PREPRINT

Not Peer Reviewed

RESEARCH ARTICLE

Using eye-tracking to examine strategies for evaluating compound lotteries

[version 1]

Narges Hajimoladarvish[1,2*], Sneha Shashidhara[1*]

[1]Centre for Social and Behavioural Change, Ashoka University, New Delhi, New Delhi, 110057, India
[2]Tehran Institute for Advanced Studies, Khatam University, Tehran, Iran

* Equal contributors

v1 **First published:** 26 Aug 2025, **2**:255
https://doi.org/10.12688/verixiv.1840.1

Latest published: 26 Aug 2025, **2**:255
https://doi.org/10.12688/verixiv.1840.1

Abstract

-->We use eye-tracking to investigate how participants evaluate compound lotteries, differentiating between two competing strategies: Backward induction, which aligns with the compound independence (CI) axiom, and forward induction, which supports the reduction of compound lotteries (ROCL) axiom. Combining eye-tracking and behavioural data, our results support the ROCL axiom as the predominant strategy. In a second, non-lottery task, eye-tracking further supports the forward induction strategy, offering a broader application of our findings.

Keywords

Compound lotteries, Reduction of compound lotteries axiom, Compound independence axiom, Eye-tracking, Backward induction strategy, Forward induction strategy.

Corresponding author: Narges Hajimoladarvish (narges.hajimolahdarvish@ashoka.edu.in)

Author roles: Hajimoladarvish N: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Shashidhara S**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The study was funded by Bill and Melinda Gates Foundation (INV026871).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Hajimoladarvish N and Shashidhara S. **Using eye-tracking to examine strategies for evaluating compound lotteries [version 1]** VeriXiv 2025, **2**:255 https://doi.org/10.12688/verixiv.1840.1

First published: 26 Aug 2025, **2**:255 https://doi.org/10.12688/verixiv.1840.1

1. Introduction

Understanding how people evaluate compound lotteries, especially when faced with complex decisions, can provide insights into broader decision-making processes, with implications for economic theory, policy design, and behavioural science. In decision theory, lotteries model risky choices with probabilistic outcomes. A compound lottery is a lottery where the outcomes themselves are lotteries, adding multiple stages of risk. Reducing such compound lotteries to simpler ones—the Reduction of Compound Lotteries (ROCL) axiom—is a central tenet of expected utility theory. ROCL axiom posits that a compound lottery can be simplified into an equivalent simple lottery, preserving the same expected utility. This reduction is based on the expected utility theory, where decision-makers rely on the final-stage outcomes and their probabilities (Von Neumann & Morgenstern, 1944), akin to a forward induction approach to problem-solving.

On the other hand, the Compound Independence (CI) axiom proposes an alternative approach to simplifying compound lotteries, where each stage is evaluated independently. CI axiom, compatible with many decision theories, suggests that people first evaluate the final-stage lotteries separately and use those evaluations to simplify the overall lottery process like backward induction (Segal, 1990). Eye-tracking data can help determine which strategy participants predominantly follow by analysing the sequence and timing of their visual fixations. The difference between these two strategies is depicted in Figure 1.

1.1 Mathematical formulation

A simple lottery, denoted as S, is a list $S=(x_1,p_1;x_2,p_2;\ldots;x_n,p_n)$ where $p_i \geq 0$ for all i and $\sum_{i=1}^{n} p_i = 1$. Here, p_i represents the probability of outcome x_i occurring. Formally, these outcomes could take many forms. However, we consider them monetary payoffs for simplicity and denote the finite set of all outcomes as X. For convenience, assume $x_1 < x_2 < \ldots < x_n$. We denote the set of simple lotteries by φ. A compound lottery is a lottery that allows the outcomes to be lotteries, represented as $C=(S_1,q_1;S_2,q_2;\ldots;S_n,q_n)$ where $q_i \geq 0$ for all i and $\sum_{i=1}^{n} q_i = 1$, and $S_1,\ldots,S_n$ are simple lotteries in φ.[1]

Under the expected utility (EU) theory, the *reduction of compound lotteries* (ROCL) axiom ensures that compound lotteries can be reduced to simple lotteries.

ROCL axiom. *Let $S_1=(x_1,p_1;\ldots;x_i,p_i;\ldots;x_n,p_n)$ and $S_2=(y_1,q_1;\ldots;y_i,q_i;\ldots;y_n,q_n)$. Then, under the ROCL axiom:*

$$C=(S_1,r;S_2,1-r) \sim (x_1,rp_1;\ldots;x_n,rp_n;y_1,q_1(1-r);\ldots;y_n,q_n(1-r)) = AE(C)$$

For any compound lottery, we can reduce the lottery to a simple lottery that generates the same probability distribution of outcomes. Hence, the ROCL axiom ensures that a compound lottery can be reduced to a simple one, which has been called the actuarially equivalent lottery (AE). Thus, decisions are made based on the reduced lottery over the final level of outcomes (Von Neumann & Morgenstern, 1944). For example, consider a two-stage compound lottery $A_1 = (L,r;x_1,1-r)$ where $L=(y_1,q;y_2,1-q)$. Under the ROCL, the lottery A_1 is reduced to its actuarially equivalent lottery $AE(A_1)=(y_1,rq;y_2,r(1-q);x_1,(1-r))$. Panel A of Figure 1 depicts both of these lotteries for $A_1=(L,0.5;4000,0.5)$ and $L=(1000,0.5;2000,0.5)$.

The compound independence (CI) axiom is an alternative to the ROCL axiom in which compound lotteries are reduced into simple ones by successive substitution of certainty equivalents (CE) of last-stage lotteries (Segal, 1990).[2] It implies that decision-makers first evaluate the last-stage lotteries of a compound lottery, then reduce and evaluate the lottery.

CI axiom. *Consider the two-stage compound lotteries $A=(S,r;SA,1-r)$ and $B=(S,r;SB,1-r)$. The preference relation $\succcurlyeq$ satisfies the CI axiom if for all $S,SA,SB \in \varphi$ and $r \in [0,1]$ we have:*

$$A \succcurlyeq B \text{ if, and only if, } SA \succcurlyeq SB.$$

If subjects follow the CI axiom, they reduce the above-mentioned compound lottery A_1 as $(CE(L),r;x_1,1-r)$, where $CE(L)$ stands for the certainty equivalent of lottery L. Hence, subjects should first evaluate the lottery L and then focus on r and the other outcome to evaluate a reduced simple lottery. Panel B of Figure 1 depicts this mechanism. Under expected value and EU, these two axioms have similar predictions. However, the predictions differ for rank-dependent utility or prospect theory preferences.

---

[1]In the real world, we can have multi-stage lotteries where the outcome of compound lotteries are compound lotteries as well. However, in our experiments, we only deal with two-stage lotteries.

[2]Certainty equivalent is a certain amount that makes the decision maker indifferent between that amount and the lottery. These certainty equivalents can be determined by any theory and as such the CI axiom can be used with non-EU theories as well.

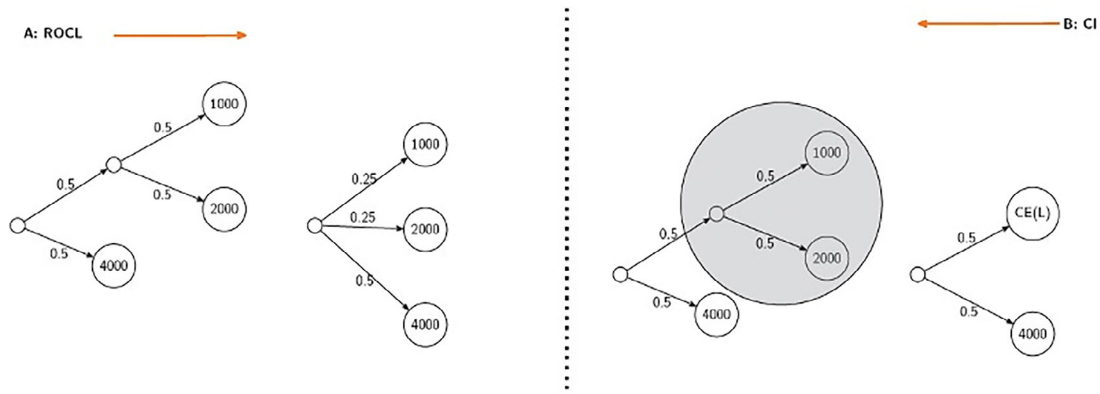

Figure 1. ROCL and CI in action. Panel A depicts the reduction of compound lotteries under ROCL, and Panel B shows how the CI strategy works.

To compare forward and backward induction strategies, we focus on these two key axioms in decision theory, ROCL and CI. These axioms represent two competing strategies for evaluating compound lotteries. While both aim to simplify the evaluation of compound lotteries, they differ in how the lottery stages are handled. The ROCL axiom posits that a compound lottery can be reduced to a simpler, actuarially equivalent lottery[3]. In contrast, the CI axiom advocates for evaluating each stage of the compound lottery separately. We aim to provide the first test of the ROCL and CI axioms through eye-tracking data. For the purpose of this study, we use forward induction to mean the ROCL and backward induction to mean the CI. Eye-tracking is helpful because it is difficult to distinguish between the two axioms by choice or valuation data alone (Hajimoladarvish, 2018).

In a valuation task, the evaluation occurs in isolation for each lottery, whereas in a choice task, two options are assessed concurrently. More importantly, if subjects follow the expected value or EU when reducing compound lotteries, the two axioms and usual tests will both hold. Thus, we also test the validity of these axioms through a novel test that combines the conventional methods used in the literature with choice data and valuation data. Furthermore, we examine whether results from eye-tracking data and self-reported valuation of lotteries are consistent. An ideal outcome would be for both datasets to indicate one axiom. The ROCL or CI axioms justify the validity of the *random lottery incentive mechanism* in experiments[4]. The random lottery incentive mechanism is incentive-compatible only when either of these two axioms holds (Starmer and Sugden, 1991).

Eye-tracking provides an innovative method for examining decision-makers' strategies when evaluating compound lotteries. By capturing participants' gaze patterns, we can directly observe the sequence in which they process information about lotteries. This allows us to distinguish between the forward-induction strategy associated with the ROCL axiom and the backward induction strategy related to the CI axiom. Unlike traditional choice or valuation data, which may not capture the cognitive processes underlying decision-making, eye-tracking provides real-time insights into how individuals interact with complex risk scenarios.

Eye-tracking data offers valuable insights into various aspects of lottery decision-making. Firstly, gaze patterns provide information about attention allocation during decision-making (Harrison and Swarthout, 2019). By analysing fixations on different attributes of lottery options, researchers can assess the salience of specific features and their impact on choice behaviour (Frydman & Mormann, 2018; Glaholt et al., 2009; Wedell & Pettibone, 1996). Moreover, eye-tracking data elucidate information processing strategies, such as the order in which individuals attend to different attributes and the duration of fixation on each attribute (Krajbich et al., 2010; Glöckner et al., 2011). This information shows how individuals integrate information and form preferences in lottery decision-making contexts. For more basic evidence on eye-tracking, attention, information processing and reading, see reviews (Rayner, 1998; Rayner, 2009; Hoffman, 1998).

Decision-making under risk has been explored by eye-tracking with choices over lotteries (Rosen and Rosenkoetter, 1976; Russo and Dosher, 1983; Arieli et al., 2011; Glöckner and Herbold, 2011; Fiedler and Glöckner, 2012;

[3]The term 'actuarially equivalent lottery' refers to a simple lottery that, when reduced from a compound lottery, produces the same expected outcome. In other words, it's a simplified version of a compound lottery that still reflects the same risk profile, making it easier to evaluate.
[4]See Supplementary section for demonstration of this claim.

Janowski, 2012; Su et al., 2013; Stewart et al., 2016; Harrison and Swarthout, 2019; Zhang et al., 2024). While all these studies used data while subjects were making choices over alternatives, we used data while they were evaluating lotteries individually, which has relatively been less studied. To our knowledge, no study has used eye-tracking to directly test the ROCL and CI axioms, despite their foundational importance in decision theory. Our study fills this gap by using eye-tracking to test these competing axioms in the context of compound lotteries, offering new insights into how decision-makers process complex risk information.

To extend our strategy test beyond numeric lotteries, we added a matrix-analogy task in which participants infer the missing tile of a 3×3 pattern before choosing among four options. This paradigm poses an analogous strategic contrast—constructive matching/forward induction (derive the rule from the pattern, then consult options) versus response elimination (retrofits the answer options to the pattern)—well documented in reasoning research (e.g., Bethell-Fox et al., 1984; Vigneau et al., 2006). Eye-tracking again lets us trace sequence and timing, complementing prior work on attention and information processing (Rayner, 1998; Rayner, 2009; Krajbich et al., 2010). Using the same subjects and a parallel difficulty manipulation (easy/hard) allows us to verify that harder problems slow responses and reduce accuracy, mirroring the lottery task.

