Using eye-tracking to examine strategies for evaluating compound lotteries

WORKING PAPER 2401

April 2024

Narges Hajimoladarvish

Sneha Shashidhara

Using eye-tracking to examine strategies for evaluating compound lotteries

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

*Both authors contributed equally to this work

April 2024

Abstract

We use eye-tracking to investigate how participants evaluate compound lotteries. We use the order of information acquisition to differentiate between the reduction of compound lotteries axiom of expected utility and the compound independence axiom compatible with any decision theory. We also test these axioms through a new test that combines valuation data with methods used to test these axioms with choice data. Evidence from both eye-tracking and our new test supports the reduction of compound lotteries axiom. We then use a second task from a different domain to generalise these strategies and find eye-tracking and behavioural data support the forward induction approach for problem-solving.

Keywords: Compound lotteries, Reduction of compound lotteries axiom, Compound independence axiom, Eye-tracking, Backward induction strategy, Forward induction strategy.
JEL Classification: C91, D81

1 Introduction

When assuming a holistic approach to lotteries evaluation, there are two ways to reduce and evaluate compound lotteries. One strategy is to start from the end of a sequence of events and reason backwards to determine the value of a lottery, a process like backward induction (Segal, 1990). Another strategy is to compute the probabilities of the final level of outcomes, reduce a compound lottery to a simple one, and then evaluate it (Von Neumann & Morgenstern, 1944). This process is akin to a forward induction approach to problem-solving. We aim to compare these two approaches by using eye-tracking and behavioural data.

Risky situations are often depicted through lotteries. A simple lottery, denoted as S, is a list $S = \left(x_1, p_1; x_2, p_2; \ldots; x_n, p_n\right)$ 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

[1] Senior Research Fellow, Centre for Social and Behaviour Change, Ashoka University, Delhi, India. E-mail: narges.hajimoladarvish@ashoka.edu.in.

[2] Senior Research Fellow, Centre for Social and Behaviour Change, Ashoka University, Delhi, India. E-mail: sneha.shashidhara@ashoka.edu.in.

set of all outcomes as X. For convenience, assume $x_1 < x_2 < \dots < x_n$. We denote the set of simple lotteries by φ. A compound lottery refers to a lottery that allows the outcomes to be lotteries, represented as $C = \left(S_1, q_1; S_2, q_2; \dots; S_n, q_n\right)$ where $q_i \geq 0$ for all i and $\sum_{i=1}^{n} q_i = 1$, and $S_1, \dots, S_n$ are simple lotteries.[3]

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* $S_1 = \left(x_1, p_1; \dots; x_i, p_i; \dots; x_n, p_n\right)$ *and* $S_2 = \left(y_1, q_1; \dots; y_i, q_i; \dots; y_n, q_n\right)$. *Then under the ROCL axiom:*
$$C = \left(S_1, r;\ S_2, 1 - r\right) \sim \left(x_1, rp_1; \dots; x_n, rp_n; y_1, q_1(1 - r) \dots; y_n, q_n(1 - r)\right) = AE(C)$$

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 $A_1 = \left(L, r; x_1, 1 - r\right)$ where $L = \left(y_1, q; y_2, 1 - q\right)$. Under the ROCL, the lottery A_1 is reduced to its actuarially equivalent lottery $AE\left(A_1\right) = (y_1, rq; y_2, r(1 - q); x_1, (1 - r))$.

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 of last-stage lotteries (Segal, 1990).[4] 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* $S, SA, SB \in$ φ *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 $\left(CE(L), r; x_1, 1 - r\right)$, 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. Under expected value and EU, these two axioms have similar predictions. However, the predictions differ for rank-dependent utility or prospect theory preferences.

[3] In 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.

[4] 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.

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 when reducing compound lotteries, the two axioms and usual tests will both hold. 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 if results from eye-tracking data and self-reported valuation of lotteries are consistent. An ideal outcome would be for both data to indicate one axiom. The ROCL or CI axioms explain the validity of the *random lottery incentive mechanism* used in experiments[5]. The random lottery incentive mechanism is incentive-compatible only when either of these two axioms holds (Starmer and Sugden, 1991).

