2020

01

Working Paper

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS (IPP)

USING EYE-TRACKING TECHNIQUES TO UNDERSTAND THE ROLE OF ATTENTION ON CHOICE AND REVERSALS

Raúl López-Pérez and Eli Spiegelman

INSTITUTO DE POLÍTICAS Y BIENES PÚBICOS – CSIC

Instituto de Políticas y Bienes Públicos
Consejo Superior de Investigaciones Científicas
C/ Albasanz, 26-28
28037 Madrid (España)

Tel: +34 91 6022300
Fax: +34 91 3045710

http://www.ipp.csic.es

How to quote or cite this document:

López-Pérez, R. and Spiegelman, E. (2020). Using eye-tracking techniques to understand the role of attention on choice and reversals. Instituto de Políticas y Bienes Públicos (IPP) CSIC, Working Paper. 2020-01

Available at: digital.csic.es

Using eye-tracking techniques to understand the role of attention on choice and reversals[*]

Raúl López-Pérez[†] and Eli Spiegelman[‡]

Abstract: A preference reversal (PR) refers to behavior that violates revealed preference or is simply incoherent – i.e., not explainable by a rational ordering. In a classical PR experiment, for instance, participants often exhibit greater risk aversion in a Choice-based revelation procedure than in an Evaluation-based one, i.e., choose the safer of two gambles but express a higher monetary valuation for the riskier. We conjecture that PRs are partly due to the interaction between attention and task mode, and explore three compatible explanations using eye-tracking techniques. Explanation 1 says that difficult tasks require more time to be performed without 'mistakes'. Those who pay scarcely more attention on Evaluation than Choice, therefore, are more likely to act incoherently. Our data corroborates such a prediction. Explanation 2 assumes that the subjective value of (i) a bet or (ii) any of its attributes (prize and winning probability) depends on the attention paid to it. PRs occur when people allocate attention to different elements across tasks. In line with recent models of drift-diffusion, we find evidence consistent with point (i): a higher focus on the safer bet during Choice predicts PRs. In contrast, little evidence supports the idea (ii) that the share of fixations on probabilities versus prizes influences behavior or PRs. Explanation 3, finally, states that the nature of the tasks may affect the comparisons people make between the options, which are relevant for behavior. For instance, the cognitive difficulty of pricing a bet in Evaluation could distract attention from the relative risk across bets, thus reducing risk aversion. In our design both bets are visible on the computer screen in *both* tasks, and subjects make substantially more transitions between bets in Choice. Yet this is observed among all participants, not only the reversers.

Keywords: Eye-tracking, Limited Attention, Mental Effort, Rationality, Reversals.

JEL Classification: D01; D81; D83; D91.

---

[*] We are grateful to […], and participants at seminars in the IPP-CSIC […] for helpful comments and suggestions. We also gratefully acknowledge financial support from the Spanish Ministry of Economics, Industry and Competitiveness through the research project ECO2017-82449-P (Raúl López-Pérez), and helpful research assistance by […].
[†] Institute of Public Goods and Policies (IPP), Spanish National Research Council (CSIC), C/ Albasanz, 26–28, 28037, Madrid, Spain. E-mail address: raul.lopez@csic.es
[‡] LESSAC Research Group, CEREN, EA 7477, Burgundy School of Business - Université Bourgogne Franche-Comté, 29 rue Sambin, Dijon 21000, France. E-mail: eli.spiegelman@bsb-education.com

1. Introduction

Consider a decider facing two problems with *exactly* the same choice set, although her preferences are elicited by means of different tasks in each case, e.g., directly choosing between the options or assigning a monetary value to each one. Ruling out multiple optima, a preference reversal (PR) occurs when she chooses a different option in each problem. Reversals are incoherent choices, and nearly 50 years of experimental evidence has shown that they are not a negligible phenomenon (Lichtenstein and Slovic, 1971, 1973; Tversky and Thaler, 1990; Seidl, 2002). This seems to mean either that preferences are not stable with respect to task mode, or that behavior is not systematically reflective of underlying preferences. Neither option is good, as both are fundamental challenges for rational choice theory and are also key from a policy perspective. If choices do not reliably reflect stable preferences over the alternatives, the justification for such measures as surveys, marketing focus groups, or even elections is called into question.

Our essential vision is that preferences are stable *given a set of stimuli to which the decider pays (sufficient) attention*. Different task modes, however, can focus or demand attention in different ways, and the resulting changes in attention generate divergent choices across tasks. We test several variants of this idea in a PR lab experiment. Subjects are faced with six simultaneous pairs of binary bets, and asked to perform two tasks for each, in random order: (i) Choice and (ii) Evaluation. In (i), they choose between the two bets comprising each pair, while they assign a monetary value to each individual bet in (ii). They then play either the bet chosen in task (i) or the bet priced higher within the pair in (ii), randomly selected. The bets are such that in each pair one of them, called the *P-bet*, has a higher probability of winning, while the *$-bet* has a higher prize; e.g., the P-bet ($-bet) gives 8 (24) Euros with a probability of 0.97 (0.33), and nothing otherwise.[1] Since visual attention arguably correlates with mental attention in tasks (i) and (ii), further, we use eye-tracking equipment to measure the number and length of "fixations" on various parts of the available bets.[2] Thus we can study in detail the attention patterns behind the PRs.

Following our essential vision stated above, our analysis checks two questions, **Task**: attention is focused differently across tasks, and **Behavior**: attention correlates with revealed preferences in each task. If we find that both Task and Behavior are validated, then we find some ground for our

[1] A robust phenomenon in PR experiments is the so-called standard PR, where the P-bet is chosen instead of the $-bet, but then the latter bet receives a higher value. That is, pricing, a task that economic agents are supposed to perform frequently in (say) financial markets, is often incoherent with direct choice. We note also that the opposite pattern to the standard PR, i.e., choosing the $-bet but assigning it a lower value, is relatively infrequent in the literature.

[2] This kind of equipment physically measures where and for how long participants look on the computer screen. As its costs and obtrusiveness have fallen over the past two decades or so, it has become a more appealing empirical tool in behavioral sciences (Lahey and Oxley, 2016).

Reversals conjecture, i.e., (in)attention is a mediating variable for reversals.[3] We also check this directly by comparing attention patterns among reversing and non-reversing bet pairs. To organize the analysis of our research questions, we consider three specific, nonexclusive accounts of the link that attention may represent between Task, Behavior and Reversals.

A first account (**Depth of Reflection**) stresses the effect of what we call the extensive margin of attention, i.e., the total effort devoted to a task. Our vision of preferences as depending on the details of attention implies that if agents don't pay "enough" attention to a task, their responses may change. Other relevant individual characteristics constant (e.g., intelligence), that is, insufficiently attentive agents should make more *mistakes*, in the sense of revealing preferences different from those they would upon further reflection. At the level of the bet pair, even random "spherical" mistakes among those agents could lead to systematic inconsistencies if they affect one task more than the other, say because it is more difficult and so behavior "converges" faster in the other.[4]

We find evidence in our experiment in line with Depth of Reflection. As seems perhaps natural, the overall level of attention to the bets in a pair differs across tasks, as people dedicate significantly more time on average to Evaluation than to Choice, i.e. to the more cognitively demanding task. Further, the overall attention during Evaluation (but not Choice), correlates with reversals. By way of indication, the "perfect score" (zero reversals out of six pairs of bets) participants spent 19.8s on each pair in the Evaluation task, compared to 10.4s by those with at least one reversal, nearly doubling the time spent. In the Choice task, in contrast, those with no reversals focused an average of 8.9s on each pair, compared to 7.8s for those committing at least one reversal, a mere 14% increase. These data hint that insufficient attention in Evaluation partially drives PRs, and hence *apparently* contrast with studies on response times (RT), e.g., Alós-Ferrer et al. (2016), where PRs are associated to higher RTs in Choice; we discuss this point in more detail throughout the paper.

A second account (**Salience**) predicts that a decider's choice is influenced by the so-called intensive margin of attention, i.e., the way in which she allocates her limited attention budget to each

[3] A key feature of our experimental design is that the decision contexts, i.e., the screens for both Choice and Evaluation, are as similar as we found possible. When choosing and pricing a pair of bets, in particular, *both* bets are presented in the screen (in most experiments on PRs, in contrast, the bets are evaluated in isolation). In other words, the attention set, i.e., the set of stimuli to which the decider can pay attention, is basically identical across tasks. In this sense, we do not restrict or "force" subjects to display some specific attention patterns in a task. This is critical to explore the differences across tasks in the allocation of attention, preventing confounds with a potential effect of the context. Our results suggest that PRs are partly an effect of the attention patterns each task mode motivates, even keeping the attention set constant.

[4] As an aside, note that not all mistakes will be observable, so that the PRs represent a lower bound. Consider, for instance, an agent who prefers the P-bet to the $-bet: mistaken evaluations that make it seem even more preferred will not be observable. Since subjects often choose the P-bet in the easier Choice task (perhaps due to risk aversion), it follows that, consistent with the literature, the standard PRs should occur more frequently than the non-standard ones. Such an effect could be strengthened by other mechanisms that distort the evaluations directly, for instance the "anchoring and insufficient adjustment" explanation suggested by Lichtenstein and Slovic (1971).

part of the decision screen in our computerized experiment. This is motivated by a psychological literature on salience, e.g., Taylor and Thompson (1982), which claims that when people pay more attention to some (salient) stimuli than to others, the information contained in the former will have a greater weight in subsequent decisions. The idea of salience is also related to the eye-mind hypothesis of Just and Carpenter (1980), which presumes that individuals cognitively process whatever their eyes are currently directed at, and the drift diffusion model (DDM) in Krajbich et al. (2010), predicting a choice bias for the options that are fixated on more.

If for some reason the different tasks in the PR experiment drive attention in divergent ways, making different parts of the decision salient for the two choices, salience could explain the PR phenomenon.[5] One possibility in this line is that reversers pay relatively more attention to one bet in one of the tasks. Another, subtler possibility is that reversers pay different attention to the bet attributes, i.e., winning probabilities and prizes, across tasks. For instance, a subject who focuses on probabilities should plausibly be attracted to the safer bet as a consequence; similarly, focus on the prizes should favor the bet with the higher prize. If Choice encourages the former and Evaluation the latter, then reversals may result. As we discuss in Section 2, some influential psychological theories of PRs including Prominence and Compatibility, Tversky et al. (1988), are in line with several specifications of this attention-on-different-attributes mechanism.

Our results suggest some support for salience, particularly across bets in a pair, rather than across attributes of a bet, and particularly through the Choice task. On one hand, the task influences relative focus across bets: subjects focus more on the P-bet (relative to the $-bet) in Choice than in Evaluation. The difference, while not large in numerical terms (46.1% of attention is on the P-bet in Evaluation, compared to 49.1% in Choice), is significant. Furthermore, focus correlates with behavior in the Choice task but not in Evaluation. Those who choose the P-bet over the $-bet spend 53.6% of their attention on it, compared to 45.4% for those who choose the $-bet. This difference is also significant, and perhaps tellingly, the modal attention is on the chosen bet, on average, as predicted by the DDM models. It also seems related to consistency. Regression analysis shows that a higher proportion of attention to the P-bet *during Choice* significantly increases the probability of a standard reversal. On the other hand, the task does not affect the allocation of attention between the prizes and probabilities in our data, and neither do we find much evidence that relative focus on probabilities or prizes preferentially determines behavior in either task. The salience-based driver of

[5] This would not be at odds with the Depth of Reflection account, as the differences in attention due to salience could well be conditional on the "insufficient" total attention described above.

reversals seems to come mainly from between-bet factors, which are more relevant to Choice than to Evaluation.

The object of attention is a single bet or attribute in the Salience story. Yet the cognitive processes that drive behavior may involve comparisons between those objects, and it is hence interesting to look at sequences of fixations, which may indicate a broader concept of attention. This is what we do with the analysis of **transitions** (Devetag et al. 2016) which, loosely speaking, occur when the decider shifts attention from one point of the problem to another.[6] Several theories motivate this approach. In the model of context-dependent preferences by Bordalo et al. (2012), for instance, reversals occur when the bets are evaluated separately, as people are assumed to compare each bet with a reference bet giving zero with certainty.[7] While this specific mechanism arguably plays little role in our experiment, where evaluations are made in pairs, it hints that PRs might have something to do with the type of comparisons that people make when deciding. It may also relate to the evaluability hypothesis (Hsee et al. 1999), which posits that some kinds of data are harder for people to assess in isolation than others. For instance, a single probability is an unfamiliar object to most people untrained in statistics, and may become more meaningful to them – and hence take more weight in the decision – when assessed in comparison with another probability.