We contribute to the evaluation of compound lotteries in three ways. First, we test the consistency prediction of both the ROCL and the CI axioms using certainty equivalents of compound lotteries. Second, we use eye-tracking for further evidence of the axioms used as strategies in decision-making. Last, we compare problem-solving strategies in tasks of different natures and domains.

The paper is organised as follows. Section 2 outlines the experimental design. Section 3 describes the methodology used to test the ROCL and CI axioms. Section 4 presents the data and main findings. Section 5 provides a discussion of the results and concludes. Experimental instructions, additional tables and figures, robustness checks, and details of the follow-up experiment are provided in the Supplementary section (https://osf.io/z7s5j/?view_only=8a0188951d734430bdf64eb052aa6524).

2. Experimental design

The experiment consists of two main tasks. A lottery evaluation task, where participants evaluate both simple and compound lotteries, and a pattern completion task, which serves as an analogical reasoning task to test general problem-solving strategies. The primary aim of the lottery task is to investigate how participants process compound lotteries in relation to the ROCL and CI axioms, with both tasks providing evidence for either forward or backward induction strategies. We use eyetracking for both tasks. The order in which these two tasks were completed was counterbalanced across participants.

2.1 Task 1

To test the ROCL and CI axioms, we elicited certainty equivalents of compound lotteries, their corresponding actuarially equivalent lotteries, and the second-stage lotteries on their own (simple lotteries). We used two sets of compound lotteries, A and B, where A had a greater expected value than B. Each lottery was presented separately, and participants indicated their certainty equivalent by typing the amount in a text box below the lottery description.[5] Each trial had no time limit, and a one-second fixation screen appeared between trials to reset attention.

Compound lotteries are denoted by A and B, and simple lotteries are marked as $S, SA, SB, AE(A)$ and $AE(B)$. Simple lotteries can have either two or three possible outcomes. The lotteries S, SA, and SB are examples of two-outcome simple lotteries, while most $AE(A)$ and $AE(B)$ are three-outcome simple lotteries. $AE(A)$ and $AE(B)$ lotteries are actuarially equivalent lotteries corresponding to compound lotteries A and B. Compound lotteries, on the other hand, are structured in two stages. They begin with a simple two-outcome lottery, and additional complexity is introduced by modifying the outcomes. A compound-up lottery replaces the upper outcome with a two-outcome simple lottery, while a compound-low lottery replaces the lower outcome with a two-outcome simple lottery. In a compound-2x lottery, both the upper and lower outcomes are replaced with two-outcome simple lotteries.

To employ eye-tracking technology, we elicited certainty equivalents by directly asking subjects to evaluate lotteries. This approach ensures that subjects concentrate on the lottery without shifting their gaze between different amounts, as in multiple price lists, or consistently fixating on sure amounts, as in the bisection procedure. Participants typed in their

[5]We did not constrain participants' valuations to fall within the range of lottery outcomes. Nonetheless, 97.66% of responses were within this range.

values for each lottery in the text box displayed below the lottery in each trial. They could delete and retype, and the screen moved to the subsequent trial only when they pressed the Enter key.

To test the ROCL axiom, we must elicit the certainty equivalent of compound lotteries and their associated actuarially equivalent lotteries. To test the CI axiom, we need to elicit certainty equivalents of four lotteries, two compound lotteries like *A* and *B*, and two distinguishing simple lotteries *SA* and *SB* embedded in compound lotteries *A* and *B*. Following Harrison et al. (2015), we used compound lotteries with no more than 3 distinct outcomes, ensuring manageable complexity in participant decision-making and experimental design.

The battery of lotteries is illustrated in Figure 2 using the Marschak-Machina triangle. This device represents three-outcome lotteries, where each point within the triangle corresponds to a specific combination of outcome probabilities. In our design, outcomes are drawn from the set {1000, 2000, 4000}, and probabilities span a wide range to facilitate the detection of violations of expected utility theory. Moreover, these lotteries mainly cover the southeast of the triangle, where most EU violations happen. The horizontal axis indicates the probability of the lowest outcome (1000), while the vertical axis corresponds to the probability of the highest outcome (4000). The left panel displays the coverage of the Marschak-Machina triangle by the simple distinguishing lotteries SA and SB, highlighted in red and blue, respectively. The right panel shows the triangle coverage for compound lotteries A and B, highlighted in red and blue, respectively. Numbers within the triangles refer to the lottery pairs listed in Table S1 in the Supplementary section describing the battery of compound lotteries. Table S2 of the Supplementary section presents the associated actuarially equivalent lotteries. Note that lotteries B are first-order stochastically dominated by lotteries A, meaning that lottery B has a lower expected value.

All lotteries were represented as decision trees. We opted for horizontal decision trees as Arieli et al. (2011) found that horizontal transitions are more frequent than diagonal ones. Figure 3 (panels A and B) demonstrates a compound lottery and how we refer to its different outcomes and probabilities. We explicitly displayed all lottery information and randomised the order of the lotteries for each subject. Moreover, we displayed the lotteries such that 1/3 of them appeared on the left of the screen, 1/3 centred around the middle and 1/3 to the right. We also counterbalanced outcomes and probabilities across positions. This was critical for getting unbiased eye-tracking data. This counterbalancing was omitted for the *AE*(*A*) and *AE*(*B*), which always followed the rest of the lotteries.

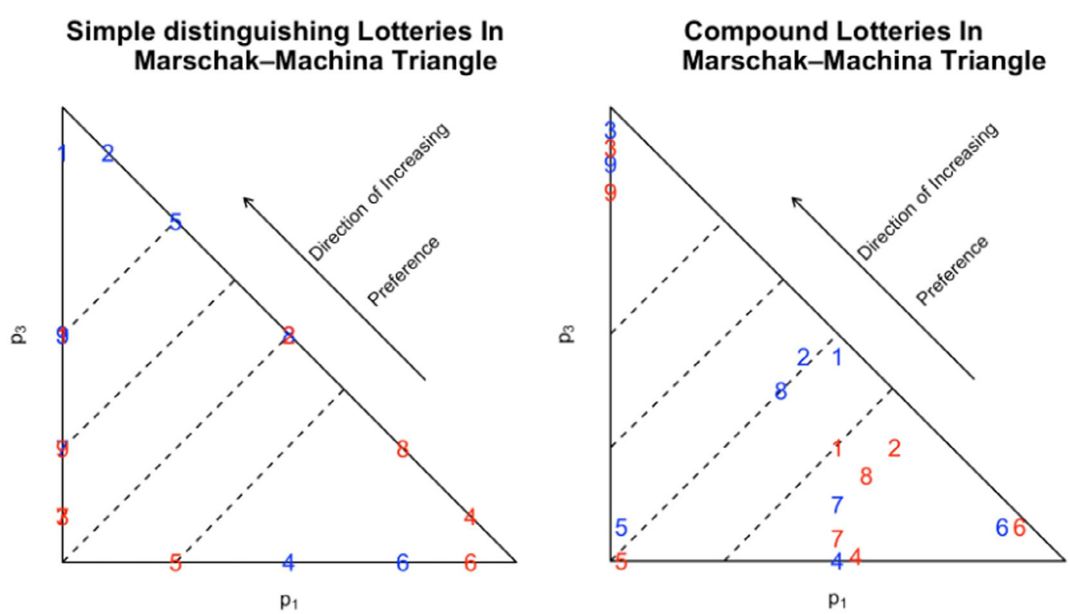

Figure 2. Battery of Lotteries in the Marschak-Machina Triangle. The triangles depict the probabilities of different outcomes in three-outcome lotteries. The horizontal axis represents the probability of the lowest outcome, while the vertical axis indicates the probability of the highest outcome. The left panel illustrates the coverage of the Marschak-Machina triangle by the simple distinguishing lotteries (*SA* and *SB*) in compound lotteries *A* and *B*. The right panel shows the triangle coverage for compound lotteries *A* and *B*. The numbers within the triangles correspond to the lottery pairs described in Table S1 in the supplementary section.

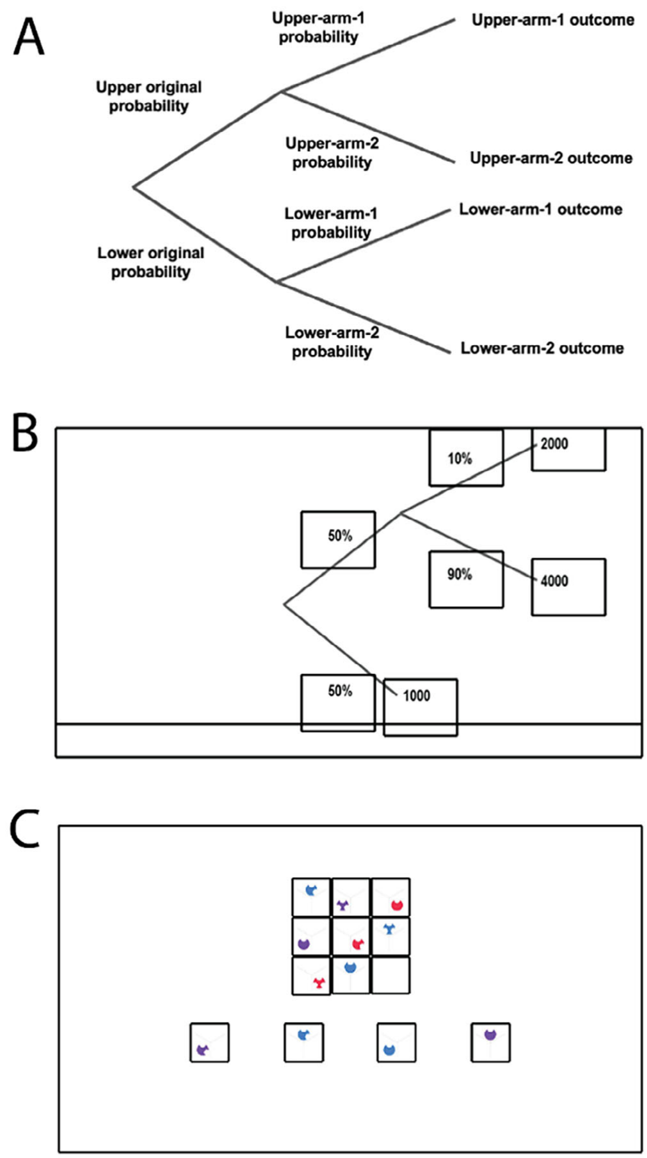

Figure 3. Task paradigm. A. Decision tree of a compound-2x lottery shown with the nomenclature followed for each probability and outcome. These are consistent across lotteries. B. An example of a compound-up lottery. The segment at the bottom of the screen illustrates the area where subjects entered the value of the lottery. C. An example of a hard trial of the reasoning task. The boxes in B and C indicate the area of analysis for eye-tracking and are not displayed during the task.

2.2 Task 2

The pattern completion task involved participants inferring the missing piece of a 3×3 matrix (Figure 3C shows an example trial) of geometric shapes before selecting from four possible options. Participants selected one option that fit the missing tile by pressing the keyboard number keys 1, 2, 3, or 4. There were 18 easy and 18 hard trials. In Easy, the missing tile was determined by one feature, like colour or shape, and in Hard, it was determined by a conjunction of two or more features. There was no time constraint and a 500ms fixation screen between trials. This task is modelled after Raven's progressive matrices (Carpenter et al., 1990; Vigneau & Bors, 2008; Raven and Raven, 2003).

In visuo-spatial reasoning tasks or relational reasoning tasks, there are two main strategies that one can use - constructive matching, which involves mentally constructing the missing piece of the pattern, or response elimination, where participants systematically compare the options given with the pattern to find one that seems to fit (Bethell-Fox et al., 1984; Vigneau et al., 2006; Snow, 1980; Gonthier & Roulin, 2020; Gonthier, C., & Thomassin, N., 2015). These strategies could be thought of as forward and backward induction strategies, respectively, as defined in the decision theory literature. Bethell-Fox et al.'s (1984) study used fixation counts and fixation sequence orders to show that subjects used these two strategies: constructive matching/forward induction strategy correlated with higher intelligence and response elimination/backward induction strategy correlated with lower intelligence. One does not need to follow the same strategy throughout, as evidenced by Gonthier and Roulin (2020), who show that subjects tend to use more response elimination in hard trials. They suggest constructive matching is effective but costly for hard trials.