We contribute to the evaluation of compound lotteries in three ways

1. We test the consistency prediction of both the ROCL and the CI axioms using certainty equivalents of compound lotteries.
2. We use eye-tracking for further evidence of the axioms used as strategies in decision-making.
3. We compare problem-solving strategies in tasks of different natures and domains.

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 Roisenkoetter, 1976; Russo and Dosher, 1983; Arieli et al., 2011; Glöckner and Herbold, 2011; Fiedler and Glöckner, 2012; Janowski, 2012; Su et al., 2013; Stewart et al. 2016; Harrison and Swarthout, 2019; Zhang et al., 2024). While all of these studies used data while subjects were making choices over alternatives, we used data while they were evaluating lotteries individually. Moreover, our study extends this growing literature to compound lotteries and focuses on the underlying strategies

[5] See appendix for demonstration of this claim.

used in the valuation process. Hence, eye movements offer a method for distinguishing between different approaches to processing information.

The paper is organised as follows. Section 2 explains the experimental design. Section 3 describes the methods used to test the ROCL and the CI axioms. Section 4 reports the data and findings. Section 5 explores the correlation between cognitive abilities and ROCL and CI consistent choices. Section 6 discusses the results, and section 7 concludes.

2 Experimental Design

The experiment consisted of two tasks, and eye-tracking was used for both. The order in which these two tasks were completed was counterbalanced. Subjects were placed such that their eyes were approximately 156.5 cm away from the monitor. They were instructed to stay as still as possible during the task and were not provided external fixators like a chin rest.

2.1 Task 1

In the first task, we elicited the certainty equivalent of a series of simple and compound lotteries. Compound lotteries are denoted by A and B, and simple lotteries are marked as$S$, SA, SB, $AE(A)$, $AE(B)$.

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.

To test the ROCL axiom, we need to elicit the certainty equivalent of compound lotteries and their associated actuarially equivalent lotteries. Participants typed in their 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. There was no time constraint, and there were one-second fixation screens between trials where the participants were instructed to look at the dot on the screen.

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. The battery of compound lotteries is presented in Table 1. The associated actuarily equivalent lotteries are presented in Table 2. Note that lottery B's are first-order stochastically dominated by lottery A's (they have lower expected value).

All lotteries were represented as decision trees. Figure 1 (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.

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.

Table 1: Battery of compound lotteries

		$A = (S, r; SA)$						$B = (S, r; SB)$		
Pairs	r	$S = (x_1, p_1; x_2)$			$SA = (x_1, p_1; x_2)$			$SB = (x_1, p_1; x_2)$		
		x_1	p_1	x_2	x_1	p_1	x_2	x_1	p_1	x_2
1	0.5	1000	1		2000	0.10	4000	2000	0.5	4000
2	0.5	1000	0.75	2000	1000	0.10	4000	1000	0.5	4000
3	0.9	4000	1		2000	0.5	4000	2000	0.9	4000
4	0.9	1000	0.5	2000	1000	0.5	2000	1000	0.9	4000
5	0.9	2000	1		1000	0.25	4000	1000	0.25	2000
6	0.75	1000	0.9	4000	1000	0.75	2000	1000	0.90	2000
7	0.5	1000	1		2000	0.75	4000	2000	0.90	4000
8	0.25	2000	1		1000	0.5	4000	1000	0.75	4000
9	0.75	4000	1		2000	0.5	4000	2000	0.75	4000

2.2 Task 2

Participants had to solve a pattern completion task in the second task (Figure 1C shows a Hard trial). In a 3X3 pattern, the ninth tile was missing. They were provided with four options below, of which they selected one option that fit the missing tile by pressing the number keys 1, 2, 3, or 4 on the keyboard. 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 (Carpernter, 1990; Vigneau & Bors, 2008; Raven and Raven, 2023).

