 <> 0.9)**	**(1.8 <> 0.6)**
Before revision	87.5%	73.4%	75.8%	68.8%
After revision	87.5%	82.0%	88.3%	89.8%
Consistency improvement	0.0 pp.	+8.6 pp.	+12.5 pp.	+21.1 pp.
	Time			
	RETEST-Consistency		STAT-Consistency	
Assessment	**(1.2 – 2.3)**	**(0.8 – 1.4)**	**(1.2 – 1.7)**	**(0.9 – 1.8)**
Before revision	80.0%	75.2%	70.4%	82.4%
After revision	88.8%	86.4%	72.8%	86.4%
Consistency improvement	+8.8 pp.	+11.2 pp.	+2.4 pp.	+4.0 pp.
	Social			
	RETEST-Consistency		ANON-Consistency	
Assessment	**(1.2 – 1.2)**	**(2.4 – 0.7)**	**(1.8 <> 0.9)**	**(1.8 <> 0.6)**
Before revision	90.4%	83.2%	50.4%	52.0%
After revision	92.0%	85.6%	51.2%	55.2%
Consistency improvement	+1.6 pp.	+2.4 pp.	+0.8 pp.	+3.2 pp.

H Qualitative Self-Reports

We selected a set of representative self-reports for each domain: risk (Table 36), time (Table 37) and social (Table 38). These are based on: (i) clarity and explicitness of the stated behavioral rule, (ii) internal consistency of the explanation, and (iii) frequency of similar responses in the dataset. Note that we are not claiming exhaustiveness. We are presenting illustrative prototypes. In total, 25 out of 128 subjects ($\approx$ 20%) expressed themselves about their heuristics in the risk domain, 33 out of 125 ($\approx$ 26%) in the time domain, and 49 out of 125 ($\approx$ 39%) in the social domain. For the simplified self-reports, we translated the original verbal statements from French into English, condensed them to 1–2 sentences, removed examples and digressions, and preserved only the core behavioral rule. Counts (n) indicate the number of participants whose self-reports were categorized under a behavioral rule. Therefore, the sum may exceed the total number of observations.

Table 36: Risk | Heuristics of representative self-reports

Behavioral rule	**Simplified self-report**	**n**
Insurance strategy	*"I aimed to guarantee a minimum gain (around €3–€4) in all possible outcomes, even if this reduced potential higher gains."*	6
Adaptive strategy	*"I adapted my strategy depending on whether outcomes were similar or very different, balancing safety and risk accordingly."*	6
Maximize average payoff	*"I distributed tokens to maximize expected earnings, allocating more tokens where their monetary value was higher."*	5
Accept variance for high upside	*"When the difference in potential gains between options was large, I accepted more risk in order to pursue higher rewards."*	5
Prioritize high-value outcomes	*"I concentrated my bets on outcomes where each token was worth more, even if this increased the risk of earning less."*	4
Balance outcomes	*"When possible, I distributed tokens evenly to obtain similar payoffs regardless of whether the outcome was heads or tails."*	4
Ignore low-value outcomes	*"When the value per token was too low, I ignored that option and focused entirely on higher-value outcomes."*	3

Table 37: Time | Heuristics of representative self-reports

Behavioral rule	**Simplified self-report**	**n**
Maximize total payoff	*"I focused on maximizing the total gain, even if it required waiting 2–4 weeks."*	14
Exclusive and high-stakes focus	*"I concentrated all tokens on the option with the highest payout."*	9
Immediate reward	*"I prioritized collecting money immediately rather than waiting for larger future gains."*	6
Adaptive strategy	*"I adapted my choices depending on payoff differences or waiting time."*	5
Balance outcomes	*"I distributed tokens evenly between periods when payoffs were similar to smooth returns."*	5
Threshold strategy	*"I chose the period with the higher payoff if the difference exceeded a certain threshold. Otherwise, I chose sooner."*	4