Applying our three questions, Task, Behavior and Reversals, to transitions reveals a strong effect of task type, but perhaps surprisingly little in terms of behavior or reversals. Choice greatly amplifies the share of across-bet transitions, compared to the Evaluation task. This could not be tested in the previous literature, where evaluations were done in isolation and therefore across transitions could not be observed. On the other hand, differences across individuals in transition patterns are largely orthogonal to the behavior in each task, predicting neither choice of the P- or $-bet, nor evaluations of either. (Nor, consequentially, reversals.) In summary, the reversers, i.e., those who make inconsistent choices, are characterized by (a) paying scarcely more attention in Evaluation than Choice, and (b) focusing on the P-bet during Choice (and not in Evaluation). In contrast, they are not characterized by, say, making relatively few comparisons across bets in Evaluation.

[6] For instance, for any pair of bets, within-bet transitions can be defined as attention shifts between the probabilities and the prizes of a particular bet, while across-bet transitions move from an attribute of one bet to one of the other.

[7] Bordalo et al. (2012) implicitly postulate that the decision weight of a bet payoff depends on the attention it attracts, which in turn depends, roughly speaking, on how different it looks with respect to the average payoff yielded by the other lotteries in the same state of the world. Since the $-bet has a very salient payoff, an over-valuation entails. See Mormann and Frydman (2018) for a test of the hypothesis that payoffs exhibiting greater variation (in percentage terms) attract more attention, a point we do not explore here. See also Kőszegi and Szeidl (2013) for an alternative model and Gabaix (2019) for an excellent review of the behavioral theories incorporating inattention.

In the following we first review some related literature, both empirical and theoretical. Then we describe and discuss our experimental design. Section 4 presents the data and results. Section 5 concludes with a summary and a discussion of some implications of our findings.

2. Literature review

Different explanations have been suggested for the reversals. Broadly speaking, these can be categorized into three groups, following Loomes and Pogrebna (2017). First, explanations based on *intransitive preferences* such as the regret theory of Loomes and Sugden (1982) posit that PRs are due to intransitivity inherent in preferences, and so reflect a natural feature of otherwise stable preference maximization.[8] These theories are not coherent with the accounts that we explore in this paper and in fact would predict that PRs are independent of attention patterns. We note that Tversky et al. (1990), Bostic et al. (1990), and Loomes and Pogrebna (2017) offer experimental evidence at odds with the idea that intransitivity is the major cause of PRs.

Second, a variety of psychological explanations focus on *procedural biases*, largely based on the idea of *contingent weighting* (Tversky et al. 1988). This posits that some parts of a decision problem take greater weight in the deliberative process, and that the attribute weights depend in systematic ways on the task context. Tversky et al. (1988) also argue that weighting is closer to lexicographic in Choice compared to other tasks; due to risk-aversion, specifically, probabilities are more heavily weighted in Choice, a phenomenon called the prominence effect. For illustrative purposes only, suppose that the psychological value of binary bet (p, z), where p denotes the probability of winning prize z, is $V(p, z) = \alpha \cdot v(p) + u(z)$, given some monotonic subjective value functions v(p), u(z), and parameter $\alpha > 0$, where a "high α" essentially indicates high risk aversion. Prominence amounts to saying that α varies with the task, and more specifically that it is higher in Choice than in Evaluation.

An additional and closely related mechanism is the compatibility effect, which states that in pricing a bet (response), people will put too much weight on its money payoffs, that is, the stimulus most 'compatible' with the response. Using our terminology above, the effect basically implies that α is lower in Evaluation. This leads to the overpricing of the $-bets, of which Tversky et al. (1990) offer abundant evidence, e.g., a person who prefers, say, $10 for sure over a 1/3 chance to win $40, assigns to this bet a cash equivalent that exceeds $10 – see also Loomes and Pogrebna (2017). In the Task goal effect (Fischer et al. 1999), the prominent attribute depends on whether the task is qualitative (differentiation, as in choice) or quantitative (equalization or matching, as in evaluation).

[8] Regret theory assumes that people evaluate the outcomes of the risky choices based not only on what they yield given the action and the state, but also on what they would have had in that state given an alternative action. They experience a "rejoicing" utility benefit from "having chosen right" if the alternative action would have given them less than they actually got, and a "regret" cost from "having chosen wrong" in the converse case.

The former encourages lexicographic evaluation, effectively increasing α in Choice. The evaluability hypothesis mentioned above (Hsee et al. 1998) is psychologically different, suggesting that probabilities are more difficult to comprehend in isolation than are money quantities, and thus α diminishes in Evaluation relative to Choice if the evaluation of each bet is performed separately.

These psychological theories can be naturally combined with the Salience idea, so that the weight of each bet attribute correlates with the attention paid on it (which is in turn mediated by the task). Discussing compatibility, for instance, Tversky and Thaler (1990, p. 208) speak of the hypothesis that "people focus their attention on the stimulus components that are most compatible with the response mode". Along these lines, the psychological theories would predict a focus on payoffs (probabilities) in Evaluation (Choice), particularly among the reversers. As we noticed above, however, the task does not affect the allocation of attention between prizes and probabilities in our data, and neither do we find much evidence that relative focus on one attribute or another determines behavior in either task. Indeed, reversers and non-reversers are similar in this respect.

It must be noted that the prior theories are fundamentally different from Depth of Reflection, which amounts to saying that the weight α depends on the overall attention paid on the task, particularly in difficult tasks. Specifically, α should be the same across tasks if 'enough' effort is devoted to each one, but dissimilar otherwise.[9] Depth of reflection, that is, predicts reversals even if people pay similar attention to any attribute, provided that they exert insufficient effort in the difficult tasks. This idea coheres well with our finding that reversers pay significantly less attention in Evaluation, but also with some studies showing that reversals decrease when the task mode demands attention. In the experiments in Bostic et al. (1990), for instance, certainty equivalents are elicited either by (i) asking subjects the amount of money *x* that leaves them indifferent between *x* and the gamble, or by (ii) an iterated procedure in which subjects are first asked to give their preference between the gamble and a fixed sum of money; if the subject prefers the gamble (money), the question is made again, but the amount of money is increased (decreased) by $0.04; this series of questions is repeated until the preference changes. Bostic et al. reported that the frequency of PRs dropped substantially when procedure (ii) was used. Similarly, PRs decrease with the number of repetitions of experiments –see Seidl (2002) for references in this respect. Experienced subjects tend to increase the evaluation of the P-bets, but also choose them less often.

The psychological theories cited before are also different than another potential specification of Salience, according to which the weight α in a bet depends on the overall attention paid to that bet,

[9] More precisely, the value of α need not be equal, nor must it be the same across individuals for any task. But to avoid reversals, the task form itself should not influence the relative value of α.

independently of how attention is allocated among its attributes. If a subject pays more attention to the P-bet in some task, say, then the α of that bet is larger than that of the $-bet. This account is compatible with Depth of Reflection if the positive relationship between α and the attention paid on the bet is restricted to hold only when the decider pays insufficient attention in the difficult task, i.e., Evaluation. Our data is more consistent with this specification of Salience, as (i) people tend to choose the bet they look at more overall, and (ii) reversers tend to exhibit dissimilar patterns of attention to the bets across tasks, i.e., pay more attention to P-bet in Choice, but not in Evaluation.[10] Without rejecting the role of regret, prominence, or compatibility, in summary, our and prior results suggest that other forces could be also at play in explaining PRs.[11]

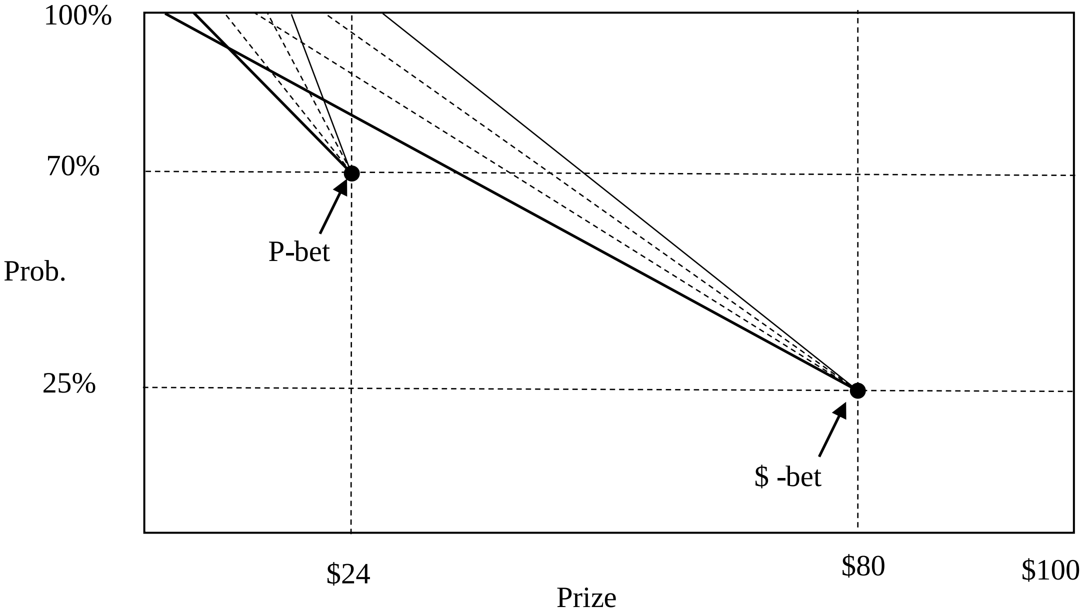
The third variety of explanations includes more recent theories invoking *imprecision and probabilistic preferences* in the decision maker's deliberative process. Building on a model developed intuitively by MacCrimmon and Smith (1986), Butler and Loomes (2007) assume that the computation of certainty equivalents is affected by imprecision, in that given bet b people find it hard to be sure whether any particular value within some "imprecision interval" is preferable to or less preferred than b –see also Schmidt and Hey (2004).[12] Further, the width of that interval increases as b becomes more dissimilar from a sure positive payoff, as happens with the $-bet compared to the P-bet. Implicitly assuming that people report the certainty equivalent of b randomly within its imprecision interval (possibly assigning more probability to the higher values), a bias in the direction of standard PRs follows. Although Butler and Loomes (2007) make no assumption on attention, one could intuitively expect higher imprecision in those who do not make the 'extra effort' when evaluating bets in our experiment. Figure 1 below adapts Figure 1.a from Butler & Loomes (2007) to show how increased attention might reduce the scope for possible reversals, by narrowing the range of uncertainty intervals for the pair of bets. Solid rays emanating from each bet show uncertainty zones with "low" attention, while dashed ones show zones with "high" attention. The latter are narrower ranges, as increased attention allows decision makers to more clearly identify their preferences. The overlap in the former yield the possibilities for reversals in Butler and Loomes (2007); we see this is greatly reduced in the latter. Thus our study of attention adds a nuance to

[10] Our Transitions account would say that the weight of an attribute depends on how it compares to the same attribute in the other bet(s). That is, relative comparisons are also weighted, conditional on attention. We do not develop this idea further as the data does not seem to support it.

[11] Cubitt, Munro and Starmer (2004) provide evidence at odds with intransitivity and any of the psychological explanations in isolation, although they point out that a combination of prominence and compatibility organize their data to some extent.

[12] Blavatskyy (2009) develops a closely related formal model of probabilistic choice. Note that the three accounts described in the introduction differ somewhat from that of imprecise preferences. We suppose that preferences are precise, given attention, but that at least for low overall levels of attention, the attributes of the problem considered may be random. The resulting randomness of behavior will be observationally equivalent, however.

theories of imprecision; conversely, imprecision can be thought of as another way of looking at the effect of low attention.



Note: Adapted from Butler and Loomes (2007, p. 280). Participants are unsure whether any alternative in a horizontal cut of the rays is more or less preferred to the bet where they originate. The "tighter focus" of dotted rays illustrates the effect of greater attention. Prob. = probability of winning prize.