Table 2: Battery of actuarily equivalent lotteries

$AE(A)$					$AE(B)$				
$(x_1, p_1; x_2, p_2; x_3, p_3)$					$(x_1, p_1; x_2, p_2; x_3, p_3)$				
x_1	p_1	x_2	p_2	x_3	x_1	p_1	x_2	p_2	x_3
2000	0.05	4000	0.45	1000	2000	0.25	4000	0.25	1000
2000	0.125	1000	0.425	4000	2000	0.125	1000	0.625	4000
2000	0.05	4000	0.95		2000	0.01	4000	0.99	
2000	0.50	1000	0.50		2000	0.45	4000	0.01	1000
4000	0.075	1000	0.025	2000	1000	0.025	2000	0.975	0
4000	0.075	1000	0.8625	2000	4000	0.075	1000	0.90	2000
1000	0.50	2000	0.375	4000	1000	0.50	2000	0.45	4000
2000	0.25	4000	0.375	1000	2000	0.25	4000	0.1875	1000
4000	0.875	2000	0.125		4000	0.8125	2000	0.1875	

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 & Bors, 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, with constructive/forward induction strategy correlated with higher intelligence and response elimination/backward strategy correlated with lower intelligence. One subject does not need to follow the same strategy throughout, as evidenced by Gonthier and Roulin (2019), who show that subjects tend to use more response elimination in hard trials. They suggest constructive/forward induction is effective but costly for hard trials.

In the reasoning task, the pattern was 384 pixels along width and height, and each option was 128 pixels along width and height. This is 2.51° and 0.84° visual angle along the width and height for pattern and options, respectively. The tasks were written and presented using Psychtoolbox-3 (Brainard, 1997) and MATLAB (The MathWorks).