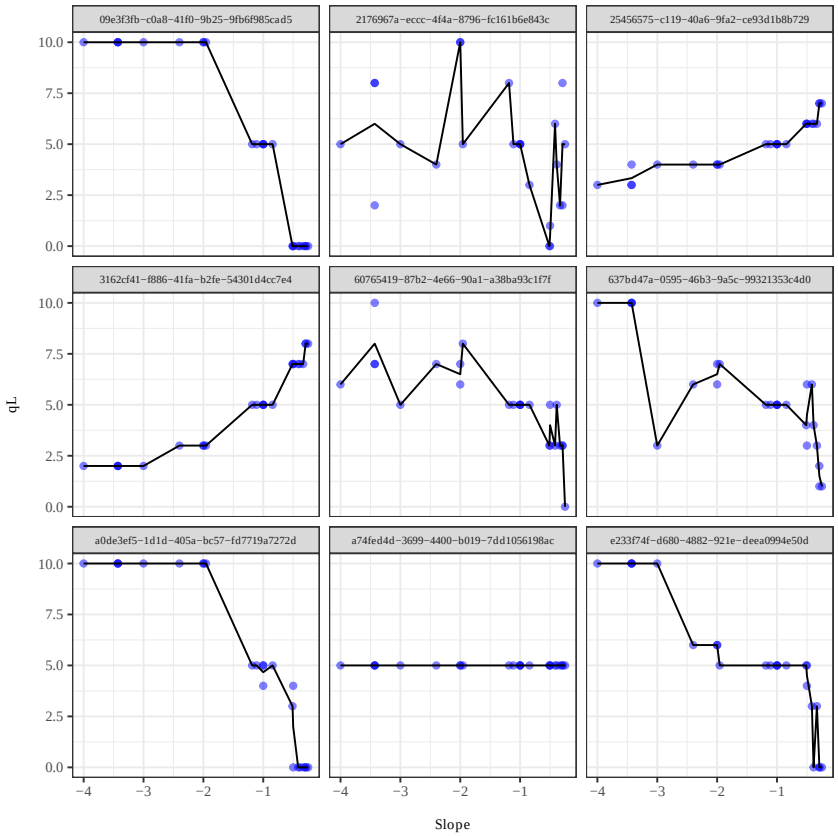
Table 38: Social | Heuristics of representative self-reports

Behavioral rule	**Simplified self-report**	**n**
Balance outcomes	*"I distributed tokens to keep gains roughly equal between myself and the other person, minimizing differences whenever possible."*	15
Maximize total payoff	*"I focused on maximizing total gains, either for myself or the other, depending on which option offered the highest outcome."*	12
Compassion	*"I sometimes gave more to the other participant to ensure they received a reasonable amount, even if it slightly reduced my own gain."*	10
Selective self-interest	*"I allocated tokens to optimize my own gain while leaving a minimum for the other participant, or adjusted depending on relative token value."*	9
Internal coherence	*"I applied consistent rules for distributing tokens, adjusting based on prior choices or payoff differences."*	8
Altruism	*"I prioritized giving more to the other participant when it improved their overall gain."*	5
Selfishness	*"I sought to maximize my own gain regardless of the other participant, unless fairness required otherwise."*	5
Mixed strategy	*"I varied my allocations across rounds to balance between fairness, self-interest, and maximizing gains."*	4