Figure 1: Illustration of how attention might focus imprecision

Our paper is also related to a body of literature in Economics that analyzes the decision processes underlying PRs, using measures of *non-visual* attention. Process tracing studies by Johnson et al. (1988) and Schkade and Johnson (1988) used MouseLab, where participants move a computer mouse to uncover occluded information, allowing researchers to track the order in which information is accessed. In their studies, subjects spent more time during evaluations than choices, as in our study, but spent proportionally more time looking at probabilities during choices than during bidding, a difference across tasks that we do not find here. A potential issue with MouseLab, however, is that it introduces some artificial costs of obtaining information, which might influence information search processes, e.g., Franco-Watkins and Johnson, 2011. Some papers use response times (RTs), i.e. the time elapsed between the presentation of the relevant information and the decision.[13] Alós-Ferrer et al. (2016) explore the interaction between PRs and RTs in the Choice task, finding that PRs of either type are associated to *longer* RTs in that task. We discuss the similarities

[13] Since RTs cannot discriminate the focus of attention, however, they cannot be used to test Salience or study Transitions. A further issue concerns the comparison between the Choice and Evaluation tasks using RTs. The fact that evaluations require typing numbers, which is more time-consuming than clicking a button in Choice, increases RTs in Evaluation, but this additional time possibly does not signal any further reflection. In contrast, a comparison of visual fixations on the screen across tasks seems to reflect more accurately the differences in cognitive effort.

and differences between our results and these in Section 4.2.2. They also report that, in choices involving no PRs, RTs are longer when the $-bet instead of the P-bet is chosen, and show their findings to be in line with a simple model that combines the idea of imprecise preferences, Butler and Loomes (2007), and the compatibility hypothesis of Tversky et al. (1990).

To finish, our study contributes to a literature that studies the interaction between attention and choice, using eye-tracking techniques. This research has identified several relationships between gaze and choice. In particular, people tend to opt for those options that (i) they looked at first, (ii) focused on at the time of choice, and (iii) looked at more overall; see Orquin and Mueller Loose (2013) and Krajbich (2019) for reviews. As we have noted, our data is in line with (iii); points (i) and (ii) are out of the scope of this paper. More closely related to our use, several studies analyze gaze and risky choice. Arieli et al. (2011) examine whether individuals calculate expected values and measure what components of bets their participants were focusing on as a proxy for cognitive processes. Consistent with our data, Stewart et al. (2016) conclude that risky options receiving more attention are more likely to be selected. In contrast, however, all attributes (i.e., probabilities and outcomes) are fixated about equally in their study, whereas we observe more fixations overall on probabilities in Choice and Evaluation. Finally, Kim et al. (2012) is possibly the study closest to ours, as they use eye-tracking techniques in a PR experiment. Yet their research questions differ substantially from ours, which explains some differences in the designs. For instance, they do not explore Depth of Reflection and Transitions, two questions that we can neatly address because we do not follow the standard methodology of showing participants different stimuli across tasks. For another example, they explore first fixations in each task and calculate risk aversion for each participant independently in each task in a comparable way; these issues are out of the scope of our study. For the calculation of risk aversion, they use 50 different pairs of bets, and both Choices and Evaluations are repeated twice in each pair (in the repeated choices, the two bets in a pair switched positions in the screen), which means a total of 200 decisions –subjects were paid according to one of these decisions, randomly chosen at the end of the experiment. Note also that they elicit cardinal evaluations using a Becker-Degroot-Marschak procedure, while we follow the ordinal pricing mechanism employed by Kahneman et al. (1990). Regardless, some of their results coincide with ours, specifically subjects (N = 24) spent significantly more time on the Evaluation than the Choice tasks, took longer to evaluate the $-bets than the P-bets, and fixated more on the P-bet during Choice if they were more risk-averse, i.e., chose more often the P-bet. In contrast, they report few significant individual differences in attention patterns and behavioral measures like the number of PRs –this is perhaps due to little variability in that number, as the authors suggest. For a final difference, they report test results showing that people look at probabilities more during Choice and prizes more during Evaluation, something not replicated by our regression analysis. Therefore, while their evidence seems in line with the contingent weighting hypothesis, ours suggests that the effect is somehow fragile and hence demands further exploration.

3. Empirical context

3.1 Experimental design

During the experiment, which was computerized, subjects were faced with a total of twelve bets, grouped in six pairs (see Table I). Although these were not indicated explicitly to subjects, in each pair of bets there was one (the *P-bet*) that had a higher probability of winning, and one (the *$-bet*) that had a higher prize. All probabilities were expressed as multiples of 1/100 (for example, 13/100 to win € 24.00). The bets in each pair had similar expected values, although slightly lower in general for the P-bets so as not to favor their choice and hence have some variability.[14]

Pair	P-bet	Expected value	$-bet	Expected value
1	(97/100, 8)	7.76	(33/100, 24)	7.92
2	(73/100, 10)	7.3	(27/100, 30)	8.1
3	(87/100, 9)	7.83	(10/100, 80)	8
4	(63/100, 7)	4.41	(17/100, 40)	6.8
5	(90/100, 5)	4.5	(13/100, 35)	4.55
6	(87/100, 3)	2.61	(20/100, 13)	2.6

Note: The pair (p, x) denotes a bet with a probability p to earn x Euros, and nothing otherwise. Each expected value refers to the bet immediately on the left.

Table I: P and $-bets employed in the experimental design

For each pair, the subjects faced two different tasks. In the *Choice task*, participants simply clicked a button to show which of the bets was preferred, or to indicate indifference between them. In the *Evaluation task*, participants assigned a numerical value to each bet. The order of the tasks was randomized, and participants were presented with a button on the screen to toggle between tasks at will. Participants were told they would play one bet for real money, from a randomly chosen pair. Specifically, we implemented the ordinal pricing mechanism (Tversky et al, 1990), so that they played either the bet they chose in the Choice task for that pair, or the one to which they assigned a higher value in the Evaluation task. Within each task, importantly, all six bet pairs were presented simultaneously on the screen. The ordering of pairs, and bets within pairs, was randomized across participants, but identical across tasks for each participant. We did not switch the placement of the amounts and probabilities, which Kim et al. (2012) report not to have any significant effects on

[14] This choice, and its motivation, were inspired by Tversky et al. (1990).

choice and bidding behavior. Figure 2 shows a screenshot of the placement of the bets in Choice; the screenshot was very similar in Evaluation, except that in place of the buttons indicating a preference for bet A or bet B, subjects saw a field under each bet in which to enter a value, and the buttons for indifference and resulting arrows did not appear.

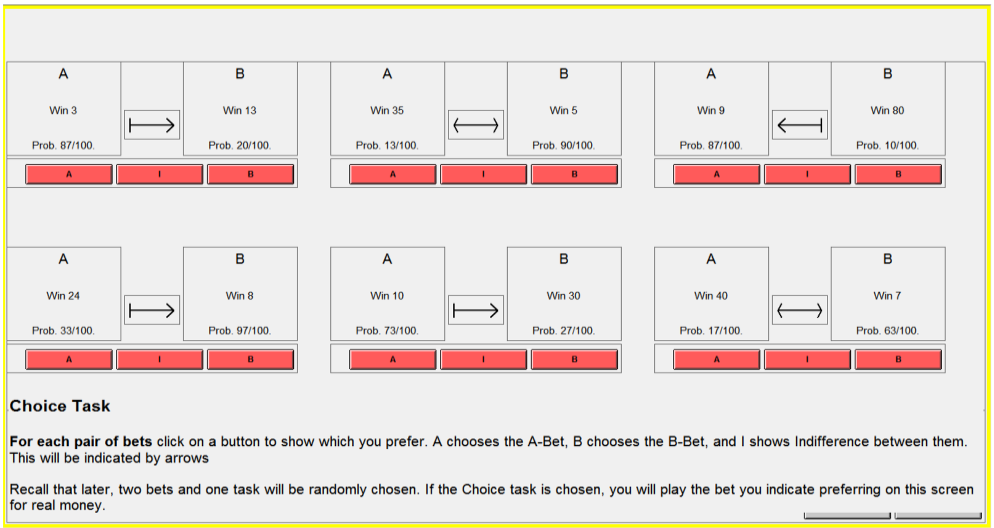


Figure 2: Interface for the experimental choice task

Our design is unusual in the literature, which most often shows bet pairs in isolation during the choice task, and individual bets one at a time in the evaluation task. We feel that this design allows a cleaner analysis of our research hypotheses than would the more standard one. First, presenting one single bet on the screen effectively constrains the intensive margin of attention in Evaluation to necessarily concentrate on the displayed bet. The results cannot therefore say whether any effect is due to the task itself, i.e., Evaluation versus Choice, or to the effect of the displayed information. Because the display is essentially constant across tasks in our experiment, we can more cleanly identify a pure effect of task mode on attentional allocation. Relatedly, we believe that our design induces more control on the comparisons that people make in each task, a point relevant for some theories. For instance, Bordalo et al. (2012) claim that reversals occur (in the standard designs) because subjects consider different reference bets across tasks: In Choice they compare the P- and $-bet, whereas in Evaluation they compare the relevant, isolated bet with a sure lottery giving a zero payoff. This confound is arguably irrelevant in our design. In this line, we wanted to reduce as much as possible any potential effect of memory on the reversal rate (via omitted referents), which explains why subjects have as much feedback as possible, including even the possibility to consult decisions in a prior task.

One of the key elements of our empirical design was the use of a Tobii eye-tracking system, overlaid on the computer screens. This system uses reflected light from the participants' eyes to determine where on the screen they are looking at any moment. In this way it measures the attention that participants are paying to any given object on the screen directly. The system records fixations, defined as gaze events longer than 16 milliseconds. We then define "areas of interest" as zones on the screen, generating binary variables indicating whether a particular fixation represented attention to something within the area of interest. Attention can therefore be measured either by the number of fixations on a particular object on the screen, or by the total amount of time the participant spent looking at the object. In our experiment, the two measures yielded very similar results, so we will generally focus on the number of fixations. Transitions (Devetag et al. 2016) occur when a participant shifts attention from one object to another.[15] To calculate transitions, we first drop all fixations that were not made on some object of interest, and then sort the remainder chronologically. We then classify the transitions according to whether they are *within-bet*, involving a switch between the probabilities and prizes of a given bet, or *between-bet*, involving a switch between some attribute (probabilities or prizes) of one bet and another.[16]

We used three eye-tracking stations at Tilburg University, with a total of 39 participants. The software used to program the tasks was z-Tree (Fischbacher, 2007). Participants were not students of the experimenters. After being seated, each participant received written instructions that described the decision problem (see Appendix I). Subjects could read the instructions at their own pace, and ask questions as they needed. Understanding of the rules and particularly of the payment procedure was checked with a control questionnaire that all subjects had to answer correctly before they could start making choices (see Appendix I). All hypothetical examples in this questionnaire were constructed by the participant, to prevent any potential suggestion of play. Once the questionnaire was completed, the eye-tracking system was calibrated, and the choice and evaluation tasks began. After all tasks had been performed by all subjects, one bet was chosen for payment for each subject and the subjects informed in this respect. Afterwards, they answered a brief questionnaire where we gathered personal information on socio-demographic characteristics, as well as the cognitive reflection test or CRT (Frederick, 2005), and a risk aversion index.[17] This ended the experiment.

---

[15] Transitions are therefore a broader class of observation than *saccades*, which represent the movement between two sequential fixation points.

[16] Little was gained empirically by further distinguishing, for example, between-probability (one probability in the pair to the other) from between-prize and between-cross-attribute.

[17] Our risk-aversion index had two bets, with prizes L_S and H_S in the safer bet, and L_R and H_R in the riskier one, such that $L_R < L_S < H_S < H_R$. It asked participants to choose a single probability for the higher prize that left them indifferent between the bets. It was not incentivized.

Subjects were paid by bank transfer the following day. Each session[18] lasted approximately 45 minutes, and on average subjects earned 16.17 Euros, including a show-up fee of 6 Euros.