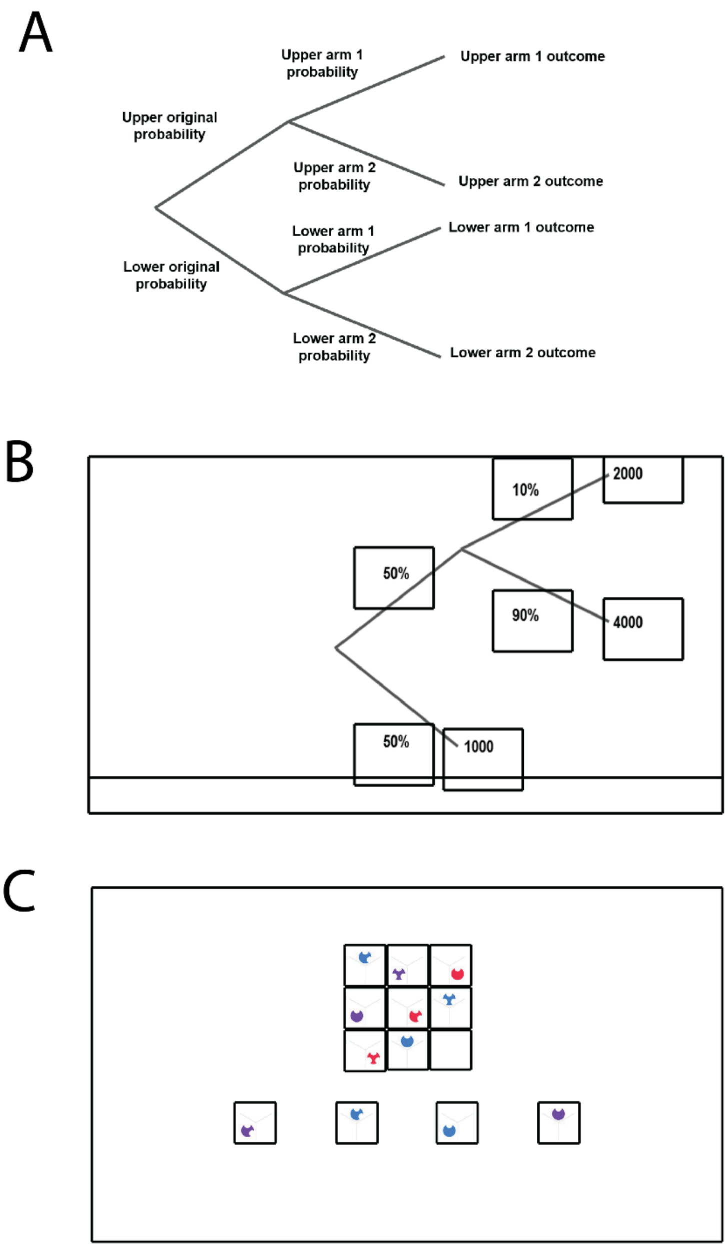


Figure 1: *Task paradigm. A. Decision tree of a compound-2x lottery shown with the nomenclature followed for each probability and outcome. These are consistent across lotteries. B. An example of compound-up lottery. The segment at the bottom of the screen illustrates the area where subjects entered the 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 not displayed during the task.*

3 Methods used to test the ROCL and CI axioms

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 S , 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 relies on the valuation of certainty equivalents. In general, evidence shows that choice data are more consistent (see, for example, Schmidt and Trautmann (2014) and Harbaugh et al. (2010)).

The methodology to test the CI axiom is based on similar methods. Hajimoladarvish (2018) tested if 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. Compared to the consistency test through 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 and not preference ordering. This presumes that if we do have well-defined preferences, the preference of A over B translates into A having higher value than B, and vice versa. Table 3 summarises our new method for testing the CI and ROCL axioms and compares it to the consistency test used with choice data.

Table 3: Methods to test the CI and ROCL axioms

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

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

3.1 Eye-tracking

The diameter and position of the participant's left and right pupils were continuously measured using a Tobi pro fusion eye tracker with a sampling rate of 120 Hz. A 5-point spatial calibration was performed before the task began. The raw data was averaged across the two pupils. Further, fixation and saccade time points were separated using a velocity threshold of 30 deg per sec (Nyström & Holmqvist, 2010; SR Research, 2007). All points below the threshold velocity were considered as fixation. In order to clean the data, only 50ms fixation (Rayner, 1998) or higher was used. Raw data and fixation data without cleaning were also analysed for robustness checks.

For each trial, we extracted the fixation points around the number in the lottery task (70 pixels in all directions) and in each tile in the reasoning task (an additional 8 pixels were added to the width and height only for the option tiles). The mean of the fixation points (lesser value indicates earlier in order) divided by the trial reaction time for each number/tile was used to determine the order. For example, to compare whether the upper original probability was looked at before the upper-arm-1 probability in the compound-up lotteries, we calculated the mean fixation points per trial in the two probability windows and averaged across the six compound-up trials for each subject. We then did a paired t-test to determine which probability was accessed first across subjects. This mean of fixation points allows for glancing at all numbers/tiles before processing them, as a lower mean would suggest the number was processed first rather than just fixated on first. To examine the pattern in individual subjects, we used the same mean fixation points averaged across trials but compared whether the mean of the upper original probability was lower than the upper-arm-1 probability without a t-test. This analysis is based on numerical differences to understand the consistency of the aggregate subject analysis. The same would hold for comparing if a particular pattern tile was looked at before an option tile.

Lastly, we calculate the number of times the fixation moved between windows. For a given window, we calculate the times there was a jump from that window to all other windows in each trial. For this analysis, we calculate the Euclidean distance from every fixation that does not fall into a window to the centre of all windows. The point is then assigned to the window with the lowest distance to have an unbroken sequence of movements in a trial.

We also computed the number of trials where no eye-tracking data was collected for each participant. There were only two subjects in the lottery task and four in the reasoning task, with no data in a few trials.