Both tasks were written and presented using Psychtoolbox-3 (Brainard, 1997) and MATLAB (The MathWorks). Subjects were paid a flat participation fee of 400 rupees, approximating \$4.82. We conducted a small incentivised follow-up (Task 1 only; see Supplementary Section) to test robustness. Participants stated their minimum selling price for each lottery using the BDM mechanism (Becker et al., 1964). Decision trees were randomly flipped, with the starting node appearing on the right in half the trials.

3. Analysis methods

3.1 Methods to test the ROCL and CI axioms using behavioural data

Two methods have directly tested the ROCL axiom of the EU. The first method uses binary choices to examine the hypothesis of equal responses between compound lotteries and their associated AE lotteries. However, this is not a precise test of the axiom. Under the EU, when there is no switching cost, any choice between a compound lottery and its AE lottery is consistent due to having identical values. Starting with Bernasconi (1994), researchers have explored the consistency of preferences for a compound lottery over a simple lottery and the associated AE lottery over the same simple lottery in pairwise choices. The logic is grounded in maintaining consistent preferences between a compound and a simple lottery when the compound lottery is replaced with its AE lottery. This procedure has been followed up by a consistency test in Harrison et al. (2015). They test for across-task consistency in pairwise choices between a compound and a simple lottery $(C-S)$ and another pair with the associated AE lottery of the compound lottery and the same simple lottery $(AE(C)-S)$. For example, if a subject prefers C to, they should also prefer $AE(C)$ to S. Empirical studies using this method find support for the ROCL (Harrison et al., 2015; Hajimoladarvish, 2018). The second method compares the elicited certainty equivalents (CE) of compound lotteries and their AEs. The ROCL axiom is violated if there is a significant difference between the elicited certainty equivalents. Empirical studies using this method do not find support for the ROCL. See, for example, Bar-Hillel (1973), Bernasconi and Loomes (1992), Miao and Zhong (2012), Abdellaoui et al. (2015), Bernasconi and Bernhofer (2020), and Hajimoladarvish (2018).

The extensive literature on the *preference reversal* phenomenon, as documented by Lichtenstein and Slovic (2006), shows that the conclusions from these two tests are different. While the former test relies on choice data, the latter depends on the valuation of certainty equivalents.

The methodology to test the CI axiom is based on similar methods. Hajimoladarvish (2018) tested whether the preference ordering of two compound lotteries, such as A over B, follows the same preference ordering as the distinguishing simple lotteries SA over SB. When both axioms are tested directly through the same methods, Hajimoladarvish (2018) cannot distinguish between them as they are both supported with choice data and violated through the valuation task. In the latter, certainty equivalents of simple lotteries were used to estimate parameters of utility and probability weighting functions, which were then applied to calculate certainty equivalents of compound lotteries according to both axioms. Using choice data, Bernasconi (1994) finds that while 57% of choices are consistent with the CI axiom, 43% are consistent with the ROCL axiom.

In this paper, to distinguish between the two axioms, we test the consistency of preference orderings through elicited certainty equivalents and check the consistency of valuations. Compared to the consistency test with choice data, this test is cognitively more demanding. This new test is instrumental for our design. Essentially, we apply the consistency test to the lotteries' valuation, not preference ordering. This presumes that if we do have well-defined preferences, the preference of A over B translates into A having a higher value than B, and vice versa.

Thus, if participants follow the ROCL axiom, the same difference between the compound and simple lottery should hold between the actuarially equivalent and simple lottery. For the CI axiom, we test the certainty equivalents of A lotteries versus B lotteries. A participant's adherence to the CI axiom is demonstrated if they assign higher values to A than B lotteries, and similarly, to their final-stage components, SA over SB. Table 1 summarises our new method for testing the CI and ROCL axioms and compares it to the consistency test used with choice data.

Table 1. Methods to test the ROCL and CI axioms.

Axiom	Consistency test with choice task	Consistency test with valuation task
ROCL	*If* $A \succcurlyeq S \Longleftrightarrow AE(A) \succcurlyeq S$	*If* $CE(A) \succcurlyeq CE(S) \Longleftrightarrow CE(AE(A)) \succcurlyeq CE(S)$
CI	*If* $A \succcurlyeq B \Longleftrightarrow SA \succcurlyeq SB$	*If* $CE(A) \succcurlyeq CE(B) \Longleftrightarrow CE(SA) \succcurlyeq CE(SB)$

Note: CE denotes certainty equivalent. These preference relations hold for the opposite pattern as well. For example, choices are consistent if $S \succcurlyeq A \Rightarrow S \succcurlyeq AE(A)$ as well.

3.2 Eye-tracking methods

The diameter and position of the participant's left and right pupils were continuously measured using a Tobii Pro fusion eye tracker with a sampling rate of 120 Hz. Participants were in a completely dark room and were seated comfortably, with their eyes approximately 80 cm away from the monitor. Before starting that task, we verified that they could see the numbers in the lottery task and the intricate patterns in the reasoning task stimuli. They were instructed to stay as still as possible during the task and were not provided external fixators like a chin rest. The Tobii advanced algorithm doesn't require the head to be fixated, which was validated in our lab. A 5-point spatial calibration was performed before each of the tasks began. The experimenter viewed the calibration result visually and repeated it if the calibration was poor. The same calibration is used throughout the task and not repeated in-between. None of our subjects took a break within tasks or stopped for any reason. The average duration of the lottery task was 14.28 minutes, and the reasoning task was 11.33 minutes; both were short enough to omit recalibration in between.

In the lottery task, the screen was divided into two regions: the upper portion (1920 × 972 pixels) displayed the lottery and areas of analysis, while the lower portion (1920 × 108 pixels) contained a centrally positioned textbox for entering the participant's valuation. For the reasoning task, the entire screen, 1920 × 1080 pixels, was used to display the task and options. The monitor width and height were 34.4 cm and 19.3 cm, respectively.

Experiments and analyses were conducted on custom-made MATLAB scripts, and an SDK was employed to connect to the eye tracker. The raw position data from the eye tracker was stored as a .mat file at the end of the experiment. This data was averaged across the two pupils. If the data from one pupil was missing, the other pupil's data point was used. The data points were scaled to represent pixels on the screen. Further, fixation and saccade time points were separated using a velocity threshold of 30 degrees per second (Nyström & Holmqvist, 2010; SR Research, 2007), a commonly used threshold. All points below the threshold velocity were considered as fixation. Clean data included fixations that lasted at least 50ms in duration (Rayner, 1998). Raw data and fixation data without cleaning were also analysed for robustness checks.

For each lottery trial, we captured fixation data around the displayed number, i.e. an outcome or probability, defining a window based on its on-screen size and extending roughly 70 pixels in every direction. Table 2 lists window/areas of interest (AOI) sizes for analysis for each trial type. Figure 3B shows the AOIs for a compound-up trial. To ensure there was no overlap between AOIs, in the compound-2x trials, the AOI for the inner outcomes was made smaller in height than the outer outcomes by 40 pixels. None of the AOIs overlapped with the textbox. All lotteries were 1530 pixels in width and 966 pixels in height. This is 10.01° and 7.01° visual angle along the width and height, respectively. In the reasoning task, the analysis AOI was the tile's width and height, plus 8 pixels in all directions, which came to 131 × 131 pixels for each tile. The entire pattern was 393 pixels along width and height, and each option was 131 pixels along width and height. This is 2.51° and 0.84° visual angle along the width and height for pattern and options, respectively.

Table 2. Created areas of analysis for each number in the lottery.

Lottery type	AOI type	Width (px)	Height (px)
Simple lotteries with 2 or 3 outcomes	Outcome	250	135
	Probability	235	140
Compound-up or compound-down lotteries	Outcome	240	125
	Probability	240	125
Compound-2x	Outcome (outer)	230	125
	Outcome (inner)	230	85
	Probability	220	125

Note: All fixation points in this AOI were used for further analysis.

3.3 Statistical methods

Data were analysed using a combination of statistical techniques for both the lottery evaluation and reasoning tasks. For behavioural data, certainty equivalents of compound lotteries, their actuarially equivalent lotteries, and simple lotteries were compared using paired t-tests and Wilcoxon signed-rank tests. For the eye-tracking data, we assessed first fixations, the sequence of fixations, and the time to first fixation (TTFF) for each AOI. Each trial was divided into 10 intervals, and the proportion of time spent on each AOI was calculated for each interval. Wilcoxon signed-rank tests were used to compare the order of fixations. At the same time, t-tests were applied to compare other metrics, including time spent on various elements of the lotteries and reasoning task components.

Additionally, we analysed the number of switches between pattern and option tiles in the reasoning task, comparing easy and hard trials using t-tests. Ordinary least squares (OLS) regression analyses were conducted to explore whether fixation patterns predicted deviations in certainty equivalents from expected values and influenced accuracy in the lottery task. The study was approved by Ashoka IRB (23-X-10044-Hajimoladarvish).

4. Data and Results

Seventy-three participants (32 female, 44%) between 18.64 and 24.98 years (M = 20.44 years) participated. Data collection took place between September 13 and 19, 2023. The sessions were conducted in the psychology lab one at a time. Recruitment of participants occurred through the SONA system. All participants are undergraduates from an Indian University, except three PhD students. A sample size of 70 was determined before data collection, as is typical for eye-tracking studies. All participants were right-handed with normal or corrected-to-normal vision and had no history of neurological or psychiatric illness. Participants with glasses were avoided to avoid interference with pupil measurement, but participants with contact lenses were allowed. The study was conducted with approval by the Research Ethics Committee of the Ashoka University on 18 July 2023 with reference number 23-X-10044-PII. All participants gave written informed consent. On average, each session took about 26 minutes, excluding the instruction.

4.1 Behavioural data: Lottery task

The overall average reaction time across all participants and lotteries was 14.26 seconds. There was a significant difference between simple lotteries (11.42) and compound (19.93) lotteries ($t_{72} = 10.46$, $p < 0.001$), with faster responses for simple lotteries. Note that by design, our compound A (AE(A)) lotteries have a higher expected value than compound lottery Bs (AE(B)). This holds for SA and SB lotteries as well. We use the percentage of declared certainty equivalents that follow the dominance ordering as a measure of accuracy. Overall, the mean accuracy across the 27 pairs was 81.33%, with higher accuracy observed in simple lotteries (82.49%) compared to compound lotteries (78.99%).

We calculated the deviation from the expected value averaged across trials for every subject. A t-test across subjects showed this deviation was not different from zero ($t_{72} = 1.44$, $p > 0.15$). This suggests that our subjects are using expected value as a guide to evaluate lotteries. Using the certainty equivalents of two-outcome simple lotteries[6], we also used maximum likelihood to estimate each individual's EU and rank-dependent utility (RDU) parameters. We assume the observed certainty equivalent of each lottery (j) can be written as $ce_j = \overline{ce_j} + \varepsilon$ with $\varepsilon \sim N(0, \sigma)$, and $j \in \{1, \ldots, J\}$ that refers to a specific lottery. Provided that the error terms ε are statistically independent of one another, then the joint density of ε which is the product of all densities, which is the likelihood function.

$$f(\varepsilon|\theta) = \prod\nolimits_{j=1}^{J} f(\varepsilon|\theta) = \prod\nolimits_{j=1}^{J} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{\left(ce_j - \overline{ce_j}\right)^2}{\sigma}}$$

The parameters of interest (θ) and the information about them are contained in observed data ce_j. Hence, we look for the parameters that given our observed data and a model of interest ($\overline{ce_j}$) is most likely to have produced the data. For utility function specification, we chose the power function $u(x) = x^{\overline{\alpha}}$ with α>0. This is suggested to be the best fit for experimental data (Stott, 2006; Wakker, 2008). For the probability weighting function, we used the two-parameter specification by Goldstein and Einhorn (1987) given by $w(p) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}}$, with γ>0, δ>0. We favour this specification because it has been shown to be a good fit for experimental data and is a common choice in the literature (Gonzalez and Wu, 1999). We then conducted log-likelihood ratio tests to determine the better fit between EU and RDU. Interestingly, all our subjects are classified as EU maximisers.[7]

[6]These include all SA and SB lotteries, along with three lotteries each from AE(A) and AE(B), for a total of 24 lotteries per participant.

[7]Assuming EU preferences, for the pooled sample, the estimated utility parameter is 1.05 with a standard error of 0.0238. Thus, we cannot reject the null hypothesis that subjects are following the expected value criterion. Assuming RDU preferences, the estimated utility parameter is 0.9819, and the parameters of the probability weighting function are 0.9407 and 1.0452, respectively — all of which are close to 1, suggesting limited distortion in probability perception.