4. Data and results

Our sample included 39 individuals and a total of 38,029 fixations. We exclude some of these from the analysis, with two different exclusion criteria. First, although the system records fixations of 16ms or longer, consistent with the literature (e.g. Manor and Gordon 2003), we only keep fixations of 80ms or longer in the data, arguing that the attention represented by fixations shorter than that is negligible. Second, noise in the calibration resulted in a very low number of recorded fixations for some participants. In particular, four individuals had 83 or fewer total recorded fixations. The next smallest number of fixations by a single subject was 311, and the average number was just over 1029. We assume that these low outliers were due not to participants actually playing without looking, but rather to a technical problem with the recording. The missing fixation data would add noise to the interpretation of their results, so we removed them from the sample, leaving 35 individuals. Not all these 35 individuals had data for all bets, tasks and/or attributes, but excluding more participants made little qualitative difference to the results, so for the remainder we focus on these 35 individuals.[19]

4.1 Revealed preferences & attention: Summary statistics

Table II below shows, for each bet pair, the values of the P- and $-bets, as well as the proportion of participants choosing the P-bet in the choice task, and giving a higher evaluation to the $-bet. Somewhat unusually for this kind of study, we see that in several pairs the P-bet was not preferred on average in the Choice task (45% chose the P-Bet overall), but the $-bet was evaluated higher in two thirds of bets (66%). It must be noted though that indifference in Choice is often not allowed in PR experiments –e.g., Tversky et al. (1990), Kim et al. (2012), Alós-Ferrer et al. (2016). This has arguably minor implications when the evaluations are cardinal, e.g., bids or minimum selling prices, because a higher evaluation of one bet over another indicates in theory strict preference. In our OPM case, in contrast, the evaluation of a bet is not to be interpreted as its certainty equivalent, and if bet 1 is assigned a larger number than bet 2, we can conclude only that 1 is weakly preferred to bet 2, i.e., not worse, but not necessarily better. The fact that we allow indifference in Choice was therefore important to discriminate which decisions were (weakly) consistent but at the same may partially

[18] In order to run the choice experiment within the eye-tracking software, it was necessary to run a separate instance of zTree for each subject, rather than one central server and a number of clients, as is more common. Therefore, each subject technically represents a different session.

[19] In particular, one large outlier recorded 1878 fixations on the attributes of the twelve bets. The next highest was 822. Excluding this individual does not change any results either in substantial numerical or significance terms.

explain the relatively low share of measured PRs in our study, compared with other OPM designs where indifference is not allowed.

	(a) Bets				(b) Behavior						(c) Outcomes		
Pair	P-bet		$-bet		Choice		Evaluation		Valuations		Reversal		
	Prize	Prob.	Prize	Prob.	Chose P-bet	Ind.	Rated $-bet	Ind.	P-bet	$-bet	Strong	Non-standard	Weakly Consistent
1	8	97	24	33	0.60	0.03	0.60	0.14	14.78	17.35	0.37	0.06	0.57
2	10	73	30	27	0.57	0.09	0.74	0.11	13.33	17.85	0.43	0.00	0.57
3	9	87	80	10	0.40	0.09	0.63	0.11	15.90	28.05	0.23	0.09	0.69
4	7	63	40	17	0.34	0.11	0.74	0.03	11.74	17.93	0.23	0.03	0.74
5	5	90	35	13	0.54	0.11	0.69	0.09	12.37	16.33	0.31	0.03	0.66
6	3	87	13	20	0.26	0.31	0.57	0.06	8.34	8.08	0.11	0.11	0.77
Mean	7	82.83	37	20	0.45	0.12	0.66	0.09	12.74	17.60	0.28	0.05	0.67

Note: N = 35. For each bet pair, panel (a) shows the prize for each bet and the probability of winning it. Panel (b) shows for each pair of bets the proportion of subjects choosing the P-bet and rating the $-bet higher in the choice and evaluation tasks, respectively; it also shows the proportion who indicated indifference in Choice or evaluated equally both bets in Evaluation, for information's sake. Panel (c) shows the percentage of participants who committed a strong standard reversal or non-standard reversal, as well as the remainder, who were, therefore, weakly consistent. This group, for instance, includes all those who indicated indifference in at least one task.

Table II: Summary statistics on preferences

In the attention data, the average recording length of the eye-tracking software was just under 25 minutes, with a variation between 14 and 56 minutes. Of this, an average of 7.8 minutes (and a minimum and maximum of 4.1 and 25.7 minutes, respectively) were spent on Choice and Evaluation. The eye-tracking software recorded a total of 36,020 fixations of 80 ms or greater across our 35 retained participants during this time. Several "areas of interest" were defined, generating binary variables that indicate whether the fixation was on a prize, probability, answer, elsewhere in a particular bet, elsewhere in the pair of bets, or in some non-bet area of the screen. For illustration, Figure 3 below plots all fixation points in each task, together with the prize, probability and answer zones of each of the twelve bets (compare with Figure 2 above). At the bottom of the figure, the fixation points cover the instructions and buttons to move away from the screen or between tasks. A first thing to notice is that the majority of fixations cluster around the six pairs of bets that were visible on the screen. This hints that the degree of noise in the recording process was arguably low. In this vein, there are relatively more fixations on the answer zone in Evaluation, which is natural because the evaluations were entered there in that task. In turn, Choice presents relatively more fixations between the two bets of each pair, again natural because the arrow indicating the choice appeared in that place. Comparing the Choice and Evaluation panes, finally, we observe noticeably more fixations in Evaluation than Choice. We will return to this issue throughout the paper.

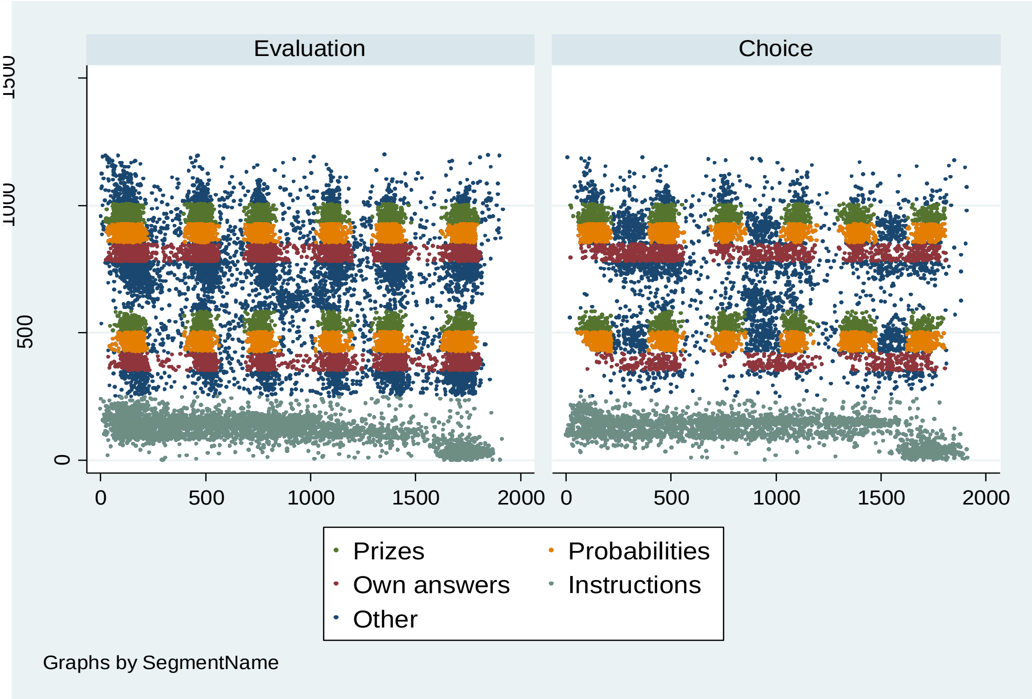


Figure 3: Fixation points on the Choice and Evaluation task interfaces

For some additional detail, Table III shows the numerical breakdown of the recorded fixations across both tasks. We see that the largest targeted component is the probability, with the other elements about equally split. Nearly a quarter of fixations were not on any particular area of interest. Unless otherwise noted, in the remainder of the paper we focus on the 14,806 fixations that occur on attributes of the bets, investigating attention allocation and the three potential accounts for reversals more rigorously.

Fixation on	Freq.	Percent	Average length (ms)
Probabilities	8,736	24.25	290.16
Prizes	6,070	16.85	271.78
Answer	6,548	18.18	261.82
Instructions	5,941	16.49	243.24
Elsewhere	8,725	24.22	249.47
Total	36,020	100.00	265.92

Table III: Overall fixations by area of interest

4.2 Depth of reflection

4.2.1 Task, Behavior and Reversals

We begin our detailed analysis with an investigation of what we called the extensive margin of attention, or depth of reflection, i.e., the whole mental effort spent on a task. Recall from the introduction that our analytical framework has three main questions, i.e., Task, Behavior, and Reversals; one reflecting the effect of the task form on attention, the second the relationship between attention and behavior in each task, and the third tracing the effect of attention on reversals. We begin with some evocative figures on the first question. Table IV below gives the number of recorded fixations over the entire sample in each Task.

Variable	Type of task		Total
	Evaluation	Choice	
Total number of fixations	9,145	5,661	14,806
Average fixation length (ms)	297.5	278.8	290.4
Fixations on probabilities	5,308	3,428	8,736
Fixations on prizes	3,837	2,233	6,070
Fixations on $-Bet	4,875	2,892	7,767
Fixations on P-Bet	4,270	2,769	7,039

Table IV: Some statistics on fixations conditional on type of task

There are many similarities but also some differences. The average fixation length was similar in the two tasks, differing by less than twenty milliseconds. The share of fixations on probabilities on the one hand (about 60%), and on the P-bets on the other, (just under 50%) was also similar. Most relevantly for Depth of Reflection, however, subjects made many more fixations in Evaluation than in Choice, which probably reflected the fact that it was more difficult or unfamiliar. As a result, including fixations outside of the bets, participants spent more than twice as much time looking at the evaluation screens as the choice screens, consistent with Figure 3 above. Per bet, and focusing on the probabilities and prizes only, the average number of fixations was 24.4 in Evaluation, and 15.2 in Choice, which represents about three seconds' less fixation time in the latter. This difference of just over 9 fixations per bet happens to be significant ($p = 0.037$), as shown by a linear regression of total fixations at the bet level on task type, as well as subject characteristics and dummies controlling for the specific pair, with standard errors clustered at the subject level (i.e., 35 independent clusters).

Result 1 (depth of reflection & task): The task affects the total number of fixations on each pair of bets, which is significantly higher in Evaluation.

While on average, subjects pay more attention to the bets in Evaluation than Choice, this *difference* is particularly marked among those subjects who exhibit more coherent behavior. This seems to be the most remarkable finding regarding Depth of Reflection, as applied to the **Behavior** question. For some preliminary evidence in this regard, Table V shows the average fixations per bet pair in each Task, broken down by behavior in each.

		Gave higher rating to					
		P-bet		$-bet		Total	
		Eval	Choice	Eval	Choice	Eval	Choice
Chose	P-bet	39.06*	30.03*	35.15S	29.14S	36.63	29.47
	$-bet	42.06N	41.06N	65.09*	25.89*	58.08	30.26
	Total	40.54	35.30	52.38	27.28	48.38	29.90
Notes: * Consistent behavior; S Standard reversal; N Non-standard reversal							

Table V: Average fixations per pair of bets in each Task, by behavioral pattern

The key point from Table V involves the horizontal differences between the numbers in each underlined cell. Specifically, the differences between fixations in Evaluation and Choice are greater among participants who were consistent (identified with stars) than among those who were not. Among the "non-reversers" who chose the P-bet (upper left cell), for instance, we observe in average 30% more fixations in Evaluation, whereas the increase amounts to 20% among the corresponding "reversers" (upper right cell). The difference is more dramatic though when comparing those reverses with the non-reversers choosing the $-bet (lower right cell). Overall, these differences are in line with our conjecture that maintaining consistency requires more attention in Evaluation than in Choice; those who do not spend the extra effort are those who answer inconsistently. Moreover, while the data in Table V are pooled over all individuals who behaved in the same way, the phenomenon remains within-subjects. Comparing the number of fixations of each individual on each pair across Tasks with a Wilcoxon Ranked Sign test, we find that those who behaved consistently across tasks fixated significantly more in Evaluation than in Choice, regardless of whether they preferred the P-bet ($p = 0.045$, $N = 35$) or the $-bet ($p = 0.000$; $N = 79$). On the other hand, those who behaved inconsistently had statistically indistinguishable fixations across tasks, whether they committed a standard ($p = 0.541$; $N = 59$) or non-standard ($p = 0.667$; $N = 32$) reversal. To construct Figure 4 we calculate the total fixations for each participant on each pair for each task separately, then subtract the numbers, finding the difference in fixations across tasks on each pair for each participant. The figure shows the average difference within-subjects in number of fixations across the two Tasks, according to (weak) consistency of choice or (pooled) reversal.