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 by one. Recruitment of participants occurred through the SONA system. All participants are undergraduates from Ashoka University, except three PhD students. A sample size of 70 was determined before data collection, as typical for eye-tracking studies. All participants were right-handed with normal or corrected-to-normal vision and had no neurological or psychiatric illness history. 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 Ashoka Research Ethics Committee. All participants gave written informed consent. On average, each session took about 26 minutes, excluding the instruction. Subjects were paid a flat participation fee of 400 rupees, approximating $4.82.

4.1 Test of the ROCL axiom

We use the Cochran Q test coupled with the Bonferroni-Dunn (B-D) correction procedure to test the consistency prediction of the ROCL axiom. We test the hypothesis that subjects evaluate lottery A higher than lottery SA in the same proportion as the actuarily equivalent simple lottery $AE(A)$ over SA lotteries. Hence, we are testing if the proportion of subjects indicating higher certainty equivalents for lottery As as compared to lottery SAs is the same as the proportion of subjects having higher certainty equivalent for $AE(A)$ lotteries as compared to SA lotteries.

We do not find evidence to reject the ROCL 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 SAs in the same proportion as the simple $AE(A)$ lotteries over SAs. This implies that our data support the ROCL consistency prediction.

We could test the same hypothesis with B lotteries and test if subjects evaluate lottery B higher than lottery SB in the same proportion as the actuarily equivalent simple lottery $AE(B)$ over SB lotteries or pool both sets and test across 18 pairs. We get similar results with the pooled data as there is no evidence of inconsistent choices. There is only one exception with lottery Bs; for the second pair, we observe a significant difference in π_1 and π_2. These results are reported in the Appendix. With the pooled data, 73% 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_{71} = -0.281, p > 0.779) or gender effect (t_{62} = 0.086, p > 0.931) was observed in the number of choices consistent with the ROCL.

Table 4: The B-D procedure on linked certainty equivalents of A-SA lottery pairs and AE(A)-SA lottery pairs

Pairs	Proportion indicating higher CE for A as compared to SA (π_1)	Proportion indicating higher CE for $AE(A)$ as compared to $SA(\pi_2)$	$\lvert\pi_1 - \pi_2\rvert$
1	0.22	0.16	0.05
2	0.26	0.32	0.05
3	0.93	0.96	0.03
4	0.79	0.88	0.08
5	0.32	0.27	0.04
6	0.92	0.85	0.07
7	0.44	0.22	0.22
8	0.62	0.58	0.04
9	0.84	0.90	0.07

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 0.24. Lottery A has a higher expected value than SA in pairs 3, 6 and 9 and equal expected values in pairs 4.

4.2 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 lottery A higher than 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 As as compared to lottery Bs 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 5 shows the results of the B-D method for each of the 9 comparisons. Table 5 provides evidence that with a 5% familywise error rate, subjects evaluated the A lotteries higher than Bs 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 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_{70} = -0.783, p > 0.435) or gender effect (t_{69} = 0.089, p > 0.929) influencing the frequency of choices consistent with the CI.

Table 5: The B-D procedure on linked certainty equivalents of A-B lottery pairs and SA-SB lottery pairs

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 $\lvert\pi_1 - \pi_2\rvert > d$. 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 0.27.

Is there a bias towards compound lotteries?

We test if the elicited certainty equivalent of compound lotteries and their associated actuarially equivalent lotteries are equal on average. We have 18 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 actuarily 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 lotteries 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 statistically significantly smaller than the objectively 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) as compared to the compound version. Hence, their data suggest a bias towards 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 show a bias towards compound lotteries, while 3 subjects indicate a bias towards actuarily equivalent lotteries.[6] Hence, the finding of bias towards the compound or reduced lottery is a confound associated with reporting average results.

4.3 Eye-tracking data

4.3.1 Lotteries

The eye-tracker records gaze location, time spent on each location and the order in which subjects look at them. We consider the order of average fixation 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 Roisenkoetter (1976), Glöckner and Herbold (2011), and Fiedler and Glöckner (2012). We could use fixation counts or durations as the atoms of analysis. However, our choice is dictated by the axioms that we are testing.