4.1.1 Test of the ROCL axiom

We examine two hypotheses central to the ROCL axiom. The first, the consistency hypothesis, predicts that individuals will maintain their preferences when a compound lottery is replaced with its actuarially equivalent simple lottery. ROCL is considered violated if there is a systematic discrepancy in the proportion of cases where the elicited value of the simple lottery (S) exceeds that of the compound lottery (C) in S–C pairs, compared to the proportion of cases where the value of S exceeds that of actuarially equivalent simple lottery (AE) corresponding to the compound lottery in S–AE pairs. As participants did not make direct choices between these lotteries, this is a constructed test based on elicited valuations rather than observed choices. The second hypothesis concerns indifference, suggesting that individuals should not systematically prefer a compound lottery and its actuarially equivalent simple lottery.

We use the Cochran's Q test, coupled with the Bonferroni-Dunn (B-D) correction procedure, to evaluate the consistency prediction of the ROCL axiom. Specifically, we test the hypothesis that the proportion of cases in which the elicited value of a compound lottery exceeds that of a simple lottery is equal to the proportion in which the compound lottery exceeds its actuarially equivalent simple lottery. The analysis includes nine constructed choices involving the pairs [A-$AE(B)$ vs. $AE(A)$-$AE(B)$] and nine pairs involving [B-$AE(A)$ vs. $AE(B)$-$AE(A)$]. We aim to determine whether subjects assign a higher value to a compound lottery (A/B) than to a simple lottery in the same proportion as they value its actuarially equivalent simple lottery ($AE(A)/AE(B)$) relative to the same simple lottery. The simple lottery used in these pairs is the actuarially equivalent lottery corresponding to the other compound lotteries, as compound lotteries A and B are comparable.

We do not find evidence to reject the ROCL axiom consistency prediction. Table 3 shows the results of the B-D method for each of the eighteen comparisons. Table 3 provides evidence that with a 5% familywise error rate, subjects evaluated compound lotteries higher than simple lotteries in the same proportion as actuarially equivalent lotteries over simple lotteries. This implies that our data support the ROCL consistency prediction.

Table 3. The B-D procedure on linked certainty equivalents of eighteen pairs.

Pairs	Proportion indicating higher CE for compound lottery as compared to Simple ones (π_1)	Proportion indicating higher CE for the actuarially equivalent lottery as compared to Simple ones (π_2)	$\lvert\pi_1 - \pi_2\rvert$
1	0.56	0.51	0.05
2	0.71	0.64	0.07
3	0.14	0.05	0.08
4	0.33	0.34	0.01
5	0.68	0.56	0.12
6	0.49	0.33	0.16
7	0.52	0.33	0.19
8	0.71	0.66	0.05
9	0.32	0.26	0.05
10	0.26	0.14	0.12
11	0.23	0.08	0.15
12	0.18	0.37	0.19
13	0.48	0.41	0.07
14	0.05	0.05	0.00
15	0.38	0.16	0.22
16	0.45	0.29	0.16
17	0.11	0.12	0.01
18	0.23	0.25	0.01

Note: The test rejects the null hypothesis of $\pi_1 = \pi_2$ if $\lvert\pi_1 - \pi_2\rvert > d$. Calculating the critical value d requires first defining *ex-ante* a familywise Type I error rate (α_{fw}). For $\alpha_{fw} = 0.05$ the corresponding critical value is 0.1927.

With the pooled data, 78% of choices are consistent with the ROCL. Almost all subjects (71/73) made more than 50% (> 9/18) ROCL consistent choices. Furthermore, no order effect ($t_{69} = 0.259$, $p > 0.796$) or gender effect ($t_{72} = -0.341$, $p > 0.734$) was observed in the number of choices consistent with the ROCL.

For the second hypothesis concerning indifference, we test if the elicited certainty equivalent of compound lotteries and their associated actuarially equivalent lotteries are equal on average. We have eighteen pairs for this comparison. The paired t-test result indicates a significant difference between the mean of elicited certainty equivalents ($t_{1313} = 3.8084$, $p < 0.001$). Wilcox signed rank test also rejected the null hypothesis that the median of pairwise differences in elicited certainty equivalents is zero ($p < 0.001$). Comparable results were obtained when averaging across 18 choices for each subject and conducting a paired t-test ($t_{72} = 3.1735$, $p < 0.003$). On average, subjects evaluated compound lotteries slightly higher than their associated actuarially equivalent simple lotteries. The mean of elicited certainty equivalents for compound lotteries is 2457, while the mean of actuarially equivalent lotteries is 2318. This is consistent with the findings of Hajimoladarvish (2018), Harrison et al. (2015) and Bar-Hillel (1973). Moreover, 38% of the pairs show equal certainty equivalents for compound and actuarially equivalent lotteries. While 35% show higher values for compound lotteries than actuarially equivalent lotteries, 27% show lower values for compound lotteries.

Kaivanto and Kroll (2012) find that the certainty equivalent of the compound lotteries is significantly smaller than the equivalent reduced form (AE) in St. Petersburg gambles. Bansal and Rosokha (2018) find that subjects' perceived value of projects is higher when projects are described in the reduced version (AE) than the compound version. Hence, their data suggest a bias towards a simple reduced form.

These mixed results can be due to heterogeneity among subjects. Inspecting individual data, we find that only for 13 subjects, there is a significant difference between the certainty equivalent of compound lotteries and their associated actuarially equivalent lotteries. Only 5 subjects are biased towards compound lotteries, while 3 subjects indicate a bias towards actuarially equivalent lotteries.[8] Hence, the finding of bias towards the compound or reduced lottery is a confound associated with reporting average results.

4.1.2 Test of the CI axiom

We use the Cochran Q test coupled with the Bonferroni-Dunn (B-D) correction procedure to test the consistency prediction of the CI axiom. We test the hypothesis that subjects evaluate a lottery *A* higher than a lottery *B* in the same proportion as the distinguishing simple *SA* lotteries over *SB* lotteries. Hence, we are testing if the proportion of subjects indicating higher certainty equivalents for lottery *A*s as compared to lottery *B*s is the same as the proportion of subjects indicating higher certainty equivalents for *SA* lotteries as compared to *SB* lotteries.

We do not find evidence to reject the CI axiom consistency prediction. Table 4 shows the results of the B-D method for each of the 9 comparisons. Table 4 provides evidence that with a 5% familywise error rate, subjects evaluated the *A* lotteries higher than *B*s in the same proportion as they evaluated the simple *SA* lotteries higher than *SBs*. This implies that our data support the CI consistency prediction. 74% of choices are consistent with the CI axiom. Most subjects (64/73) made more than 50% (> 4/9) CI consistent choices. Additionally, we found no evidence of an order effect ($t_{72} = -0.251$, $p > 0.802$) or gender effect ($t_{72} = -0.175$, $p > 0.861$) influencing the frequency of choices consistent with the CI.

4.2 Eye-tracking data

4.3.1 Lottery task

The eye-tracker was used to estimate gaze location, time spent on each area, and the order in which subjects looked at them for each lottery trial. We consider the fixation order across all probabilities and outcomes to be the order of information acquisition. This is consistent with using all lottery characteristics for valuation, supported by Rosen and Rosenkoetter (1976), Glöckner and Herbold (2011), and Fiedler and Glöckner (2012).

Figure 4A illustrates a single trial of a compound lottery task. Here, we see that the fixation points are concentrated around numbers, with a few in the region below (the keyboard) where they enter the value. We see that each number is fixated upon multiple times. The colour scheme (the redder the colour, the later in the order) suggests the original probabilities were looked at before the outcomes. As expected, the last fixations are related to entering the answer. The three compound

[8]A bias towards compound lotteries is established if subjects evaluated compound lotteries higher than the actuarially equivalent ones in more than 9 occasions out of 18.

Table 4. The B-D procedure on linked certainty equivalents of *A-B* lottery pairs and *SA-SBs*.

Pairs	Proportion indicating higher CE for *A* as compared to *B* (π_1)	Proportion indicating higher CE for *SA* as compared to *SB* (π_2)	$\lvert\pi_1 - \pi_2\rvert$
1	0.77	0.89	0.12
2	0.77	0.92	0.15
3	0.81	0.89	0.08
4	0.56	0.62	0.05
5	0.97	0.93	0.04
6	0.74	0.90	0.16
7	0.77	0.76	0.00
8	0.90	0.86	0.04
9	0.82	0.90	0.08

Note: The test rejects the null hypothesis of $\pi_1 = \pi_2$ if $|\pi_1 - \pi_2| > d$. Calculating the critical value *d* requires first defining *ex-ante* a familywise Type I error rate (α_{fw}). For $\alpha_{fw} = 0.05$ the corresponding critical value is 0.1140.

lottery types are illustrated as Compound-2x and Compound-up in Figure 3A and 3B, respectively, and Compound-down in Figure S2B.

Each participant completed six trials for the three lottery types, compound-up, compound-down and compound-2x. We determined the first time each outcome or probability was fixated upon for each lottery type, using 7AOI in single compound lotteries (3 outcomes and 4 probabilities for compound-up and down) and 10 (4 outcomes and 6 probabilities) for the compound-2x lotteries. We then identified each participant's most common first fixations by taking the mode across their trials. A first fixation refers to the *initial point where the eye lands* and briefly pauses after a visual stimulus appears. Studies consistently show that these first fixations tend to occur near the geometric centre or top-centre of the screen or visual scene (Tatler, 2007), as seen in our data.

The most frequent initial fixations in compound-up lotteries showed 39 participants (53.42%) first fixated on the upper-arm-2 probability; in compound-down lotteries, 32 participants (43.84%) initially fixated on the lower-arm-1 probability; and in compound-2x lotteries, 30 participants (36.99%) first fixated on the lower-arm-1 outcome. Because the mode did not capture all trial-to-trial variation, the fraction of trials each AOI was fixated on first, second, and so forth was also calculated. Tables S3 and S4 present the percentage of trials in which each outcome and probability appeared in each fixation position, averaged across participants. These values reflected only the first time participants fixated on a given AOI (i.e., subsequent returns were ignored). The primary objective was to determine whether participants tended to fixate on second-stage outcomes before or after examining the original probability, and the data suggest it is the latter.

Next, we compare the order of fixations using Wilcoxon tests. In compound-up trials on average, the upper original probability appeared more frequently among the first three fixation positions than the upper outcomes, which tended to be fixated on in the fourth through seventh positions. A Wilcoxon test across participants indicated that the upper original probability was ranked first and second more often than the upper-arm-1 and upper-arm-2 outcomes ($z > 4$, $p < 0.001$). In the third position, it was fixated on more often than the upper-arm-2 outcome ($z = 3.77$, $p < 0.001$), but no significant difference emerged for the fourth position. In the fifth, sixth, and seventh positions, this pattern reversed. The upper original probability was fixated on less frequently than both upper-arm-1 and upper-arm-2 outcomes ($z < -3.54$, $p < 0.001$). These findings suggested that participants generally examined the original probability before viewing the second-stage outcomes.

Similarly, in compound-low lotteries, Wilcoxon tests showed that the lower original probability was ranked in the first four positions more often than the lower-arm-1 and lower-arm-2 outcomes ($z > 2.42$, $p < 0.05$). The one exception occurred in the first position, where participants fixated on the lower-arm-1 outcome more frequently (z = -4.76, $p < 0.001$). The lower original probability was ranked fifth less often than the lower-arm-1 outcome ($z = -3.23$, $p < 0.01$) and was ranked sixth and seventh less often than the lower-arm-2 outcome ($z < -5.34$, $p < 0.001$). All other comparisons were not significant.

In compound-2x lotteries, the two second-stage lotteries were analysed separately, and each set of second-stage outcomes was compared to its corresponding original probability. The upper original probability was ranked first and second more often than the upper-arm-1 outcome ($z > 4.23$, $p < 0.001$), but did not differ significantly from the upper-arm-2 outcome. It was also ranked third and fourth more often than both upper-arm-1 and upper-arm-2 outcomes ($z > 2.16$, $p < 0.05$). However, from the sixth to tenth positions, the upper original probability was fixated on less frequently than both outcomes ($z < -2.91$, $p < 0.05$), except that it did not differ from the upper-arm-2 outcome in the ninth position. For the lower original probability, it was ranked first less often than the lower-arm-1 outcome ($z = -5.14$, $p < 0.001$). By contrast, it was ranked more often than the lower-arm-1 outcome in the third and fifth positions ($z > 3.74$, $p < 0.05$) and was fixated on less frequently than the lower-arm-1 outcome in the seventh, eighth, and ninth positions ($z < -2.31$, $p < 0.05$). Overall, from ranks one to five, the lower original probability exceeded the lower-arm-1 outcome ($z > 2.86$, $p < 0.01$), whereas from ranks eight to ten, it was lower ($z < -2.34$, $p < 0.05$).