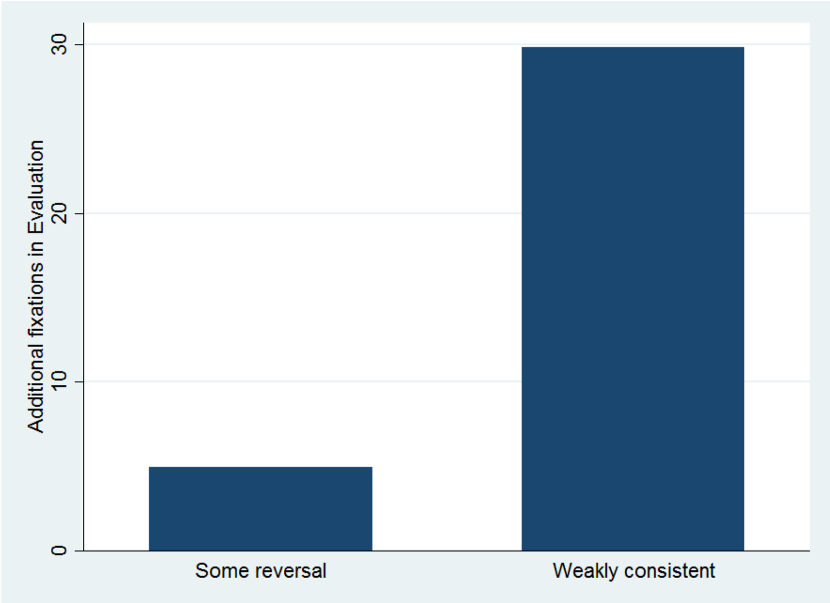


Figure 4: Within-subject differences in attention from Evaluation to Choice, conditional on weakly consistent behavior

The difference is dramatic. On average, those who behave consistently look nearly 30 extra times at the Evaluation Task compared to Choice, more than six-fold the difference of just under 5 found among those who exhibit some kind of reversal. These numbers are significantly different from each other (Mann-Whitney $p = 0.0003$), and only the former is significantly different from zero ($t = 5.101$ for consistent; $p = 0.0000$; $t = 1.105$ for inconsistent, $p = 0.2720$).

Result 2 (depth of reflection & behavior): Subjects who exert relatively little additional effort to complete the Evaluation task are more likely to give evaluations contradicting the choice between bets.

An interesting point is that, while consistent behavior seems to require a significant difference in attention between tasks, it does not depend on the level of attention in Choice. This is in line with the Depth of Reflection story. That is, those who do not exert sufficiently more attention are more likely to make a "mistake" in the more difficult Evaluation task. As a result, the form of the reversals will be driven by the most common pattern of behavior in the task that is easiest to complete, i.e., Choice. For further detail, Table VI below reports several regressions at the individual subject level, measuring the relationship between the total number of reversals committed across the six pairs and the overall number of fixations. Hence the outcome is a number between 0 and 6, and the main explanatory variable is an individual-level characteristic measured for the session. We note that results are qualitatively similar if attention time (in seconds) is used instead of the number of fixations.

	(1)	(2)	(3)
VARIABLES	Number of PRs	Number of PRs	Number of PRs
	Overall	Evaluation	Choice
Number of Fixations	-0.00149	-0.00216	0.000791
	(0.000793)	(0.000788)	(0.00445)
CRT score	1.086	1.046	1.072
	(0.534)	(0.551)	(0.592)
Risk aversion	1.422	1.623	1.311
	(0.995)	(0.993)	(1.057)
Male	-0.0695	-0.271	0.0937
	(0.557)	(0.587)	(0.630)
Age	-0.270	-0.233	-0.292
	(0.107)	(0.113)	(0.115)
Constant	7.215	6.702	7.237
	(2.341)	(2.358)	(2.670)
Observations	70	35	35
R-squared	0.313	0.355	0.293

Note: Robust standard errors in parentheses.

Table VI: effect of overall attention levels on number of reversals

Table VI shows, first, that the overall level of attention, measured as the number of fixations recorded, significantly predicts a lower number of reversals at the individual level. Column 1 estimates this over both tasks, using the attention to a pair of bets in one task as an observation (N = 70 in 35 independent clusters). We then split the sample, running separate regressions for the data from the Evaluation (col. 2) and Choice (col. 3) Tasks. We see that the number of reversals committed correlates with our measures of attention only during the Evaluation task. Thus attention appears to exhibit the kind of asymmetric effect across tasks that is necessary for it to be involved in generating systematically skewed patterns of inconsistency.[20]

Result 3: At the participant level, more attention during Evaluation reduces the number of reversals committed; differences of attention in Choice do not affect reversal rates.

4.2.2 Relationship to Indifference

[20] Among the controls, we note with surprise that a higher CRT score marginally predicts *more* reversals. We have no particularly convincing story for why this is the case. We suspect it may have to do with previous exposure of some participants to those questions. In this vein, we note that those who got perfect scores on the CRT completed the page nearly one full minute earlier than those who missed at least one question (average 95s versus 155s, Mann-Whitney p = 0.0001). This is suggestive that the differences in CRT score are more linked to mathematical skill and experience, both of which should reduce the time taken, than to the penchant for "cognitive reflection" the test is designed to assess, which should if anything increase it.

On its surface, Result 3 seems to conflict with the finding by Alós-Ferrer et al. (2016) that PRs are associated to *longer* response times (RTs) in the Choice task, and hence deserves a slight digression. We make several remarks in this respect. First, the rationale given by Alós-Ferrer et al. (2016) for the correlation between PRs and RTs is twofold: (I) people close to indifference in Choice face a relatively "hard" decision and hence take longer to choose, (II) at the same time, an imprecision interval argument predicts that these subjects are most likely to exhibit PRs. Note that this idea (I) that some choices are "hard" is not actually at odds with Depth of Reflection. Our results, therefore, do not necessarily contradict Alós-Ferrer et al. (2016); it is conceivable that the two mechanisms operate in parallel. In terms of our vision, the "sufficient'" attention required to avoid reversals would simply be higher when the choice was "harder". For any level of difficulty, it will be those who focus longer in Evaluation who do not make reversals, even if overall, the level of attention is higher for those closer to indifference.

If both factors are operative – hard choices take more work and generate reversals on the one hand, but more effort in Evaluation reduces reversals on the other – the question remains why the former effect dominated in Alós-Ferrer et al. (2016), while the latter is in evidence in our data. We believe the difference can be traced to differences in design. In our design, for instance, subjects can express indifference in Choice, and by construction those who do so cannot commit a reversal. Thus mechanically, some who would have faced the hard choices that Alós-Ferrer et al. (2016) find lead to reversals, can "escape" by expressing indifference in our design, which will obviously reduce the effect they study. The structure of our experiment is also substantially different. In Alós-Ferrer et al. (2016), participants evaluate 40 different bets, and so there is substantial variation possible in the objective difficulty of the tasks. This difference *across bets* generates the effect those authors investigate. In our study, by contrast, there are only six pairs of bets available, and so it seems reasonable to suppose that the variation observed in our study will be more likely *across individuals* than across bets. Unless differences in preferences across individuals greatly outweigh differences in attention, this will also tend to bring out the latter effect in our results. Both of these design factors will amplify the "hard choice effect" in Alós-Ferrer et al. (2016), and highlight depth of reflection in our study.

Nevertheless, if both factors are indeed operative, we should be able to find some indication of the hard choice effect in our data. And indeed we do. Table VII reports the results of a series of regressions on attention, measured both by number of fixations and by time (since we are comparing explicitly with the reaction time in Alós-Ferrer et al. (2016)), as determined by indifference. We see in regressions (1) and (2) that overall, across both tasks, there seems to be little effect. However, this hides a strong effect in the direction predicted by Alós-Ferrer et al. (2016) in the Choice (3) and (4), behind an effectively zero effect in Evaluation (5) and (6).

VARIABLES	Overall		Choice		Evaluation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Fixations	Time	Fixations	Time	Fixations	Time	Fixations	Time
Indifferent	6.946	3,027	14.58	5,707	-3.157	-356.5		
	(5.420)	(1,781)	(4.605)	(1,558)	(10.83)	(3,300)		
Pshare							238.3	66,433
							(83.33)	(24,998)
PShare²							-279.5	-80,535
							(94.19)	(28,727)
TaskType	-16.27	-5,416						
	(6.873)	(2,035)						
Constant	-8.722	-67.57	21.23	6,271	-53.68	-11,390	-88.44	-20,858
	(31.28)	(10,077)	(19.55)	(5,309)	(61.96)	(19,634)	(62.94)	(19,857)
Controls	ALL	ALL	ALL	ALL	ALL	ALL	ALL	ALL
Observations	413	413	204	204	209	209	209	209
R-squared	0.136	0.126	0.093	0.146	0.154	0.119	0.219	0.184

Notes: Robust standard errors in parentheses, clustered at the individual level. Controls include gender, age, CRT score, risk aversion index, and dummies for each bet. Indifference is measured in the Choice task by clicking on the appropriate button, and in Evaluation by entering equal values for each bet. PShare is defined as EvalP/(EvalP + Eval$), where EvalB is the monetary evaluation given for bet B, B = P, $.

Table VII: Indifference increases attention

On the other hand, the models in regression (5) and (6) may be an overly stringent test. The “hard choices” argument is built on the idea of imprecision; the “indifference” indicated in regressions (5) and (6) is, on the contrary, very precise, defined by identical evaluations of each bet. To get a better measure, therefore, we calculate the PShare variable, defined as the evaluation of the P-bet, divided by the sum of the evaluations. The closer this is to 0.5, the closer the evaluations are to each other, and presumably the closer the participant is to indifference. In line with the hard choice effect, one could expect attention to peak when this “degree of indifference” was around 0.5. We therefore run two more regressions, using PShare and its square as regressors. The prediction would be that the peak of the resulting parabola should be around 0.5. Indeed, the parabola is shaped in the right way (positive direct and negative squared coefficients), and focusing on regression (7) the turning point is at –(–279.5/2*283.3) = 0.586. The 95% confidence interval for this point goes from 0.482 to 0.691. Hence under the assumption that relative evaluations accurately reflect relative value (with error) we cannot reject the idea that attention peaks at true indifference. It would therefore appear that despite initial appearances to the contrary, we in fact replicate the effect of indifference on attention times from Alós-Ferrer et al. (2016), although not the correlation with the PRs, perhaps for the reasons given above.

4.3 Salience

4.3.1 P-bet versus $-bet

We now turn to the second element of our analysis, the salience hypothesis, or intensive margin of attention, i.e. how attention is allocated across the attributes of the bets and the bets themselves. There are two different dimensions of allocation that present themselves: allocation to the P- versus the $-bet, and allocation to Prizes of the bets versus Probabilities. We will investigate each in turn, beginning with the allocation across bets.[21]

The observation level is a bet pair, and the main outcome variable is the fraction of fixations on the P-bet in the pair –results are similar if time in seconds is used instead. Some preliminary evidence is reported in Table VIII. Panels (a) through (d) show the comparisons across variables representing our Task, Behavior and Reversals questions. Columns (1)-(6) show the comparisons for their application. Column (6), for instance, corresponds to the Depth of Reflection hypothesis, measuring differences in total fixations. Reading down Column (1) addresses the same questions for the Salience hypothesis across bets, Columns (2)-(4) for Salience across the attributes of the bets; and (5) for the type of transition, an issue that we check in more detail in 4.4. We report Mann-Whitney test p-values for indication of the magnitudes of the comparisons.[22]

	(1) P-bet vs $-bet	(2) Prob vs Prizes overall	(3) Prob vs Prizes $-bet only	(4) Prob vs Prizes P-bet only	(5) Across transitions	(6) Total fix
(a) Task type						
Evaluation	0.461	0.584	0.572	0.593	0.137	48.376
Choice	0.491	0.566	0.563	0.566	0.326	29.898
M-W p-value	0.038	0.481	0.947	0.271	0.000	0.000
(b) Choice						
Chose $-bet	0.454	0.592	0.581	0.593	0.325	30.261
Chose P-bet	0.536	0.534	0.542	0.534	0.326	29.468
M-W p-value	0.000	0.057	0.175	0.079	0.816	0.958
(c) Evaluation						
Rated $ higher	0.457	0.614	0.616	0.611	0.123	40.535
Rated P-higher	0.462	0.568	0.550	0.584	0.144	52.381
MW p	0.884	0.083	0.061	0.366	0.562	0.926
(d) Reversal						
No reversal	0.461	0.599	0.587	0.608	0.207	82.748

[21] Little additional insight arose from the interaction of the two, and the analysis becomes quite cumbersome, so we do not report it here.