Figure 2 A 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, 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. For analysis, the fixation points from areas around each outcome and probability (see Figure 1 B) are averaged to determine the order in which these numbers are processed.

All compound lotteries used in this study are two-outcome two-stage lotteries where either one or both outcomes of the first-stage is a simple two-outcome lottery. Panel B of Figure 1 shows a case where the upper first-stage outcome is a lottery. We call these types *Compound-up*. *Compound-low* denotes compound lotteries where the lower first-stage outcome is a lottery. *Compound-2x* denotes lotteries where both upper and lower first-stage outcomes are 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 average accuracy was 81.33% (across 27 pairs).

The overall average reaction time 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 simple lotteries.

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 lotteries, we used maximum likelihood to estimate expected utility and rank dependent utility parameters for each individual (see Hajimoladarvish(2018) for details). We then conducted

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

log-likelihood ratio tests to determine the better fit. Interestingly, all of our subjects are classified as expected utility maximisers.

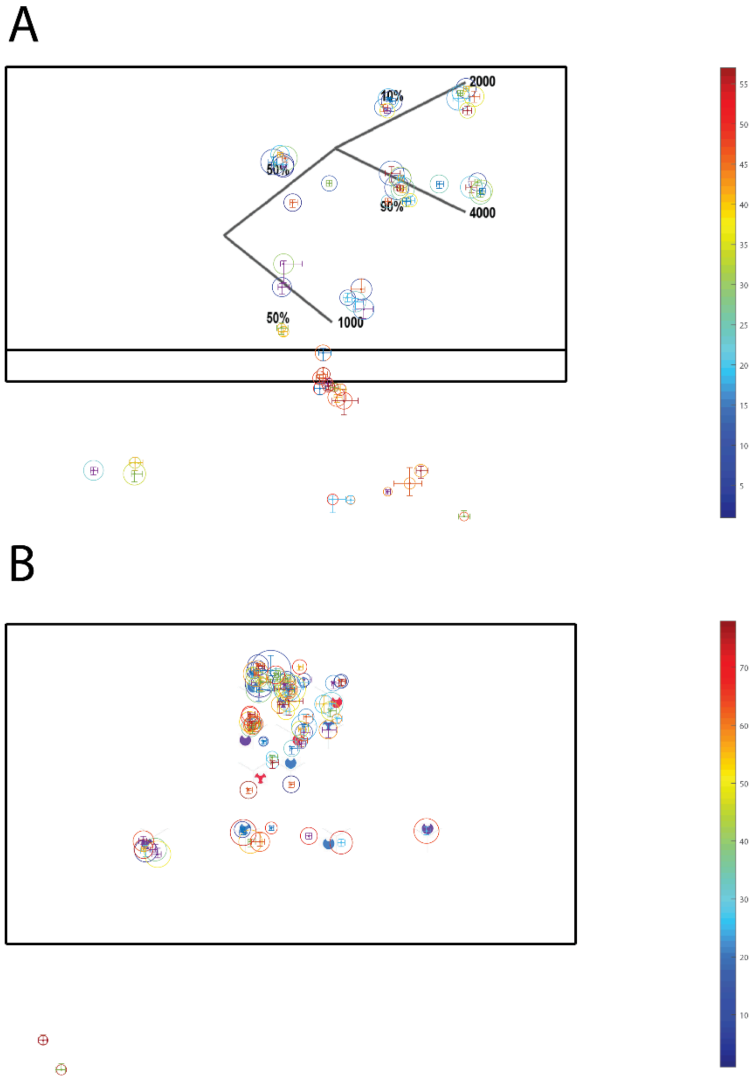


Figure 2: Fixation areas in both tasks. A. This panel shows the average fixation points of a subject in one trial of the compound lottery task. B. This panel shows the average fixation points of a subject in one hard trial of the reasoning task. The size of the points indicates the time spent. The error bars show the mean and standard deviation of movement during each fixation event. The colour scheme indicates the order of fixation, with blue being the first and red being the last. The trials shown here are the same as those in Figure 1.