To further distinguish the order of fixations, we divided each trial into ten equal-length time intervals. Within each interval, we calculated the number of fixations on each AOI as a proportion of the interval. For example, if an interval lasted one second (120 time points) and 60 fixation points were recorded in the upper original probability AOI, the measure would be 0.5. This process was repeated for each participant across all 10 time intervals and the relevant AOI in each type of compound lottery. As shown in Figure 5, across all compound trials, the original probability was fixated on more frequently in the earlier intervals of the trial compared to either or both second-stage outcomes. In compound-up lotteries, the proportion of time spent fixating on the upper original probability in the first interval was significantly higher than on the second-stage outcomes ($t_{72} > 4.10$, $p < 0.001$). This pattern persisted for the upper-arm-2 outcome in the second and third intervals ($t_{72} > 2.09$, $p < 0.05$), but no significant differences were found in the remaining intervals. In compound-down lotteries, the proportion of time fixating on the lower original probability was significantly greater than on the lower-arm-2 outcome in intervals 1–4 and 8 ($t_{72} > 2.12$, $p < 0.05$), while all other comparisons were not statistically significant. For compound-2x lotteries, the proportion of time spent on the upper original probability was greater than on both the upper-arm-1 and upper-arm-2 outcomes in the first and second intervals ($t_{72} > 2.76$, $p < 0.01$), and greater than on the upper-arm-1 outcome only in the 10th interval ($t_{72} = 2.01$, $p = 0.048$). Similarly, the time spent on the lower original probability was significantly higher than on the lower-arm-1 outcome in the first interval and higher than on the lower-arm-2 outcome in intervals 1–4 and 6 ($t_{72} > 2.05$, $p < 0.05$). While in the ninth interval, fixation time on the lower original probability was significantly lower than on the lower-arm-1 outcome ($t_{72} = -2.39$, $p = 0.019$).

Time to First Fixation (TTFF) is the elapsed time (in milliseconds) between the onset of a visual stimulus and the moment the viewer's gaze first lands within a defined AOI. We estimated the time within each trial before participants made their first fixation on each outcome or probability. T-tests across subjects were then conducted to examine differences in these TTFFs, which were averaged over trials. In the compound-up lotteries, participants fixated on the upper original probability (mean = 1.46 s, SD = 1.62) earlier than they fixated on upper-arm-1 (mean = 3.39 s, SD = 2.14) or upper-arm-2 (mean = 4.38 s, SD = 4.23; $t_{72} > 5.65$, $p < 0.001$) outcome. In the compound-down lotteries, the lower original probability (mean = 2.44 s, SD = 2.13) was viewed earlier than the lower-arm-2 outcome (mean = 5.42 s, SD = 3.20; $t_{72} = 6.89$, $p < 0.001$) but was not significantly different in timing from lower-arm-1 outcome (mean = 2.83 s, SD = 1.99). This pattern was replicated in the compound-2x lotteries. Participants first looked at the upper original probability (mean = 1.99 s, SD = 2.30), before upper-arm-1 (mean = 4.30 s, SD = 3.80) and upper-arm-2 outcomes (mean = 3.11 s, SD = 3.13; $t_{72} > 3.04$, $p < 0.01$). Meanwhile, the lower original probability (mean = 3.67 s, SD = 3.36) was viewed before lower-arm-2 (mean = 7.21 s, SD = 4.46; $t_{72} = 6.80$, $p < 0.001$) but did not differ significantly from lower-arm-1 (mean = 4.39 s, SD = 3.73) outcome.

Finally, we examined the order of fixations in each trial. Trials were classified as following a CI strategy if both second-stage outcomes and probabilities were fixated before the original probability. Across all compound lotteries, very few trials met this criterion. In compound-up lotteries, only 37 (8.4%) trials across 20 subjects matched this pattern. This suggests subjects did not consistently use this strategy across all 6 trials. To further examine this, we examined the proportion of trials that showed the CI strategy for each subject. Only 2 among 20 showed this pattern for more than half the trials, 3 for half the trials, and the remaining 15 for less than half the trials. In compound-down, there were 53 (12.1%) such trials across 24 participants, with 5 subjects showing this pattern for more than half the trials, and 19 for less than half the trials. In compound-2x lotteries, there were 33 (7.5%) trials in the upper second-stage lottery across 24 subjects, with 1 subject showing this pattern for more than half the trials, 2 for half the trials, and 21 for less than half the trials. 34 (7.8%) trials in the lower second stage lottery across 20 subjects matched this pattern, with 2 subjects showing the pattern for more than half of the trials, 1 for half the trials, and 17 for less than half the trials. 12 subjects showed this pattern in both upper and lower second-stage lotteries (one subject showed the pattern for five-sixths of the trials in upper second-stage lotteries and all lower second-stage lotteries). In all cases, no more than 5 subjects used it in half or more trials, thus less evidence of a strategy. It is worth noting that in approximately 7.5% trials of compound-up, 10.3% of compound-down, and 9.6% of compound-2x, there was no recorded fixation in the original probability AOI, or one of the second stage values when the original probability was fixated on later, i.e., fifth or later in the order of fixations.

We used regression to see if CI strategies predict the absolute deviation of the values quoted by participants from the expected value of the lotteries. A regression of this difference across 1314 trials (73 subjects × 18 compound trials) with subject fixed effect, reaction time per trial, and CI rate for each type of compound trials was performed. For compound-2x types, we averaged the CI rate across both upper and lower second-stage lotteries. Table S6 in the supplementary section reports the regression result. Across all lottery types, greater use of the CI strategy is associated with lower deviations from expected value, but this relationship is not statistically significant, and the estimates are imprecise. As expected, on average, deviation is almost 160 units higher in compound_2x lotteries than in compound-up lotteries (insignificant), as they are more complicated. Interestingly, deviations are about 184 units higher in compound-down lotteries than in compound-up lotteries ($p < 0.05$). In general, it seems the use of CI cannot explain any deviation from the expected value.

4.3.1.1 Fixation duration and pupil size

Fixation duration and pupil diameter are often analysed in eye-tracking studies. Here we test whether total fixation duration on the original probabilities differs from that on second-stage outcomes. In compound-up and compound-down lotteries, participants spent more time on the original probability than on the second-stage arm-2-outcome ($t_{72} > 3.71$, $p < 0.001$) but not the arm-1-outcome. In the compound-2x lottery, the upper original probability was fixated longer than the upper-arm-1 outcome ($t_{72} = 2.58$, $p = 0.012$), and not the upper-arm-2 outcome; the lower original probability was longer than the lower-arm-2 outcome ($t_{72} = 2.02$, $p = 0.047$), but not the lower-arm-1 outcome. This pattern of equal or greater fixation duration on original probabilities aligns with a forward induction strategy. A comparison of average durations across all outcomes and probabilities reveals no significant differences across any lottery types. Previous studies have reported either greater fixation durations on probabilities or no significant difference in multi-lottery conditions (Alós-Ferrer et al., 2021) and greater fixation durations on outcomes (Alós-Ferrer and Ritschel, 2022). That said, duration is a noisy proxy here. Larger numbers (e.g., 4,000 vs. 1,000 or 90% vs. 10%) and redundancies in the lottery information can inflate viewing time independent of strategy. Participants may also "park" their gaze at central or top-central locations while mentally combining probabilities and outcomes, artificially boosting the recorded duration for whatever AOI lies beneath. These factors caution against over-interpreting absolute duration differences.

Pupil diameter indicates both reward expectation (Fröber et al., 2020) and cognitive load (Krejtz et al., 2018). Here we see that pupil diameter is greater for outcomes than probabilities in all compound lottery types ($t_{72} > 3.38$, $p < 0.01$), suggesting that the outcome is more directly related to the reward expectation.

4.3.1.2 Quality and Robustness checks

We examined the AOIs and trials with no tracking as a quality check for eye-tracking. 8.37% of all AOIs had no fixation points across lotteries and subjects. Note that there is redundant information in the probability AOIs, i.e., in two-outcome lotteries, if one branch is 90%, participants know that the other is also 10%. Similarly, the outcomes and probabilities could be read through peripheral vision, as we used a small set of repeated outcomes and probabilities. 9 (12.3%) subjects had more than 20% AOIs without fixations, and the analyses were repeated, excluding these subjects (See Supplementary section 5). Similarly, the analysis was repeated after removing trials with more than 20% non-tracked timepoints (33% of all trials), due to the pupil not being visible (Nyström & Holmqvist, 2010).

The story is unchanged across the primary analysis and the first set of robustness checks (dropping 9 subjects with >20% untracked trials and then dropping individual trials with >20% missing data). Participants typically inspect the original (first-stage) probability early—both in rank order and TTFF—and very few individual trials follow the CI strategy. Rank patterns and early-interval fixation advantages for the original probability hold throughout, as do TTFF differences. The complete analyses for these two conditions are provided in the Supplementary Section 5.1.

The second set of the robustness checks separates trials at each position (left, right and centre). The position-specific cuts confirm that the first fixations are mostly a *where-is-the-middle?* effect. When an AOI sits in the centre, it wins the very first look (e.g., in the centre trials, 79% of first fixations are directed to the upper-arm-2 probability in compound-up types, and 86% to the upper-arm-1 probability for compound-down kinds). In contrast, original probability dominates only when it sits on the right (74% first looks), and is rarely first on the left (1–2%). These asymmetries also explain why the robustness checks produce small shifts, some Wilcoxon first-rank comparisons flip sign, certain TTFF contrasts lose/gain significance, and a few later time-interval effects drop, without altering the qualitative story that original probability information is accessed early. CI trials remain scarce (8–12% in the main set and even fewer once split), so they do not drive these results. The complete analyses for these three conditions are provided in the Supplementary Section 5.2.

Finally, we did an incentivised version of the lottery task (n = 8) with two display positions, the lottery in the middle, and the lottery flipped on the y-axis in the middle. The qualitative pattern of the main study is reproduced. Original probabilities are accessed early in the sequence (shorter TTFF and elevated fixation shares in the earliest time bins), and CI trials remain exceptionally rare. The distribution of *first fixations* is more diffuse and strongly contingent on the screen position. Disaggregating by position clarifies this. In the “left-to-right” position, original probabilities still enter early but seldom capture the very first glance; in the “mirror” position (right-to-left), they occasionally regain a first-look advantage, yet initial fixations are fragmented across multiple items. These positional contingencies account for minor divergences in rank tests, TTFF comparisons, and interval-level effects, without altering the central conclusion that original probability information is accessed early and CI behaviour is uncommon. The complete analysis of behavioural and eye-tracking data for this follow-up study is in the Supplementary Section 6.

4.3.2 Reasoning task

Overall, the average accuracy was 83.7%. There was a significant difference between easy (91.48%) and hard (76.03%) trials ($t_{72} = 8.61$, $p < 0.001$), with easy trials being more accurate. The overall average reaction time was 18.13 seconds. There was a significant difference between easy (10.16 s) and hard (26.10 s) trials ($t_{72}=16.65$, $p < 0.001$), with easy trials being faster. This is expected in all hard versus easy comparisons across various tasks and modalities (Fedorenko et al., 2013; Shashidhara et al., 2019; Vigneau et al., 2006).

In this task, participants could use two strategies: constructive matching (forward induction), which involves mentally constructing the missing piece of the pattern, or response elimination (backwards induction), where participants systematically compare the options given with the pattern to find one that seems to fit. To test the use of backwards and forward induction strategies in the reasoning task, we examined whether subjects first looked at the option tiles before viewing the pattern tiles.

Figure 4B illustrates a single trial of the hard reasoning task. Here, we see that the fixation points are concentrated around the tiles of the pattern and option, with a few in the region below (the keyboard) where they enter the value. Many tiles are fixated upon multiple times, particularly the first two tiles in the first row and the first tile in the second row (first tile, and next along the row and column). The colour scheme (the redder the colour, the later in the order) suggests the pattern tiles were looked at before the options. As expected, the last fixations are related to entering the answer. Each participant completed 18 trials of each of the easy and hard levels of the reasoning task. We estimate first fixations separately for easy and hard trials. For each trial type, we determined the first time each pattern tile and option tile were fixated upon, using 13 AOIs (including the missing tile), i.e. 9 pattern tiles and 4 option tiles. Positions are named as follows. 1st position is (row 1 column 1), 2nd position is (row 1, column 2) until position 9, the missing tile, is (row 3 column 3). The option tiles are 10-13 from left to right. We then identified each participant’s most common first fixation by taking the mode across their trials.