[22] However, it is important to note that the validity of the tests is severely compromised by the non-independence of the decisions that generate them. Our analysis is largely at the individual-pair or individual-pair-task level, so we have at most 35 independent observations. Our central analyses below will be based on regressions with corrected standard errors at the individual level and controls for several other variables.

Reversal	0.480	0.537	0.532	0.536	0.230	64.288
MW p	0.447	0.062	0.232	0.033	0.077	0.010

Table VIII: Average share of fixations on P-bet and bet items, conditional on several aspects

Focusing for the moment on column 1, we see several points. First, in (a), task affects relative attention to the bets: people pay more attention to the P-bet in Choice, by a modest but significant increase of three percentage points. Second, in (b), attention to the P-bet relative to the $-bet significantly predicts its choice in the Choice task. Those who chose the P-bet looked at it nearly 20% more in percentage terms than those who chose the $-bet, and the modal object of fixation tracks choice; on average, people choose the bet they look at most, consistent with drift diffusion models, as we explained above. On the other hand, we see in panel (c) that this relationship doesn't hold for Evaluation. Those who rated the P-bet higher than the $-bet looked at it only slightly more – the difference is an insignificant half a percentage point. And finally, panel (d) indicates that bare reversals do not correlate with relative attention to the P versus $-bet at the pair level, *pooling over both tasks*. For a deeper analysis, Table IX below shows the results of several regressions, which repeat the analysis above, controlling for individual characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	**Share on P-bet**	**Chose P over $**	**Rated $ > P**	**Reversal overall**	**Reversal Choice**	**Reversal Evaluation**
Choice task	0.0310					
	(0.0144)					
% of f. on P-bet		5.756	-0.301	0.0197	5.858	-3.868
		(1.847)	(0.821)	(1.644)	(2.196)	(2.278)
% of f. on Probs				-1.227	-1.538	-1.459
				(1.650)	(1.823)	(2.501)
% of f. in P-bet Prob				1.389	0.590	3.351
				(2.484)	(3.011)	(3.799)
Constant	0.541	-3.965	0.179	-2.899	-6.454	-1.199
	(0.124)	(1.927)	(2.016)	(1.186)	(1.770)	(1.539)
Controls	ALL	ALL	ALL	ALL	ALL	ALL
Observations	415	205	210	415	205	210
R-squared	0.036					
Note: Robust standard errors in parentheses.						

Table IX: Regression results – salience across bets

Regression (1) in Table IX corresponds to panel (a) of Table VIII ; we see that overall, Choice led to a higher share of fixations on the P-bet relative to the $-bet than did Evaluation. Regressions (2) and (3) repeat the analysis from panels (b) and (c), confirming that the imbalance that exists also correlates more with behavior in Choice than in Evaluation: people choose the bet they look at more, but do not rate it systematically higher or lower. Finally, regressions (4)-(6) unpack the relationship to reversals. As we did above, we begin in regression (4) with the entire sample (12 observations per individual; six pairs in each of two tasks). We then split the sample, re-running the analysis with the data on fixations in Choice (regression 5) and Evaluation (regression 6). Having a large share of fixations in the P-bet *in Choice* (5) was highly predictive of inconsistent behavior, and a percent of fixation *in Evaluation* (6) was in fact marginally predictive of consistent behavior; overall (4), the effect washes out, consistent with the analysis in panel (d) in Table VIII.

Taken together, the story these regressions tell is consistent with reversals being at least partially driven by differences in the allocation of attention across tasks. If in Choice people focus more on the P-bet relative to the $-bet than in Evaluation, and moreover that focus tends to influence behavior, then Choice, relative to Evaluation, seems to encourage preferring the P-bet over the $-bet through some attentional mediation effect. The fact that this pattern operates only in Choice then opens the possibility that these attention-mediated preferences generate standard reversals. And indeed, we see that reversals are more common among subjects whose focus is stronger on the P-bet during Choice.

Result 4: Preferential attention to the P-bet during Choice encourages its choice and partially drives standard reversals.

4.3.2 Prize versus probability

The parallel analysis applied to the share of fixations on probabilities versus prizes turns up surprisingly little. Columns (2)-(4) of Table VIII above investigate the correlation of a relative focus on probabilities over prizes overall, on the $-bet and on the P-bet, respectively, (a) across tasks; (b) on behavior in Choice; (c) on behavior in Evaluation; and (d) in relation to reversals. The asymmetric patterns observed with respect to attention across bets, however, fail to appear. There is some marginal evidence that those who focus more on probabilities have a tendency to prefer the $-bet over the P-bet, and one result that those who focus on probabilities in the P-bet tend not to commit reversals. However, these points are neither consistent nor particularly coherent. Our intuitions about attention to probabilities suggest that it should caution people away from the $-bet. We suspect that this is another reflection in the data of the Depth of Reflection results found above, i.e., people who pay less attention overall also pay less attention particularly to probabilities. Observe also that these results are at odds with psychological theories assuming that the weights on the decision attributes, i.e., probability and prize, depend on the task, via the attention pattern the task motivates (recall the discussion in Section 2).

The results are again largely confirmed in the regressions reported in Table X below. The analysis for the Task (i.e., Col. (1)) and Behavior (Cols. (2)-(3)) questions here is at the bet level rather than the pair level, because the phenomenon involves allocation within a particular bet. For the Reversal questions (Cols. (4)-(6)), we retain the analysis at the pair level because the outcome variable is only defined at that level. For columns (5) and (6), however, we once again break the sample into observations recorded during the Choice and Evaluation tasks, respectively. We include interactions between the percent of fixations on probabilities and the type of bet, P or \$. The interpretation of these interactions is the difference in the marginal effect of increasing attention to probabilities across the bet type on each outcome. Notice that because the dummy in the interaction indicated the P-bet, in the presence of this interaction, the interpretation of the direct effect of attention to probabilities is the effect of that attention on the \$-bet. We expected these interactions to be significant, since focus on the probabilities in the P-bet entails focus on its better characteristic, while focus on the probabilities of the \$-bet indicates attention to its worse characteristic.[23]

VARIABLES	(1) % of f. on probabilities	(2) Chose P-bet	(3) Higher eval. for \$-bet	(4) Reversal Overall	(5) Reversal Choice	(6) Reversal Evaluation
Choice task	-0.0178					
	(0.0212)					
% of f. on probabilities (\$-bet)		-0.321	-1.664	-0.124	-0.198	-0.0922
		(0.686)	(0.984)	(0.210)	(0.195)	(0.341)
P-bet – direct effect	0.0318	0.00436	-0.199	0.00925	0.225	-0.303
	(0.0130)	(0.363)	(0.432)	(0.222)	(0.260)	(0.406)
Interaction	-0.0199	-0.0202	0.350	-0.0170	-0.0894	0.266
	(0.0162)	(0.590)	(0.676)	(0.316)	(0.360)	(0.601)
Constant	0.713	5.398	3.679	1.395	1.338	1.426
	(0.189)	(2.244)	(2.263)	(0.422)	(0.420)	(0.478)
Observations	813	399	414	413	204	209
R-squared	0.063			0.183	0.197	0.183
Note: Robust standard errors in parentheses.						

Table X: salience across attributes.

The first thing to notice is that the share of fixations on probabilities versus prizes is never significant in predicting anything at conventional levels. There seems to be a slight tendency for increased focus on the probabilities to reduce the chance that the \$-bet is evaluated higher than the P-

[23] Additional regressions, not reported, leave out the interaction terms. No changes in significance are found.

bet, which goes in line with our hypotheses.[24] However, the effect is not significant at conventional levels, and given the number of hypotheses tested we are wary of making strong claims about the occasional marginal result. For the most part, increased proportional attention to probabilities correlates neither with the task or subsequent behavior. The interaction coefficients, in particular, are not significant, which implies that even controlling for the "connotation" (positive or negative) of the probability in the particular bet, differences in attention across attributes does not seem to predict behavior. Given this, it is perhaps not that surprising that it also does not correlate with reversals. These results are robust to several different specifications and controls, particularly concerning control for the characteristics of the bets in question.[25]

Result 5: Attention to prizes versus probabilities does not seem to be a driver of reversals in this experiment.

4.4 Transitions

We have found in 4.3 that differences in the relative share of attention across bets drives behavior in the Choice task, and that substantial attention is required for subjects to uncover preferences coherent with this in Evaluation (4.2). This suggests that deliberation in Choice may involve different patterns of reasoning than in Evaluation. We further corroborate this by looking at transitions. We define a transition in the data when two consecutive fixations by a particular subject are on elements of the *same bet pair* in the same task.[26] By comparison with fixations, which we take to measure "what the subject is thinking about", transitions show how subjects may be combining these elements together, along the lines of "what are they thinking about it". It seems plausible that attention to two different data in sequence might relate to paying attention to both of them, and therefore to further inferences drawn from them. It is therefore useful to investigate whether there are differences across task type in how the basic data of a pair of bets are being combined. The main patterns of transitions are shown in Table XI.

In Table XI we make a distinction between "within" transitions, representing a shift between different parts of the same bet, and "across" transitions, which occur when the subject moves from

[24] Technically, the reported result concerns the share of fixations on probabilities for the $-bet only. However, because the interaction is not significant, we see the same effect holds for fixations on the P-bet; additional regressions without the interaction – not reported here – show the same level of significance on this pooled measure.

[25] We found one specification, aggregated at the pair level instead of the bet level, and controlling for the exact values of the prizes and probabilities instead of dummies for the bets, in which the significance of the result for the evaluation task rises to conventional levels. But given its fragility to other specifications, we do not include it in the main analysis.

[26] Note that these were defined after removing all fixations not on the attributes (probabilities and prizes) of the bets. Therefore, it is possible that the subject looked elsewhere on the screen between the fixations that we record as a transition. However, if a fixation was recorded on a bet in a different pair, that broke the transition. For instance, looking at the probability of the P-bet in Pair 2, then the instructions, then the prize for the $-bet in Pair 2 would be recorded as a transition across the bets in Pair 2, but looking at the probability of the P-bet in Pair 2, then the prize for the $-bet in Pair 3, and then the prize for the $-bet in Pair 2 would not be recorded as any transitions at all.

one bet in a pair to the other. Note that the explicit tasks of Choice and Evaluation might be expected to encourage transitions across bets in a pair in the former relative to the latter. However, the instruction that the subjects would play either the higher-evaluated of the bets in a pair or the chosen one implies that the payoff-relevant decision involves a comparison of the bets in both tasks. Further, our design, which leaves all bets on the screen during both tasks, enables subjects to make these comparisons explicitly, equally well in both tasks. Therefore, any differences that we find in transition patterns represent a cleaner test of the effect of task mode than found in earlier literature, which traditionally only shows one bet at a time during Evaluation.

	Choice		Evaluation	
Transition Type	**Number**	**Percent**	**Number**	**Percent**
Within-bet	**3191**	**67.88%**	**6491**	**86.67%**
Within $-bet	663	14.10%	1654	22.09%
Within P-bet	713	15.17%	1372	18.32%
Focus	1815	38.61%	3465	46.27%
Across-bet	**1510**	**32.12%**	**998**	**13.33%**
Compare Probabilities	705	15.00%	359	4.79%
Compare Prizes	403	8.57%	288	3.85%
Best (P-prob - $-prize)	196	4.17%	184	2.46%
Worst (P-prize - $-prob)	206	4.38%	167	2.23%
Grand Total	**4701**	**100.00%**	**7489**	**100.00%**

Table XI: Switches between elements of a bet pair.