We first test for regularities in different types of compound lotteries. We find that subjects start from the left of the screen and then move forward, a pattern consistent with the ROCL. This is established by examining the order of fixations of the probabilities and outcomes in the second-stage lottery and comparing them with their corresponding first-stage probability using paired t-tests across all subjects.

In compound-up type lotteries, we find the following regularities:

1- Upper original probability is viewed earlier than upper-arm-1 and arm-2 probabilities and outcomes ($t_{72} > 2.7$, $p < 0.001$, Bonferroni corrected for four comparisons $p_{corrected} < 0.03$).

2- In the second-stage lottery, the order of fixations shows no preference between upper-arm-1 probability [$p > 0.1$] and upper-arm-2 probability followed by the upper-arm-1 outcome and then upper-arm-2 outcome ($t_{72} > 3.31$, $p < 0.001$, Bonferroni corrected for six comparisons $p_{corrected} < 0.01$).

If subjects followed the CI axiom, they should have viewed the upper original probability after the second-stage components. Johnson et al. (2002) used mouse-tracking to study backward and forward induction, which aligns perfectly with the CI axiom as a backward induction method and the ROCL as a forward induction one. They find evidence for forward induction, which is consistent with our findings.

The pattern of fixations in compound-low type lotteries is also consistent with the ROCL. Most importantly, lower original probability is viewed before second-stage components. We also find the following regularities:

1- Lower original probability is viewed earlier than lower-arm-1 and arm-2 outcomes and lower-arm-2 probability ($t_{72} > 3.38$, $p < 0.001$, Bonferroni corrected for four comparisons $p_{corrected} < 0.001$) but not earlier than lower-arm-1 probability.

2- In the second-stage lottery, the order of fixations is lower-arm-1 probability, lower-arm-1 outcome, lower-arm-2 probability and lower-arm-2 outcome ($t_{72} > 3.88$, $p < 0.001$, Bonferroni corrected for six comparisons $p_{corrected} < 0.001$).

The pattern of fixations in compound-2x type lotteries is also consistent with the ROCL. Thus, the pattern does not change as the uncertainty increases. By increase in uncertainty, we mean when both outcomes of a simple lottery are lotteries instead of just one. Upper original probability is viewed before upper-arm-1, and arm-2 probabilities and outcomes, and lower original probability is viewed before lower-arm-1 and arm-2 probabilities and outcomes, with all p-values from paired t-tests across all subjects being less than 0.01. More specifically, we find the following:

1- The upper lottery follows the same pattern as the complex-up lottery.

2- The lower lottery follows a similar pattern to the complex-low lottery with minor differences that don't affect the inference of which axioms they follow.[7]

Similar patterns are observed in individual choices. For compound-up lotteries with 7 fixations, there are 5040 possible orders. Among these possibilities, there are 68 unique

[7] Lower original probability was viewed earlier than lower-arm-1 probability. There was no preference between lower-arm-1 outcome and lower-arm-2 probability.

orders. The most stringent test is to examine if the upper original probability is seen before all the characteristics of the second-stage lottery. That is if the upper original probability is viewed earlier than the outcomes and probabilities of the upper-arm-1 and upper-arm-2. 42 subjects (57.53%) in our sample passed this stringent test. Another way of testing, if subjects have followed the ROCL axiom, is to focus only on outcomes of the second-stage and examine if the upper original probability is viewed before outcomes of upper-arm-1 and upper-arm-2. In our sample, 58 subjects (79.45%) follow this pattern. Finally, we can focus only on either of the arm's outcomes and examine if they are viewed later than the upper original probability. This is because the necessary condition for the CI axiom requires subjects to view both second-stage outcomes before the original probability. If this condition is not met, it is sufficient evidence for the ROCL axiom. 66 subjects (90.41%) viewed the upper original probability before either of the second-stage outcomes. This leaves us with 7 (9.59%) subjects who viewed the upper original probability after both second-stage outcomes and probabilities, suggesting they followed the CI axiom.