I Response Signatures

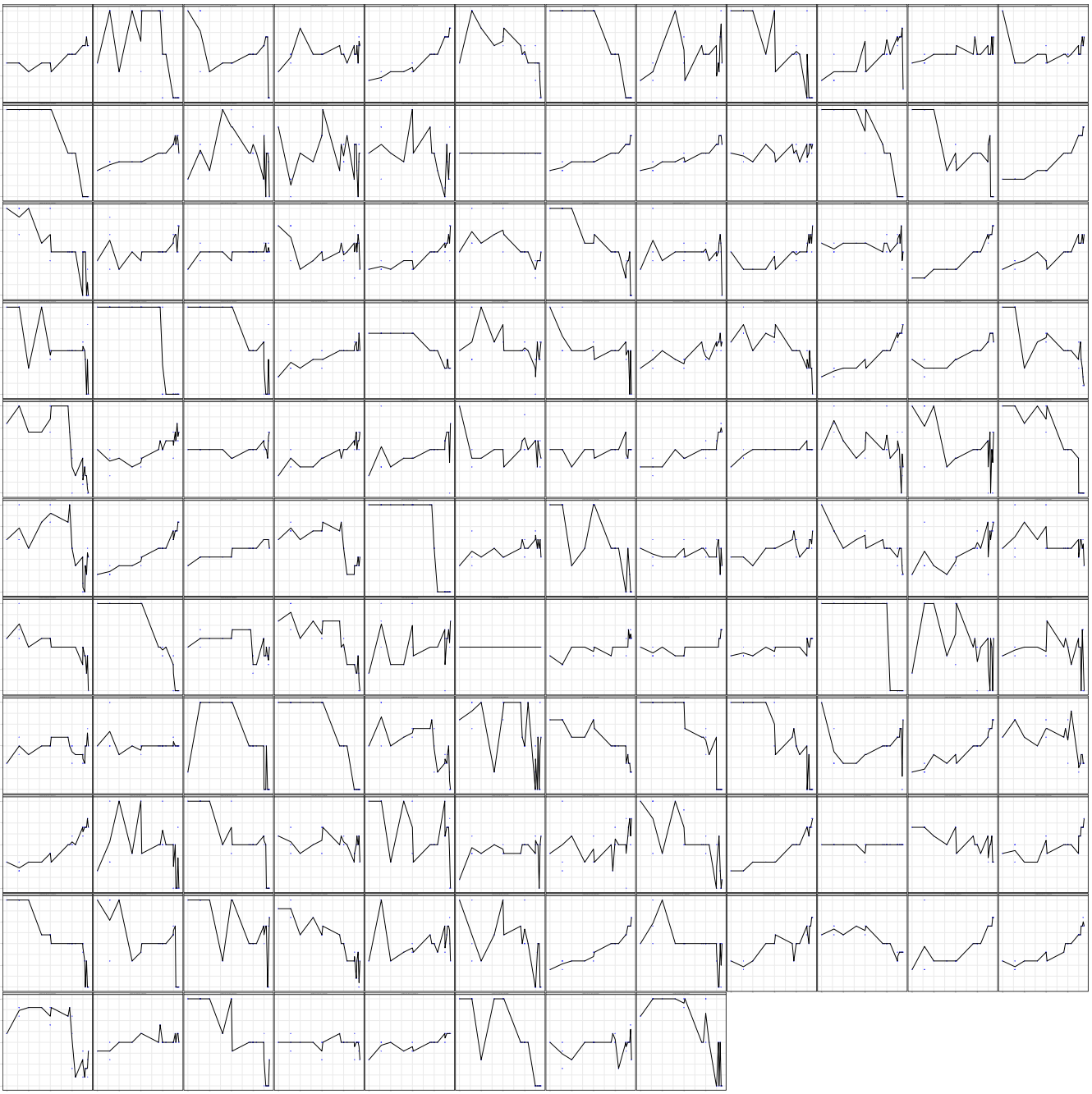
I.1 Risk Behavior

Figure 18: Risk | Response signatures of 9 selected subjects



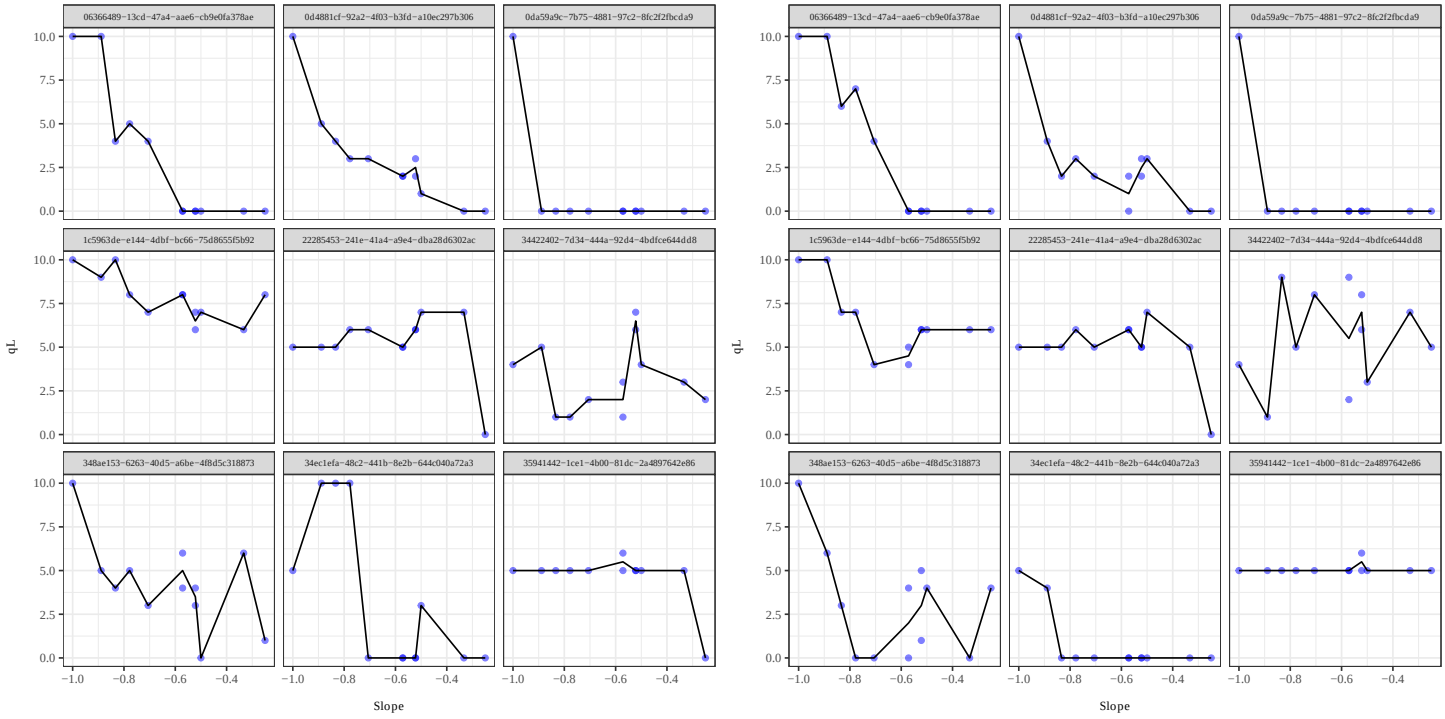
Note. A frequent response pattern is illustrated in the top-left panel and in the panels on the left and right of the bottom row. Subjects allocate more to the event with the higher return and choose more equal allocations when the returns of the two events are similar. Subjects differ in how consistently they follow this strategy. For instance, the subject in the top-left panel exhibits very consistent behavior, whereas the one in the bottom-right panel shows somewhat less consistency. This type of behavior aligns with the goal of maximizing the payoff conditional on the event occurring. A substantial fraction of subjects, however, display the opposite pattern, i.e., a positive response function. Examples are shown in the middle-left and top-right panels. These subjects appear more cautious, aiming to avoid the risk of ending up with nothing or with only a small payoff. They balance allocations more evenly across events. More rarely, we observe subjects like the one in the middle of the bottom row, who do not respond to the slope at all and instead adopt a form of naïve diversification strategy. In addition, some subjects exhibit unpredictable behavior that does not fit neatly into the above strategies. Examples can be found in the top-middle panel and in the two panels on the right of the middle row. Still, the nature of deviations varies. For the subjects in the middle row, behavior appears largely consistent but features occasional departures from it, whereas the subject in the top-middle panel exhibits "consistently" unpredictable responses.