Table XI shows that indeed, within-switches are much more common in Evaluation than in Choice, both in absolute and relative terms. In fact, nearly 87% of all transitions in Evaluation were within-bet transitions, if we include "transitions" where two consecutive fixations were on the same datum (called "focus"). On the other hand, about 68% of transitions in Choice were within the same bet. Because there were also more fixations in Evaluation, this higher percentage corresponds to an even greater difference in absolute number, with about twice the number of within-bet transitions recorded in Evaluation as in Choice. Looking at across-switches, we see the opposite relationship, with higher absolute numbers in all categories in Choice than in Evaluation. There are twice as many transitions that compare probabilities in Choice as in Evaluation, and half again as many comparing prizes; these numbers are again amplified if we look at the percentages, because of the lower number of total fixations in Choice. Figure 5 breaks these patterns down further, showing the proportion of transition in each of the sixteen possible directions for each task.

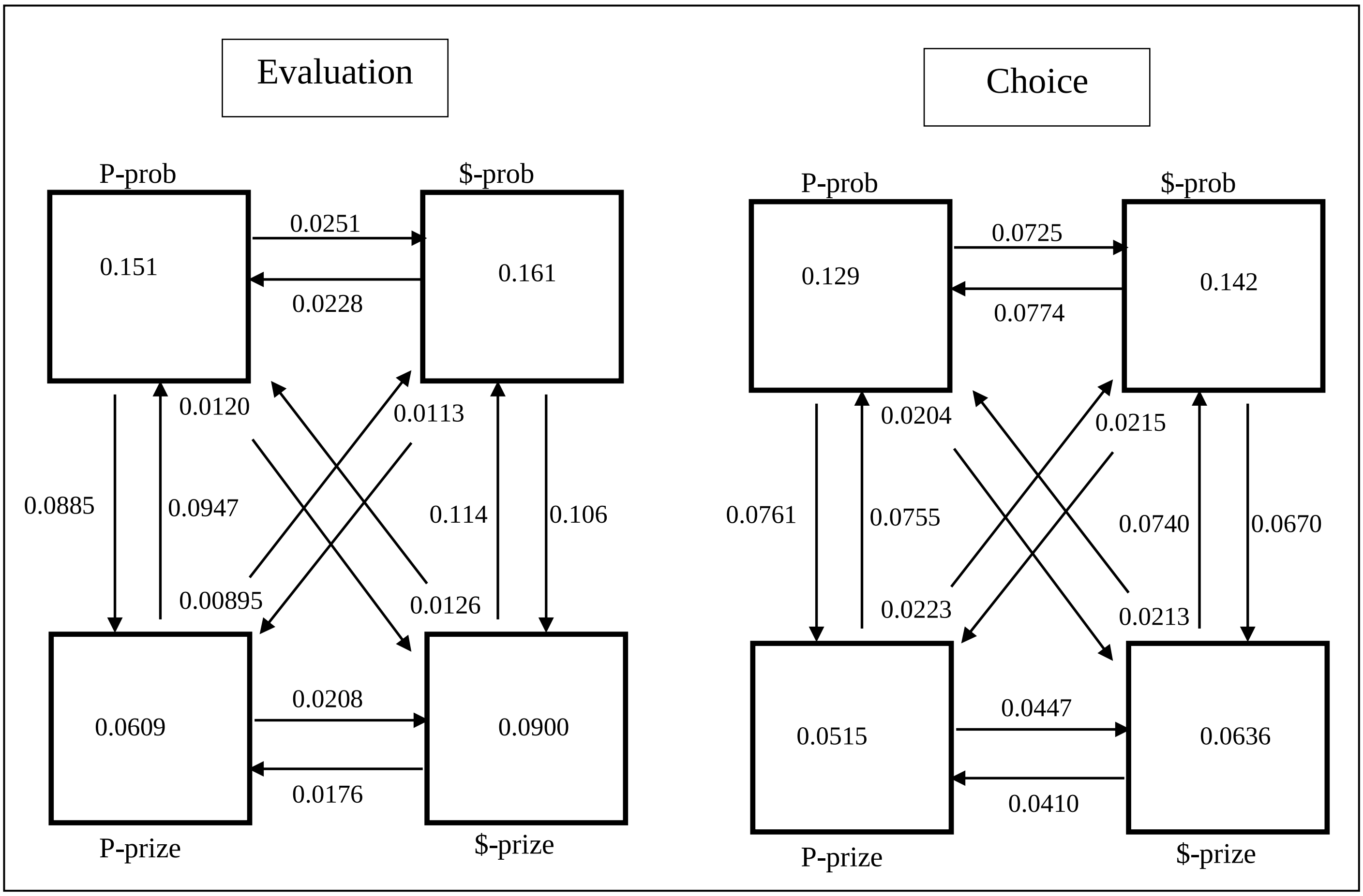


Figure 5: proportions of within-bet and across-bet transitions in each task

The figure shows a schematic for a pair of bets in each task, with the P-bet on the left and $-bet on the right, and the probabilities of each above their respective prizes. Arrows indicate the direction of the transition, with focus transitions indicated inside the boxes representing the attributes of the bets. The figure brings out several points. First, the direction of the transition (e.g. from P- to $-bet or from probability to prize) is quite evenly distributed; the two arrows between any pair of boxes are always quite similar. Second, probabilities require more attention than prizes, in that the focus on the upper boxes is substantially higher than that on the lower.[27] Third, focus and within-bet (vertical) transitions are more preponderant in Evaluation than in Choice, while across-bet (horizontal) transitions are (as arithmetically necessary) correspondingly more frequent in Choice than in Evaluation. Every one of the cross-task comparisons is statistically significant in a Mann-Whitney or proportion test. We therefore establish rather conclusively the following:

Result 6: Choice encourages cross-bet comparisons compared to Evaluation.

[27] In our computer display, probabilities were always presented above their associated prizes, and Kim et al. (2012) find that the attribute at the top receives more fixations. However, the placement was constant across tasks, so that the task differences reported in this paper should be unaffected by this aspect.

The finding that Choice leads to cross-bet comparison fits rather nicely with the previous results that relative focus on one bet or the other drives behavior in Choice but not in Evaluation (4.3) It seems plausible that the insignificant influence of relative focus on the P- versus $-bet upon behavior in Evaluation may be related to the observation that in Evaluation, only 13 percent of transitions are across bets. The transition analysis helps to explain the phenomenon, in that deliberation in Evaluation appears to be based mainly within-bets, rather than across. We interpret this as reducing awareness among subjects of relevant factors that come from comparisons across bets. There is an aspect of what subjects perceive in cross-bet evaluations that affects their behavior in Choice, which is somehow obscured by the within-bet transitions that dominate in Evaluation.

On the other hand, were it the case that across- versus within-bet transitions caused differences in perception that then affected behavior, then we might expect that between participants, those who had more Across-bet transitions during Evaluation would be less likely to commit reversals, or at least that differences in the preponderance of transition types would explain behavior. In short, transitions seem related to our Task question, but what about Behavior and Reversals? We study this with a series of regressions like those in previous sections. The results are presented in Table XII

	(1)	(2)	(3)	(4)	(5)	(6)
	Across	Chose P	Rated S	Reversal	Reversal	Reversal
VARIABLES				Both tasks	Choice	Evaluation
Choice Task	0.188					
	(0.0167)					
Across		0.638	1.008	0.0694	-0.0582	0.334
		(1.401)	(1.673)	(0.0622)	(0.150)	(0.210)
Constant	0.148	0.318	-0.691	0.00958	0.0415	-0.0424
	(0.0721)	(2.676)	(2.138)	(0.0377)	(0.0383)	(0.0932)
Controls	ALL	ALL	ALL	IND	IND	IND
Observations	409	202	207	409	202	207
R-squared	0.451			0.434	0.424	0.450

Note: Robust standard errors in parentheses.

Table XII: Effect of proportion of within- versus across transitions.

As above, we see that the Choice task greatly increases the prevalence of "across" transitions relative to "within" ones. However, individual differences in this allocation do not seem related to choice or to reversals. We use the share of fixations that are part of an across transition at the bet level, but the results are unchanged if we use the number of transitions instead. We also consider another possibility. Our results on Depth of Reflection indicated that participants with more fixations in the Evaluation task are less likely to commit reversals. It therefore seems possible given the results about transitions that those who focus more during evaluation also shift to between-bet transitions, but pooling over the whole sample dilutes the effect beyond what our statistics can measure. If this were true, it would imply that participants with more fixations in the Evaluation task would also have a greater share of those fixations in “across” transitions. However, we find no evidence of this, either. The share of across transitions is unrelated to the total number of fixations in the Evaluation task (corr. = 0.0165, p = 0.8179), and a regression of the proportion of “across” transitions on the total number of fixations and our controls does little better (coeff. = 0.00018, p = 0.690). We interpret this to mean that, while the evaluation mode (within or across bets) may be behind reversals, it is only indirectly responsible. Preferences themselves do not depend on the direction of transitions, but rather within-bet transitions in Evaluation make it difficult for subjects to “see” the cross-bet comparisons that drive behavior more transparently in Choice.

Result 7: Differences in transition patterns across task type reflect differences in attention, not differences in preferences.

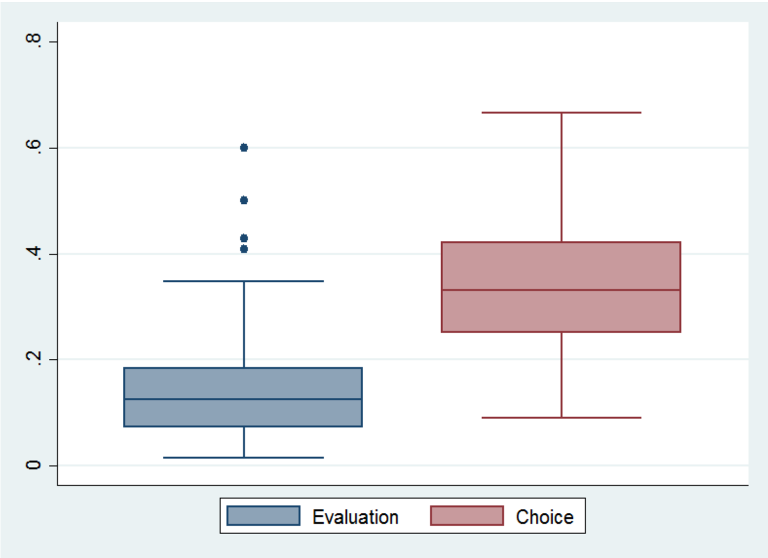


Figure 6: distribution of the share of transitions across bets in a pair, by task condition

Why do transition differences not predict behavior? One possibility is that they are not causally related. But another is that variation across individuals is insufficient to generate statistical associations. Figure 6 shows distributional graphs (box-plots) of the proportion of fixations measured as across at the pair level in each task. Beyond the marked difference in the distributions, we see that particularly in the Evaluation task, these are quite condensed. Indeed, the inter-quartile ranges (identified with the boxes) of the two conditions do not overlap, and leaving aside some outliers, the maximum value in Evaluation is similar to the median in Choice. This suggests that there may not be enough overlap in the support of the distributions to make our comparisons possible. Perhaps across-type transitions in Evaluation, for example, do not significantly affect the chance of rating the P-bet higher simply because an insignificant number of participants had a sufficiently high proportion to do so.

5. Conclusions

This paper explores several mechanisms of how inattention and differential attention allocation can generate incoherent behavior. Our analysis has revealed several facts. First, Evaluation substantially and significantly increases the depth of reflection on each bet pair, relative to Choice. Across individuals, second, those who focus more during Evaluation have fewer reversals – a relationship that does not hold in Choice. These two facts are in line with the idea that, due to the cognitive difficulty of Evaluation, substantially more attention is required in that task to prevent mistakes, which may occur otherwise. Third, the task form influences relative attention across bet pairs, with the Choice task increasing focus on the P-bet generally. In Choice but not in Evaluation, fourth, relative attention on the P-bet correlates with behavior, and specifically with reversals in the corresponding bet pair. That is, we identify an attentional asymmetry that mirrors the behavioral asymmetry exhibited in reversals, bolstering the claim of a causal link. Fifth, there is no analogous asymmetry with respect to attention on the attributes of the bets, i.e., prizes and probabilities, and it seems difficult to explain reversals based on these variables. In this respect, we find little evidence in line with the Prominence or Compatibility hypotheses.