For easy trials, 53 participants (72.6%) first fixated on the 8th tile in the pattern (row 3, column 2), roughly the centre of the screen, 7 (9.59%) each on the 4th and 5th tile, 5 (6.85%) on the first tile and 1 (1.37%) on the 7th tile. For hard trials, 47 (64.38%) first fixated on the 8th tile, 11 (15.07%) on the 4th tile, 7 (9.59%) on the 5th tile, 5 (6.85%) on the first tile, and 3 (4.11%) on the second tile. In both cases, the first fixation was on only pattern tiles and was mainly driven by what was at the centre of the screen, and the pattern was similar between hard and easy trials.

Because the mode did not capture all trial-to-trial variation, the fraction of trials each AOI was fixated on first, second, and so forth was also calculated. Table S5 presents the percentage of trials where each tile appeared in each fixation position, averaged across participants. These values reflected only the first time participants fixated on a given AIO (i.e., subsequent returns were ignored). The primary objective was to determine whether participants tended to fixate on options before or after examining the pattern, and the data suggest it is the latter.

In both easy and hard reasoning trials, on average, the pattern was looked at more frequently among the first 4-5 fixation positions than the answer options, which tended to be fixated on in the 9 through 13 positions. Note, given the easy nature of the task, subjects need not see each of the pattern tiles to figure out the answer; just one row and one column is a sufficient condition to solve the problem. Similarly, the subjects need not see all the option tiles in every trial, depending on when they find the answer. We compared fixation-order ranks with Wilcoxon signed-rank tests computed per participant on tile-pairs. Specifically, for each trial type (18 easy, 18 hard; 73 participants), we tested whether any option tile (4 total) was fixated before any of the five pattern tiles that make up the first row and first column of the 3×3 pattern grid (4×5 = 20 comparisons per Easy and Hard trial types; 40 total). Pattern tiles were more often in ranks 1–4 than the options ($z > 2.31$, $p < .05$); the same held at rank 5 in 34 of the 40 comparisons. One pattern tile (position 3) was dropped

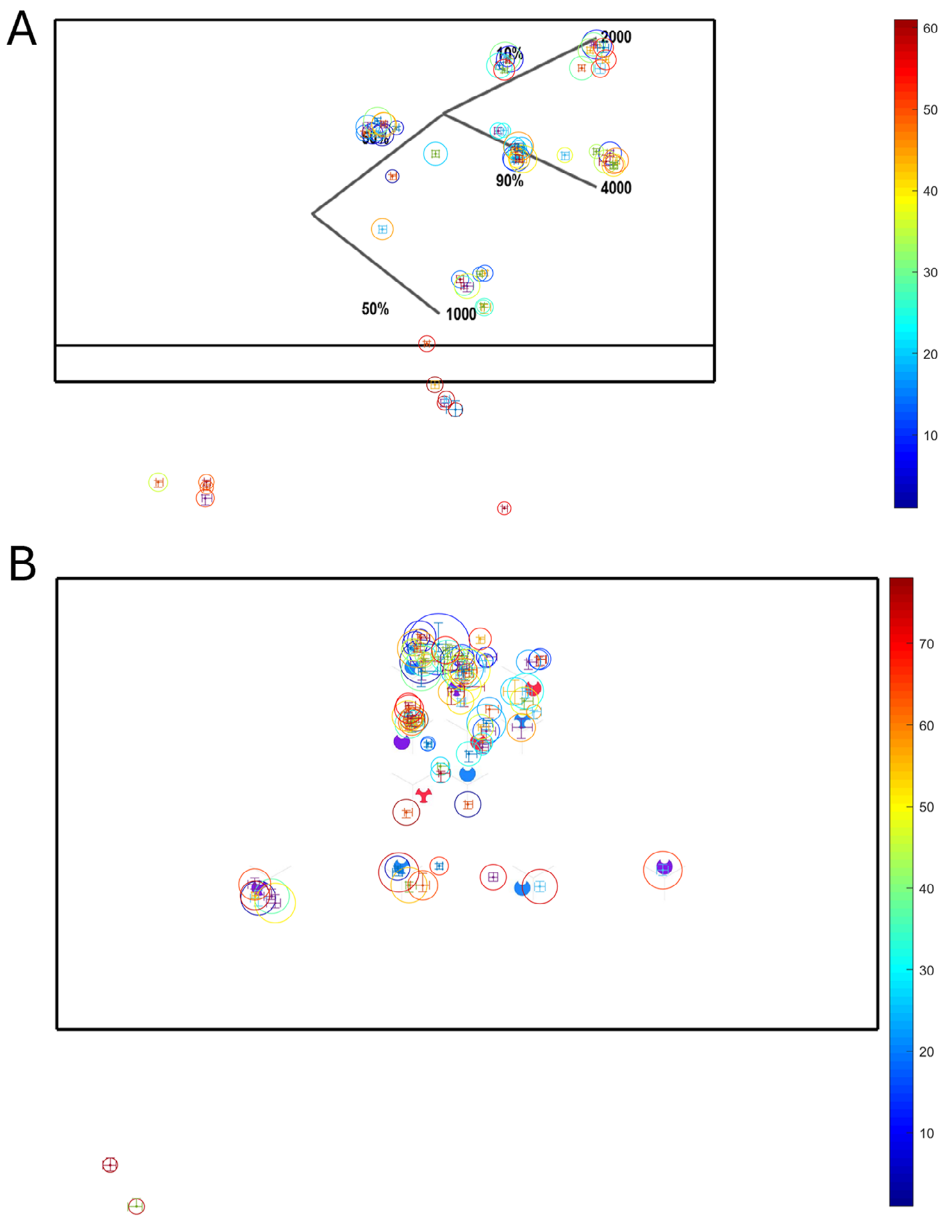

Figure 4. Fixation areas in both tasks. A. This panel shows the average fixation points of a subject in one trial of the compound lottery task (compound-up). B. This panel shows the average fixation points of a subject in one hard trial of the reasoning task. The size of the points indicates the time spent. The error bars show the mean and standard deviation of movement during each fixation event. The colour scheme indicates the fixation order, with blue being the first and red being the last. The trials shown here are the same as those in Figure 3.

from the early-rank report because too few observations were available to estimate a z-value. The reverse pattern emerged later. In 34 or more of the 40 comparisons, option tiles appeared more frequently than the first row and column tiles in ranks 9–13 ($z > 1.99$, $p < .05$).

To further distinguish the order of fixations, we divided each trial into ten equal-length time intervals. Within each interval, we calculated the number of fixations on each AOI as a proportion of the interval. This process was repeated for each participant across all 10 time intervals and the relevant AOI in the two types of trials. As depicted in Figure 6, across all trials, the first row and column of the pattern were fixated on more frequently in the earlier intervals of the trial compared to the option tiles. The proportion of time spent fixating on the first row and column of the pattern in the first seven intervals (in 137 out of 140 comparisons) was significantly higher than on the options ($t_{72} > 2.38$, $p < 0.05$). In

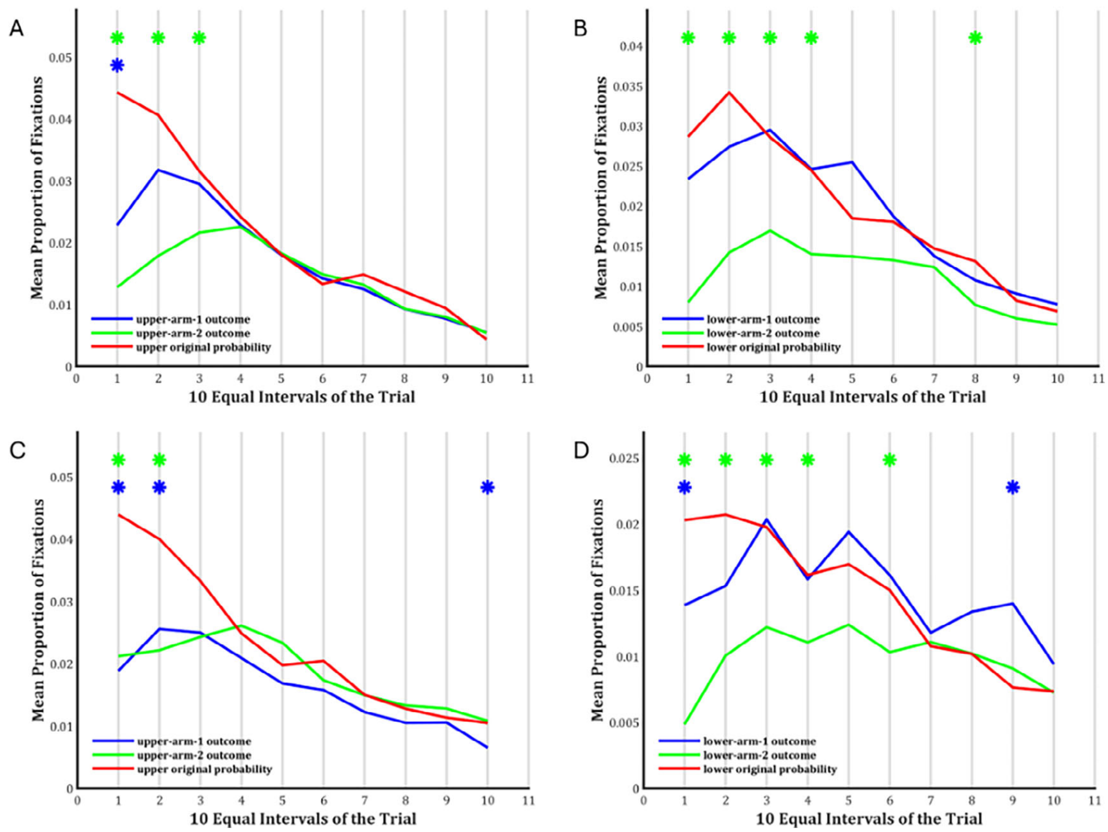

Figure 5. The trend of fixations for the second-stage outcomes, and the corresponding original probability across the trial split into 10 intervals. The proportion of the interval with fixation on an AOI is calculated per subject and then averaged. The star represents a significant difference between second-stage outcomes (stars are colour-coded to indicate the outcome) and the original probability. A. Across compound-up lotteries. B. Across Compound-low lotteries. C. Across compound-2x upper second-stage lottery. D. Across compound-2x lower second-stage lottery.

interval 10, there were more comparisons (29 out of 40) where options were fixated on more than the first row and column of the pattern ($t_{72} > 1.46$, $p < 0.05$).

We estimated the TTFF on each pattern tile and option trials. T-tests across subjects were then conducted to examine differences between these times to first fixation, which was averaged over trials. In both easy and hard trials, participants fixated on the pattern tiles (here we use all the 8 pattern tiles, the missing tile is not included) before the option tiles ($t_{72} > 4.36$, $p < 0.001$).

The proportion of the trial time that elapsed before any option was fixated on (time spent purely on the pattern) was 0.436 (std = 0.096) for easy trials and 0.318 (0.123) for hard trials; with this value being greater in easy trials than hard ($t_{72} = 9.55$, $p < 0.001$). The higher proportion of trial time spent on the pattern before the first toggle to the options and the higher average proportion of time spent on the pattern indicate the forward induction strategy (Vigneou & Bors, 2006). While they saw that the greater the difficulty, the longer the subjects examined the pattern before toggling down to inspect the response choices, we see the opposite pattern. The fact that our trials had a clear distinction between easy and hard trials might have led participants to look at the options in the hard trials. On average, participants spent 0.850 (std = 0.62) of the total fixation time across the tiles of interest on the pattern in easy trials and 0.847 (std = 0.62) in hard trials, with no significant difference. This large proportion of fixation time on the pattern again indicates a forward induction strategy.

We also examined the pupil diameter. The diameter is larger when fixated on pattern tiles than on option tiles, in both easy ($t_{72} = 2.23$, $p < 0.029$) and hard trials ($t_{72} = 6.41$, $p < 0.001$), with the estimate larger in hard trials. This suggests greater cognitive processing when looking at the pattern to decipher the rule. No significant difference exists between the average diameter of easy and hard trials across subjects. While there is a load difference between easy and hard trials, the average across the trials may not be sensitive enough to pick up the difference (Hayes and Petrov, 2016).