Figure 19: Risk | All 128 response signatures



I.2 Time Behavior

Figure 20: Time | Response signatures of 9 selected subjects for sooner and later



Note. The majority response pattern for both *sooner* (0 vs. 2 weeks) and *later* (2 vs. 4 weeks) tradeoffs is negative, as expected. Specifically, impatience predicts that subjects should allocate more to the sooner point in time when the compensation for deferring payoffs is low, and more to the later point as the compensation increases. This pattern is clearly visible in the *sooner* and *later* figures for the subjects in the top row. Once again, individuals differ in how consistently they respond to changing interest rates: depending on their degree of impatience, they display flatter or steeper response functions. Only rarely do subjects show no response to changing interest rates, such as the one illustrated in the bottom-right panel. Some subjects exhibit more unpredictable behavior, adjusting their allocations inconsistently as the slope changes back and forth. Broadly, the vast majority save more as the compensation for waiting increases, consistent with impatience. Specifically, 93 out of 125 subjects (74%) display a decreasing profile, 6 subjects (5%) show no discounting (flat profile), and the remainder exhibit unspecified or unpredictable patterns.

Figure 21: Time | All 125 response signatures for sooner

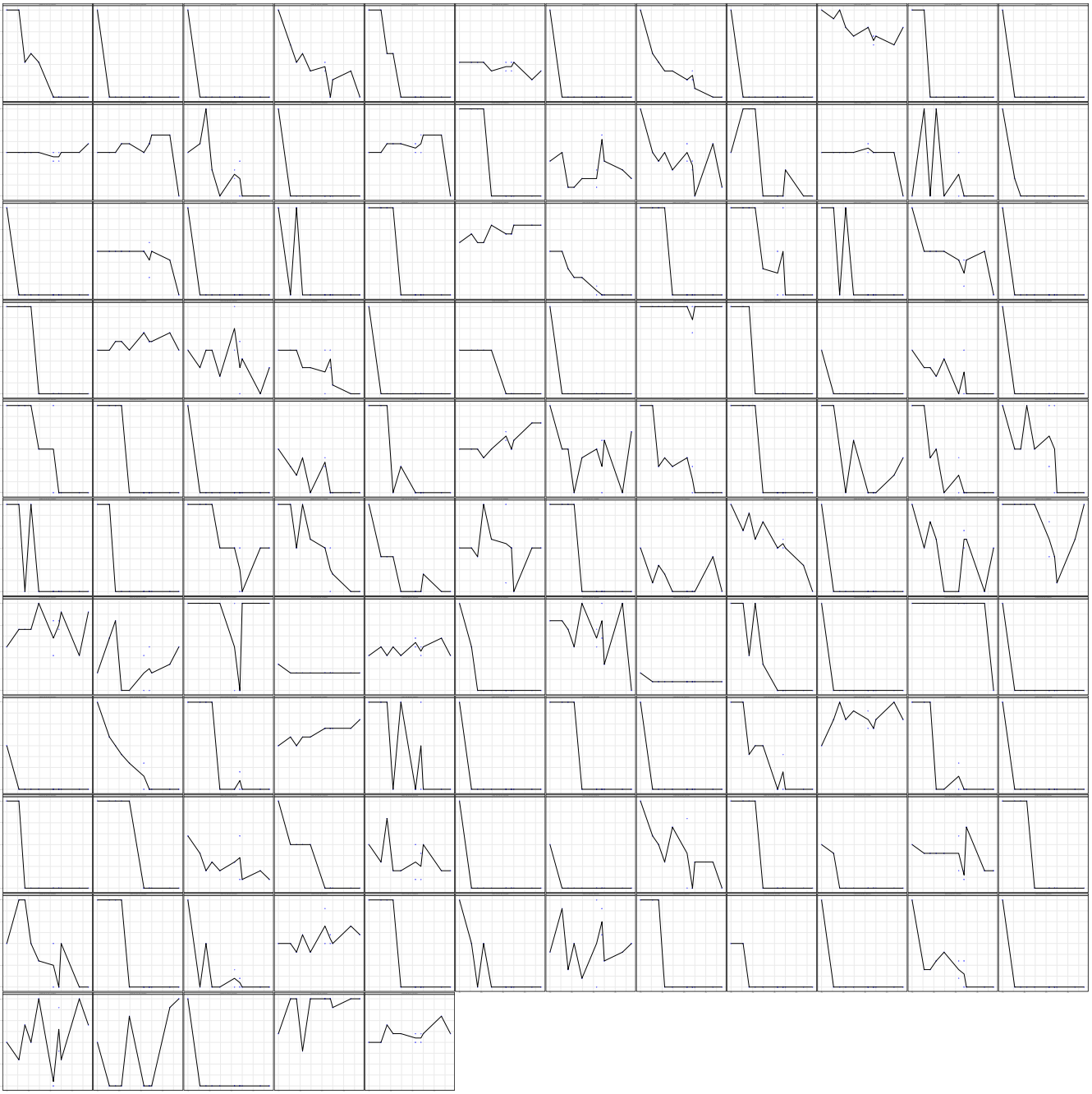
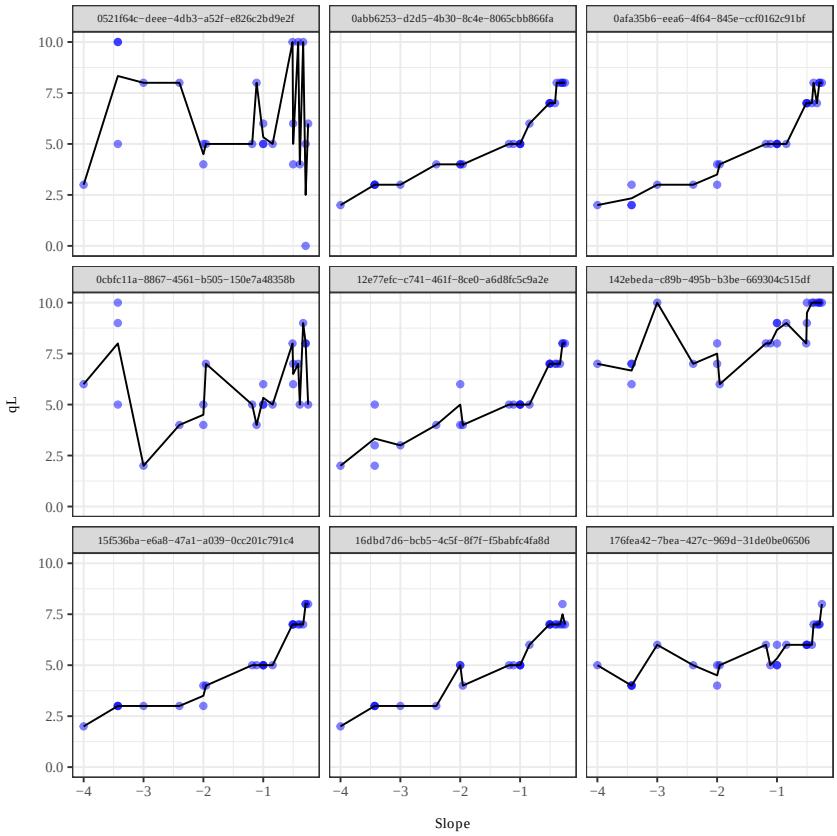


Figure 22: Time | All 125 response signatures for later

I.3 Social Behavior

Figure 23: Social | Response signatures of 9 selected subjects



Note. By far the most prevalent pattern is an increasing response profile. Subjects differ in how consistently they respond to changes in the slope of the budget line. In general, they are willing to give more to the other when the cost to themselves is small, but this willingness decreases as their own cost increases. A smaller group of subjects exhibits more unpredictable behavior, switching back and forth between giving more or less to the other.

Figure 24: Social | All 125 response signatures