A sixth key finding refers to the comparisons that people make across bets and merits a longer discussion. That is, we see that Choice nearly triples across-bet transitions relative to Evaluation –importantly, both bets in a pair are visible simultaneously in both tasks, a key and novel feature of our design. Hence this is additional evidence that patterns of attention, and thus plausibly of deliberation, are qualitatively different in the two tasks. One possible conjecture is that risk aversion is more salient in Choice, whereas Evaluation reduces risk aversion by drawing attention away from the relative risk across bets. Those with sufficient attention budgets manage to overcome this by careful scrutiny of each bet, but participants who spend less overall attention fall into reversals as a result. However, reversals are uncorrelated with the number of across-bet transitions, which sheds some doubt on this conjecture. Regardless, the conjecture suggests that risk-averse behavior may to some extent be triggered by comparisons across bets, a point for potential future research in the domain of risk elicitation, i.e., comparing calculated parameters for risk aversion across comparison-based versus evaluation-based elicitation procedures.

Preference reversals are a venerable and stubborn anomaly with important practical and policy implications. The normative basis for essentially all economic analysis rests on the assumption that behavior reflects subjective values, and reversals seem to indicate that it may not reliably do so. This casts doubt upon cost-benefit analyses in key areas such as health, safety, or environmental protection, for example, which often operate through asking for individuals' "willingness to pay" for the good in question. But it also may have implications even to market behavior. Consider behavior in auctions, for instance. E-bay, the online consumer-to-consumer auction site, had a gross merchandising sales volume of 85.5 billion USD in 2019,[28] all of which is based on individuals stating their value for different goods. Even a very small systematic bias in evaluations of this kind can have a large impact on global wellbeing. Our study did not explicitly test the various partially competing explanations for where this bias may originate. However, it does offer some reassurances, in that we find that when they pay enough attention, people seem to behave in a consistent manner. We believe our contribution to this already extremely rich field lies in our simple, intuitive explanation of the problem. If we want to elicit values that more reliably reflect the choices people make, we simply require elicitation methods that force people to pay attention, like those discussed in Loomes and Pogrebna (2017, pp. 181-182).

[28] See annual report at https://d18rn0p25nwr6d.cloudfront.net/CIK-0001065088/d33d35e7-32e8-4a9c-ad67-12baec291433.pdf

Bibliography

Alós-Ferrer, C., Granić, Đ. G., Kern, J., and Wagner A. K. (2016). Preference reversals: Time and again. *Journal of Risk and Uncertainty*, 52, 65-97.

Arieli, A., Ben-Ami, Y., & Rubinstein, A. (2011). Tracking decision makers under uncertainty. *American Economic Journal: Microeconomics*, *3*(4), 68-76.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly journal of economics*, *127*(3), 1243-1285.

Bostic, R., Herrnstein, R., and Luce, D. (1990). The effect on the preference-reversal phenomenon of using choice indifference, *Journal of Economic Behavior & Organization*, 13, pp. 193-222.

Blavatskyy, Pavlo R. (2009). Preference Reversals and Probabilistic Decisions, *Journal of Risk and Uncertainty*, 39, 3, 237.

Butler, D. J., & Loomes, G. C. (2007). Imprecision as an account of the preference reversal phenomenon. *American Economic Review*, *97*(1), 277-297.

Cubitt, R. P., Munro, A., & Starmer, C. (2004). Testing explanations of preference reversal. *The Economic Journal*, *114*(497), 709-726.

Devetag, G., Di Guida, S., & Polonio, L. (2016). An eye-tracking study of feature-based choice in one-shot games. Experimental Economics, 19(1), 177-201.

Fischbacher, Urs (2007). z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10(2), pp. 171-178.

Fischer, G.W., Carmon, Z., Ariely, D. and Zauberman, G. (1999). Goal-based construction of preferences: task goals and the prominence effect, *Management Science*, 45 (8), pp. 1057–75.

Franco-Watkins, A. M., and Johnson, J. G. (2011). Decision moving window: Using interactive eye tracking to examine decision processes. *Behavioral Research Methods*, 43, 853–863.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, *19*(4), 25-42.

Frydman, C. and Mormann, M. M. (2018) The Role of Salience in Choice under Risk: An Experimental Investigation, Available at SSRN: https://ssrn.com/abstract=2778822

Gabaix, X. (2019). Behavioral Inattention, in *Handbook of Behavioral Economics*, Vol. 2, eds. D. Bernheim, S. DellaVigna, and D. Laibson, Elsevier.

Hsee, C. K., S. Blount, G. F. Loewenstein, and M. H. Bazerman (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis, *Psychological Bulletin*, 125(5), 576-59.

Johnson, E. J., J. W. Payne, and J. R. Bettman (1988). Information displays and preference reversals. *Organizational Behavior and Human Decision Processes*, 42, 1–21.

Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension, *Psychological Review*, 87 (4), 329-354.

Kim, B., D. Seligman, and J. Kable (2012). Preference Reversals in Decision Making Under Risk Are Accompanied by Changes in Attention to Different Attributes. *Frontiers in neuroscience*, 6, 109.

Tversky, Amos; Kahneman, Daniel (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*. **5** (4): 297–323.

Köszegi, B., and Szeidl, A. (2013). A model of focusing in economic choice. *Quarterly Journal of Economics* 128 (1), 53-104.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and comparison of value in simple choice. *Nature Neuroscience* 13, 1292–1298.

Krajbich, I. (2019). Accounting for attention in sequential sampling models of decision making. *Current Opinion in Psychology* 29, 6-11.

Lahey, J. N., & Oxley, D. (2016). The power of eye tracking in economics experiments. *American Economic Review*, *106*(5), 309-13.

Lichtenstein, Sarah and Paul Slovic (1971). Reversals of Preference between Bids and Choices in Gambling Decisions, *Journal of Experimental Psychology*, 89, pp. 46-55.

Lichtenstein, Sarah and Paul Slovic (1973). Response-Induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas, *Journal of Experimental Psychology*, 101 (1), pp. 16-20.

Loomes, G. and G. Pogrebna (2017). Do Preference Reversals Disappear When We Allow for Probabilistic Choice? *Management Science* 63(1), pp.166-184.

Loomes, G. and R. Sugden (1982). Regret Theory: An Alternative Theory of Rational Choice under Uncertainty, *The Economic Journal* 92, pp. 805–24.

MacCrimmon, K. and M. Smith (1986). Imprecise Equivalences: Preference Reversals in Money and Probability, University of British Columbia Working Paper 1211.

Manor, B.R., Gordon, E., 2003. Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. J. Neurosci. Methods128 (1), 85–93.

Orquin, J. L., and S. Mueller Loose (2013). Attention and choice: A review of decision making and eye tracking studies. *Acta Psychologica*, 144, 190–206.

Schkade, D. A., and E. J. Johnson (1988). Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes*, 44, 203–231.

Schmidt, U., and Hey, J.D. (2004). Are preference reversals errors? An experimental investigation. *The Journal of Risk and Uncertainty*, 29(3), 207–218.

Seidl, C. (2002). Preference reversal, *Journal of Economic Surveys*, 6, pp. 621-55.

Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye Movements in Risky Choice. *Journal of Behavioral Decision Making*, 29(2-3), 116-136.

Taylor, S., and S. Thompson (1982). Stalking the ellusive "vividness" effect, *Psychological Review*, 89 (2), pp. 155-181.

Tversky, A., P. Slovic, and D. Kahneman (1990). The Causes of Preference Reversal, *American Economic Review* 80, pp. 204–17.

Tversky, A., Sattath, S. and Slovic, P. (1988). Contingent weighting in judgement and choice, *Psychological Review,* 95(3), pp. 371–84.

Tversky, A. and R. H. Thaler (1990). Anomalies: Preference Reversals, *Journal of Economic Perspectives* 4 (2), pp. 201-211.

Appendix I: Instructions

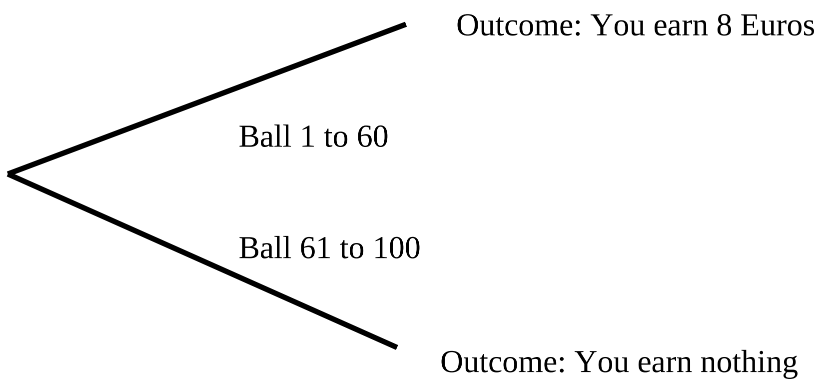
Thank you for participating in this experiment on choice under uncertainty. You will make decisions with your computer keyboard and mouse. Read all the information appearing on your screen carefully. There is no hurry to decide. At the end of the experiment, you will be paid some money; the precise amount will depend on chance and your decisions.

It is very important that you not talk to other participants. If you do not follow this rule we will have to exclude you from the experiment. If you have questions, please raise your hand and we will assist you.

Description of the Experiment

In this experiment we speak of **bets**. Each bet has two possible **outcomes**: (i) Earn some money, (ii) earn nothing. Each outcome has a **probability**, described with reference to a random draw of a ball from a bag with 100 numbered balls. Example of a bet (see figure below): You win 8 Euros if balls 1 to 60 are drawn and 0 Euros otherwise. The probability of winning in this bet is 60/100.

Example of bet:



On the other hand, during the experiment there are two types of decisions, described below.

- ***Choice***: You are asked to consider a pair of bets, and indicate your favorite, or express indifference between them.
- ***Evaluation***: You are asked to consider one bet, and indicate its value to you in Euros. In other words, you should state the amount of money M that would leave you indifferent between the bet and M.

There are a total of 12 different bets, grouped in six different pairs. For each pair of bets (A and B) you will make three decisions: (i) *Choice* between A and B, (ii) *Evaluation* of A, (iii) *Evaluation* of B.

After you have made the 18 decisions (12 evaluations and 6 choices), you will be given a paid memory test. In this test, you will be given 10 seconds to memorize an arrangement of four multi-digit numbers. You will then be asked to re-enter the numbers, in order, in the following screen. You will receive €0.25 for each number correctly entered. This will be followed by a short questionnaire. Finally, the bet you will play for real money will be determined in the following way.

Which of the bets is played for cash will be determined by the roll of a 6-sided die, the flip of a coin, and your previous choices. The roll determines one of the six bet pairs (call it P); the coin flip determines one decision type (T, either *Choice* or *Evaluation*); and then your decision of type T for pair of bets P fixes the selected bet, (call it S). More precisely:

- *Choice* is selected if the coin flip is heads. In this case the selected bet S will be that in pair P which you had indicated as your favorite in the *Choice* decision – if you expressed indifference in that decision, S will be randomly selected by the computer.
- *Evaluation* is selected if the coin flip is tails. In this case the selected bet S will be the one in P to which you gave a higher evaluation. If both evaluations were equal, the computer will again randomly select one of the options in P.

Once S has been decided, your payoff will be determined by the computer randomly generating a number from 1 to 100; you will get the corresponding prize if the number is less than or equal to the corresponding probability from the bet.

Your final payment will be computed as **(Your payoff from S) + (6 Euros Show-up Fee) + (0.25 for each number correctly recalled and entered)**. You will receive this through bank transfer.

<u>Before we start the experiment, please answer the following questions. Raise your hand when you are done so that we can verify the answers.</u>

- How many different bets are there in this experiment?_______________
- A number of these bets will be randomly chosen at the end of the experiment to play for real money. How many bets will you end up playing for real money? ______________

In a hypothetical example, consider a bet A that gives ________ Euros (choose a number to construct your own example) if a ball from 1 to ____ (choose a number from 1 to 100) is selected, and zero otherwise. Suppose that this bet A is finally selected as the pay-bet. For this hypothetical example,

- What would be your probability of winning? _______

Suppose that bet A in the previous example is paired with another bet B that gives _________ Euros if a ball from 1 to _____ is selected (choose numbers so as to construct your own example of a bet), and zero otherwise. In the *Choice* decision, you choose option A B (select your preferred bet), and in the *Evaluation* decisions, you estimate that option A is worth ______ Euros and option B is worth ______ Euros (indicate each evaluation). Assume that this pair of bets (A-B) is randomly chosen at the end of the experiment.

- If the *Choice* decision is randomly chosen, what bet will you play? ___________
- If the *Evaluation* decision is randomly chosen, what bet will you play? ____________