Lastly, we look at the number of switches or jumps between the pattern and option tiles. Many jumps or toggles point towards response elimination/backwards induction strategy, and fewer suggest constructive matching/forward induction strategy (Bethell-Fox et al., 1984; Vigneou & Bors, 2006). The easy trials' mean number of switches equals 4.05 (SD = 1.56). That is, subjects, on average, look at the option, go back to the pattern, and come back to the option and so on four times in a given easy trial. As expected, this number is higher in the hard trials, with a mean number of switches equal to 9 (SD = 3.50), given the longer trial time. When divided by the trial time, we see a difference in the opposite direction, more jumps in the easy versus hard (t_{72} = 5.50, p < 0.001). The correlation of switches in easy and hard trials across

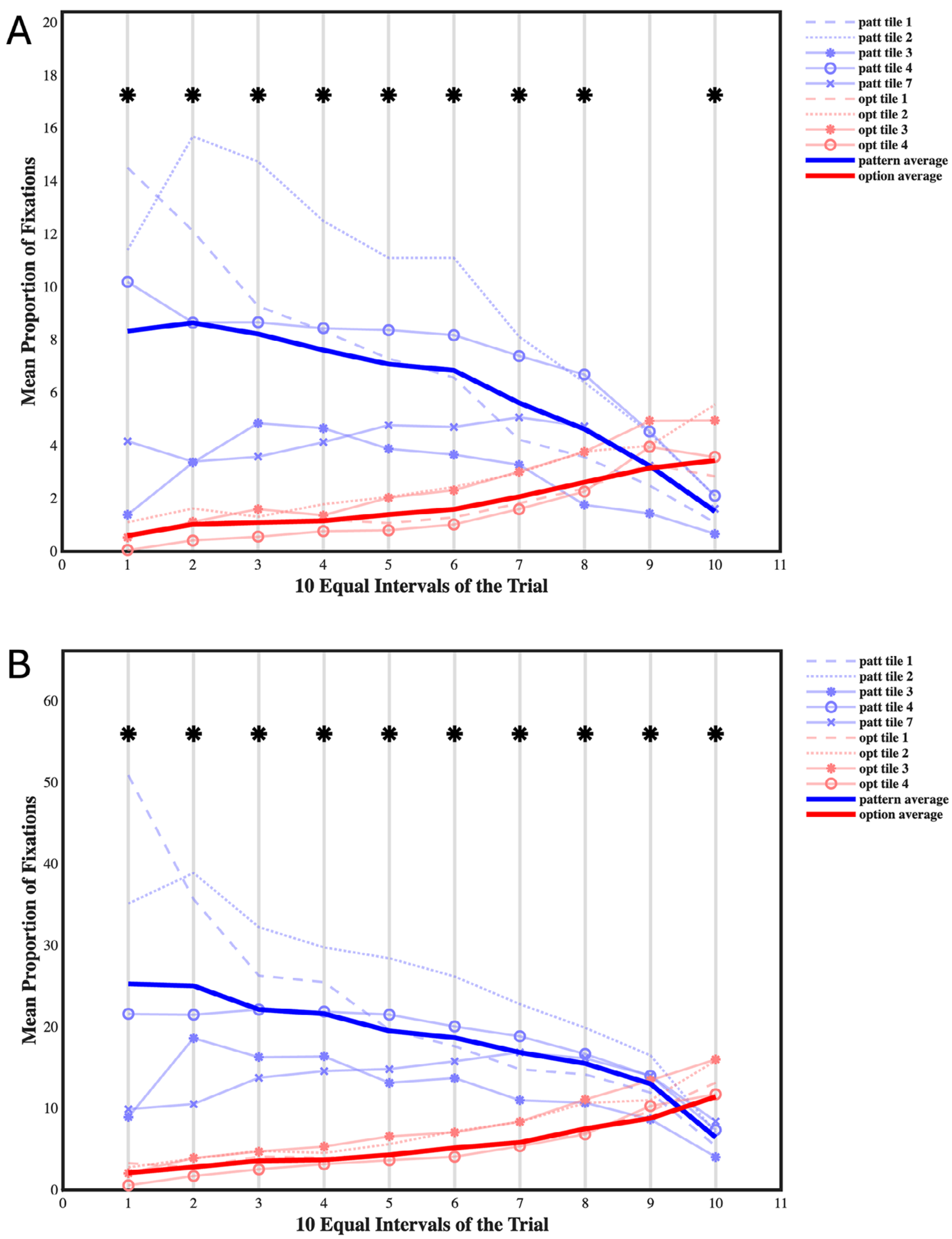

Figure 6. The trend of fixations for the pattern (first column and row), and the option tiles across the trial split into 10 intervals (blue = pattern tiles, red = hard tiles). A. Across easy trials. B. Across hard trials. The proportion of the interval with fixation on an AOI is calculated per subject and then averaged. The pattern average is calculated by averaging the tiles in the first row and first column. The option average is computed by averaging all four options. The star represents a significant difference between the average pattern (first row and column) and the options (four options).

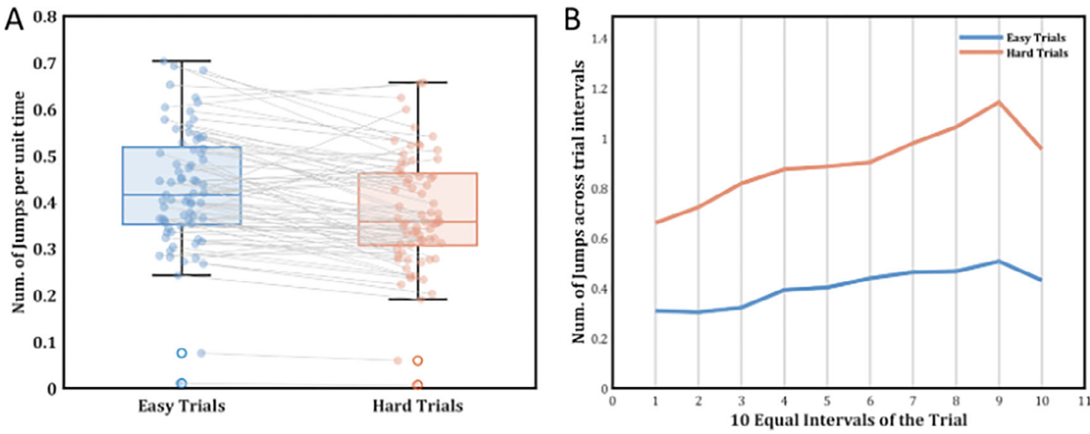

Figure 7. Shows the jumps across hard and easy trials (blue = easy trials, red = hard trials). A. The average number of jumps per unit time in easy and hard trials shows a high correlation across participants and a significant decrease in jumps in the hard trials. B. The trend of the number of jumps between the pattern and the option tiles across the trial is split into 10 intervals.

subjects is 0.72 ($p < 0.001$). The correlation increased to 0.90 ($p < 0.001$) when the jumps are adjusted for reaction time (Figure 7A), with very few subjects showing a large number of switches and no outliers. Figure 7B shows the progression of switches across trials for easy and hard trials, which follow the same pattern, with switches increasing in the second half of the trial and decreasing in the last interval where participants focus on response selection. While the initial toggle may have been slightly earlier in the complex task (perhaps to confirm the trial is the harder type), the number of toggles in hard trials is less, suggesting a more constructive matching strategy. While other studies measured strategy variation with IQ and difficulty and ensured a wide range of IQ among participants (Vigneou & Bors, 2006), we could not get an independent measure of IQ due to an already taxing session. Given that our subject pool consisted of university students, we may have an average IQ higher than that of a typical sample.

As a quality check for eyetracking, we examined AOIs and trials without tracking. 23.19% of all AOIs (13 per trial) had no fixation points across all trials and all subjects. Here, the pattern was small, and the distance between the tiles was negligible. Neighbouring tiles could be seen when fixation on one tile. They need not see all the tiles in simple trials to infer the pattern. As only 10% of trials had more than 20% non-tracked timepoints, no additional robustness checks were done due to the pupil not being visible.

5. Discussion and Conclusion

Our study looks into the evaluation of compound lotteries, utilising eye-tracking to observe the sequence of gaze patterns when assessing these lotteries. We use both eye-tracking and behavioural data to distinguish between two methods of reducing compound lotteries, ROCL and CI axioms that are akin to forward induction and backward induction approaches to problem-solving, respectively. We used a novel test for behavioural data that combines methodologies used for choice and valuation data. We find support for both ROCL and CI, confirming the limitations of traditional behavioural measures in distinguishing between these axioms. We overcome this limitation by leveraging eye-tracking and providing stronger evidence in favour of the ROCL axiom. Across all three compound lottery types, participants accessed the original (first-stage) probability early. It appeared in the top fixation ranks, attracted more looks in the first time bins, and was reached sooner in TTFF than most second-stage outcomes. Wilcoxon tests showed this early-rank advantage reverses in later ranks. TTFF t-tests confirmed earlier access for the original probability in all three lottery types, with a few nonsignificant comparisons. The use of CI strategies was rare (8–12% of trials) and scattered across subjects, indicating that CI was not the dominant strategy even among a subset of participants. The fixation duration of original probabilities was equal to or greater than that of most comparison outcomes, but we treated fixation duration cautiously, given numeric and positional confounds.

In a second task completed by the same subjects, we examined the strategies used in a general problem-solving context, specifically, a pattern completion task. Accuracy was higher and faster on easy trials than on hard trials. Eye movements converged on a constructive matching/forward induction strategy. Pattern tiles, especially the central ones, captured the first fixations, dominated early ranks and intervals, and were reached earlier in TTFF than options. Participants spent a substantial share of trial time on the pattern before ever viewing an option, and switch rates (adjusted for trial length) were lower in hard trials, again consistent with constructive matching rather than response elimination. Pupil diameter was larger on pattern tiles than on options, suggesting greater cognitive effort during rule inference.

Our findings are based on hypothetical scenarios[9], which naturally raises concerns about the validity of the results, particularly regarding the elicitation of true preferences in the absence of monetary incentives. The primary concern is that, without incentives, participants may not reveal their true preferences. However, research across various domains suggests that hypothetical and incentivised responses often do not differ significantly (see Etchart-Vincent and l'Haridon, 2011; Matousek et al., 2022). Moreover, while it is commonly argued that real incentives reduce response variance and thus enhance estimation precision, our data show no significant deviation between subjects' valuations and the expected value of lotteries on average. This suggests high-quality responses and indicates that participants were likely motivated to express their true preferences, even in a hypothetical setting. Nevertheless, incentives, historically the main distinction between experimental economics and other disciplines, continue to be viewed as important. To address this, we conducted a follow-up experiment with eight participants using real incentives (see Supplementary Section 6). The results largely replicated the main patterns observed in the primary, non-incentivised sample.

Across the robustness checks in the lottery task, data-cleaning steps (e.g., excluding poorly tracked participants/trials and re-estimating by display position) altered precision and significance levels in predictable ways but did not change the qualitative interpretation. The central findings, early access to original probabilities and the rarity of CI trials remained stable. Thus, cleaning procedures affected magnitude, not direction, reinforcing that the observed patterns are not artefacts of tracking loss or layout configuration.

Our findings could have been confounded by the decision tree representation of lotteries that promotes the multiplication of probabilities for reducing compound lotteries. This is because the probabilities are always to the left of outcomes in our lotteries, and subjects viewing the screen from left to right is a strong finding that suggests the ROCL axiom might have been used more because of the lottery representation. For example, the findings of Zhang et al. (2024) show that participants more frequently choose the lottery with higher expected monetary value in difficult choices when the payoff information is presented horizontally but not vertically, while the findings of Segovia et al. (2022) show that elicited parameters for risk preferences do not vary with presentation formats.

We randomly displayed lotteries in one of the three positions to mitigate the issue of outcomes being to the right. In the left position, the outcomes were displayed near the centre of the screen, and therefore, they were more easily fixated on before the probabilities. However, in this subset, we also observe the same pattern of forward induction. First fixations were overwhelmingly positional. Items placed in the middle captured the very first glance. By cycling the critical elements through left, middle, and right locations, we could separate spatial salience from informational priority, showing that the pooled "original probability is first" effect is largely a consequence of where it sat on the screen, not a pure strategic choice. This manipulation clarifies that strategic access to key information emerges after (or alongside) an initial, layout-driven orienting bias. To rule out the interpretation that our results are driven by a Western reading bias, in our small follow-up experiment, decision trees were randomly flipped along the y-axis, so that the starting node appeared on the right side of the screen in half of the trials. The data and results from this follow-up closely match those of the main experiment.

We introduced a second task from a different domain that involved matrix analogies with patterns rather than numbers, aiming to investigate problem-solving strategies more broadly while avoiding confounds related to lottery formatting and left-to-right numerical processing. Both tasks had two difficulty levels (simple versus compound lotteries; easy versus hard analogies), and in both, we validated difficulty with longer reaction times and lower accuracies in the harder condition. While the tasks are not directly comparable in content, they are comparable in strategic logic. In the lottery task, participants predominantly proceeded "forward" from the original probability to the second-stage outcomes (consistent with the ROCL axiom), and CI trials that solved the second stage first and only then consulted the original probability were rare. In the analogy task, participants likewise adopted a forward-induction/constructive matching approach, inferring the pattern before turning to the options, rather than an option-first elimination strategy. Across numeric (lotteries) and non-numeric (analogies) domains, initial fixations are primarily driven by spatial layout (centre bias in both tasks), not by strategy per se. Nevertheless, sequences typically move "forward" through the task structure after the initial orienting glance. The lottery task involves examining the original probability information, followed by the second stage, with CI-type sequences occurring rarely. In the analogy task, participants tend to infer the pattern first before evaluating the answer options, rather than beginning with the answer set. Thus, the strategic progression is forward, but the very first look is often positional rather than principled.

Understanding how individuals evaluate compound lotteries has broad implications for economic theory, behavioural economics, and policy design. By investigating how people make decisions involving risk, uncertainty and complexity, we can develop more accurate models of human behaviour and improve decision-making frameworks in both experimental and real-world settings.

In conclusion, our study provides insights into the mechanisms used when evaluating compound lotteries through eye-tracking. Our investigation into gaze patterns during lottery evaluation answers an open question in the literature that aims to distinguish between the ROCL and CI axioms through a given test. We find compelling evidence supporting the ROCL axiom, indicating a forward induction evaluation strategy. This pattern holds true across various types of compound lotteries and also in a reasoning task considered in our study. Thus, our subjects use forward induction strategies in uncertain and complex situations used herein. By integrating behavioural data with eye-tracking, we offer a new lens through which to examine decision strategies, paving the way for further research that combines these methodologies to deepen our understanding of human decision-making.

Data availability statement

All data and codes are available on OSF: https://osf.io/z7s5j/?view_only=8a0188951d734430bdf64eb052aa6524. Refer to the supplementary folder in the same repository for supplementary tables, figures, and text.

References

Abdellaoui M, Klibanoff P, Placido L: **Experiments on compound risk in relation to simple risk and to ambiguity.** *Manag. Sci.* 2015; **61**(6): 1306–1322.
Publisher Full Text

Alós-Ferrer C, Ritschel A: **Attention and salience in preference reversals.** *Exp. Econ.* 2022; **25**(3): 1024–1051.
Publisher Full Text

Alós-Ferrer C, Jaudas A, Ritschel A: **Attentional shifts and preference reversals: An eye-tracking study.** *Judgm. Decis. Mak.* 2021; **16**(1): 57–93.
Publisher Full Text

Arieli A, Ben-Ami Y, Rubinstein A: **Tracking decision makers under uncertainty.** *Am. Econ. J.: Microecon.* 2011; **3**(4): 68–76.
Publisher Full Text

Bansal S, Rosokha Y: **Impact of compound and reduced specification on valuation of projects with multiple risks.** *Decis. Anal.* 2018; **15**(1): 27–46.
Publisher Full Text

Bar-Hillel M: **On the subjective probability of compound events.** *Organ. Behav. Hum. Perform.* 1973; **9**(3): 396–406.
Publisher Full Text

Becker G, DeGroot M, Marschak J: **Measuring utility by a single-response sequential method.** *Behav. Sci.* 1964; **9**: 226–232.
PubMed Abstract | Publisher Full Text

Bethell-Fox CE, Lohman DF, Snow RE: **Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance.** *Intelligence.* 1984; **8**(3): 205–238.
Publisher Full Text

Bernasconi M: **Nonlinear preferences and two-stage lotteries: Theories and evidence.** *Econ. J.* 1994; **104**(422): 54–70.
Publisher Full Text

Bernasconi M, Bernhofer J: **Catch me if you can: Testing the reduction of compound lotteries axiom in a tax compliance experiment.** *J. Behav. Exp. Econ.* 2020; **84**: 101479.
Publisher Full Text

Bernasconi M, Loomes G: **Failures of the reduction principle in an Ellsberg-type problem.** *Theor. Decis.* 1992; **32**: 77–100.
Publisher Full Text

Brainard DH: **The Psychophysics Toolbox.** *Spat. Vis.* 1997; **10**(4): 433–436.
Publisher Full Text

Carpenter PA, Just MA, Shell P: **What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test.** *Psychol. Rev.* 1990; **97**(3): 404–431.
PubMed Abstract | Publisher Full Text

Etchart-Vincent N, l'Haridon O: **Monetary incentives in the loss domain and behavior toward risk: An experimental comparison of three reward schemes including real losses.** *J. Risk Uncertain.* 2011; **42**: 61–83.
Publisher Full Text

Fedorenko E, Duncan J, Kanwisher N: **Broad domain generality in focal regions of frontal and parietal cortex.** *Proc. Natl. Acad. Sci.* 2013; **110**(41): 16616–16621.
PubMed Abstract | Publisher Full Text | Free Full Text

Fiedler S, Glöckner A: **The dynamics of decision making in risky choice: An eye-tracking analysis.** *Front. Psychol.* 2012; **3**: 335.
Publisher Full Text

Fröber K, Pittino F, Dreisbach G: **How sequential changes in reward expectation modulate cognitive control: Pupillometry as a tool to monitor dynamic changes in reward expectation.** *Int. J. Psychophysiol.* 2020; **148**: 35–49.
PubMed Abstract | Publisher Full Text

Frydman C, Mormann MM: **The role of salience in choice under risk: An experimental investigation.** *Available at SSRN, 2778822.* 2018.

Glaholt MG, Wu MC, Reingold EM: **Predicting preference from fixations.** *PsychNology J.* 2009; **7**(2): 141–158.

Glöckner A, Herbold AK: **An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes.** *J. Behav. Decis. Mak.* 2011; **24**(1): 71–98.
Publisher Full Text

Gonthier C, Roulin JL: **Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition.** *J. Exp. Psychol. Gen.* 2020; **149**(3): 564–579.
PubMed Abstract | Publisher Full Text

Gonthier C, Thomassin N: **Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices.** *J. Exp. Psychol. Gen.* 2015; **144**(5): 916–924.
PubMed Abstract | Publisher Full Text

Goldstein WM, Einhorn HJ: **Expression theory and the preference reversal phenomenon.** *Psychol. Rev.* 1987; **94**(2): 236–254.
Publisher Full Text

Gonzalez R, Wu G: **On the shape of the probability weighting function.** *Cogn. Psychol.* 1999; **38**: 129–166.
Publisher Full Text

Hajimoladarvish N: **How do people reduce compound lotteries?** *J. Behav. Exp. Econ.* 2018; **75**: 126–133.
Publisher Full Text

Halevy Y: **Ellsberg revisited: An experimental study.** *Econometrica.* 2007; **75**(2): 503–536.
Publisher Full Text

Harrison GW, Martinez-Correa J, Swarthout JT: **Reduction of Compound Lotteries with Objective Probabilities: Theory and Evidence.** *J. Econ. Behav. Organ.* 2015; **119**: 32–55.
Publisher Full Text

Harrison GW, Swarthout JT: **Eye-tracking and economic theories of choice under risk.** *J. Econ. Sci. Assoc.* 2019; **5**: 26–37.
Publisher Full Text

Hoffman JE: **Visual Attention and Eye Movements.** Pashler H, editors. *Attention.* Psychology Press;1st ed. 1998.

Hayes TR, Petrov AA: **Pupil diameter tracks the exploration-exploitation trade-off during analogical reasoning and explains individual differences in fluid intelligence.** *J. Cogn. Neurosci.* 2016; **28**(2): 308–318.
PubMed Abstract | **Publisher Full Text**

Janowski V: *Computational Biases in Decision-Making, Ph.* California Institute of Technology; 2012. D Thesis Dissertation.

Johnson EJ, Camerer C, Sen S, *et al.*: **Detecting failures of backward induction: Monitoring information search in sequential bargaining.** *J. Econ. Theory.* 2002; **104**(1): 16–47.
Publisher Full Text

Krajbich I, Armel C, Rangel A: **Visual fixations and the computation and comparison of value in simple choice.** *Nat. Neurosci.* 2010; **13**(10): 1292–1298.
PubMed Abstract | **Publisher Full Text**

Krejtz K, Duchowski AT, Niedzielska A, *et al.*: **Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze.** *PLoS One.* 2018; **13**(9): e0203629.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Lichtenstein S, Slovic P: *The construction of preference.* Cambridge: Cambridge University Press; 2006.

Matousek J, Havranek t, Irsova Z: **Individual discount rates: a meta-analysis of experimental evidence.** *Exp. Econ.* 2022; **25**(1): 318–358.
Publisher Full Text

Miao B, Zhong S: **An experimental study of attitude towards compound lottery.** *Available at SSRN 2039216.* 2012.

Nyström M, Holmqvist K: **An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data.** *Behav. Res. Methods.* 2010; **42**(1): 188–204.
PubMed Abstract | **Publisher Full Text**

Raven J, Raven J: **Raven Progressive Matrices.** McCallum RS, editor. *Handbook of nonverbal assessment.* Kluwer Academic/Plenum Publishers; 2003; pp. 223–237.

Rayner K: **Eye movements in reading and information processing: 20 years of research.** *Psychol. Bull.* 1998; **124**(3): 372–422.
PubMed Abstract | **Publisher Full Text**

Rayner K: **The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search.** *Q. J. Exp. Psychol.* 2009; **62**(8): 1457–1506.
Publisher Full Text

Rosen LD, Rosenkoetter P: **An eye fixation analysis of choice and judgment with multiattribute stimuli.** *Mem. Cogn.* 1976; **4**(6): 747–752.
PubMed Abstract | **Publisher Full Text**

Segal U: **Two-Stage Lotteries without the Reduction Axiom.** *Econometrica.* 1990; **58**(2): 349–377.
Publisher Full Text

Segovia M, Palma M, Lusk JL, *et al.*: **Visual formats in risk preference elicitation: What catches the eye?.** *MPRA Paper No. 115572.* 2022.

Shashidhara S, Mitchell DJ, Erez Y, *et al.*: **Progressive recruitment of the frontoparietal multiple-demand system with increased task complexity, time pressure, and reward.** *J. Cogn. Neurosci.* 2019; **31**(11): 1617–1630.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Starmer C, Sugden R: **Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation.** *Am. Econ. Rev.* 1991; **81**(4): 971–978.

Stewart N, Hermens F, Matthews WJ: **Eye movements in risky choice.** *J. Behav. Decis. Mak.* 2016; **29**(2-3): 116–136.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Stott HP: **Cumulative prospect theory's functional menagerie.** *J. Risk Uncertain.* 2006; **32**: 101–130.
Publisher Full Text

SR RESEARCH, Inc: *EyeLink user manual. Version 1.3.0 [Computer software manual].* Ottawa, ON: Author; 2007.

Su Y, Rao LL, Sun HY, *et al.*: **Is making a risky choice based on a weighting and adding process? An eye-tracking investigation.** *J. Exp. Psychol. Learn. Mem. Cogn.* 2013; **39**(6): 1765–1780.
PubMed Abstract | **Publisher Full Text**

Tatler BW: **The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions.** *J. Vis.* 2007; **7**(14): 4–4, 417.
PubMed Abstract | **Publisher Full Text**

Vigneau F, Bors DA: **The quest for item types based on information processing: An analysis of Raven's Advanced Progressive Matrices, with a consideration of gender differences.** *Intelligence.* 2008; **36**(6): 702–710.
Publisher Full Text

Vigneau F, Caissie AF, Bors DA: **Eye-movement analysis demonstrates strategic influences on intelligence.** *Intelligence.* 2006; **34**(3): 261–272.
Publisher Full Text

Von Neumann J, Morgenstern O: *The Theory of Games and Economic Behavior.* Princeton University Press; 1944.

Wakker PP: **Explaining the characteristics of the power (CRRA) utility function.** *Health Econ.* 2008; **17**: 1329–1344.
PubMed Abstract | **Publisher Full Text**

Zhang X, Aimone JA, Alsharawy A, *et al.*: **The effects of task difficulty and presentation format on eye movements in risky choice.** *Front. Behav. Econ.* 2024; **3**: 1321301.
Publisher Full Text