We observe a similar pattern for compound-low lotteries, with 67 subjects (91.78%) viewing lower original probability before either of the second-stage outcomes. This leaves us with 6 (8.22%) subjects that viewed the lower original probability after both second-stage outcomes.

For compound-2x lotteries with 10 numbers to process, there are 3628800 possibilities. Again, we examine if the corresponding original probability is viewed before either of the second-stage outcomes. 66 subjects (90.41%) follow this pattern in the upper lottery. Similarly, we find that for 65 subjects (89.04%), the lower original probability is viewed before either of the lower second-stage outcomes. Thus, 7 (9.59%) and 8(10.96%) subjects see the original probability after the second-stage outcomes. Hence, irrespective of the type of compound lotteries considered herein, we find evidence suggestive of the ROCL axiom through averages across subjects and strategies used by individual subjects.

Lastly, we investigated the sequence of movements or jumps between the numbers in a trial to ascertain the strategy used by individual subjects. Here, if, in any trial, a given subject looked at the second-stage probabilities and outcomes in one sequence (in any order) before any fixation on the corresponding original probability; it was taken as evidence that the subject may have employed the CI axiom. We see that 21 subjects (28.77%) showed evidence of CI axiom in at least one trial in the compound-up trials. 25 subjects (34.25%) showed evidence of CI axiom in the compound-low trials. For compound-2x trials, 14 subjects (19.18%) showed evidence of the CI axiom in the upper second-stage lottery, and no subjects showed evidence of the CI axiom in the lower second-stage lottery. There are 12 subjects common between the compound-up and compound-low trials and 4 subjects common between all three compound-up, compound-low and compound-2x upper lottery.

Position differences

Looking at the subset of lotteries displayed in the left position, we again find evidence for the ROCL axiom. Interestingly, we find that the lower original probability is not different from the lower-arm-1 probability and outcome in the fixation order in both compound-low and compound-2x. The compound-up lottery remains the same. We observe a similar pattern in the middle and right positions as when using all the positions.

Simple lotteries

Here, we analyse the simplest lotteries with two outcomes and two probabilities[8]. Our general intuitive finding provides the sanity check for the order of screen viewing, which goes from top to bottom, and probabilities are viewed earlier than outcomes. We find that subjects follow a "Z" shaped movement as the order of fixations is upper probability, upper outcome, lower probability and lower outcome in two-outcome simple lotteries ($t_{72} > 4.06$, $p < 0.001$, Bonferroni corrected for six comparisons $p_{corrected} < 0.001$). Furthermore, the upper arm (mean of upper probability and outcome) was viewed earlier than the lower arm ($t_{72} = 13.49$, $p < 0.001$), and the mean of the outcomes was viewed later than the mean of the probabilities ($t_{72} = 10.94$, $p < 0.001$).

The same pattern is observed in individual choices as 35 (47.95%) subjects follow the same "Z" shape order. There are 9 other unique orders among 24 possibilities of fixations in our data, with 22 subjects (30.14%) following the order of upper probability, lower probability, upper outcome, and lower outcome.

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) and hard (26.10) trials ($t_{72}=16.65$, $p < 0.001$), with easy trials being faster. This is expected in all hard versus easy comparisons across various different tasks and modalities (Fedorenko et al., 2013; Shashidhara et al., 2019; Vigneau & Bors, 2006).

To test the use of backward and forward induction strategies in the reasoning task, we examined whether subjects first looked at the option tiles before viewing the pattern tiles. In the case of both easy and hard trials, on average, the pattern tiles were viewed before the option tiles ($t_{72}>25.32$, $p < 0.001$, Bonferroni corrected for two comparisons $p_{corrected} < 0.001$). We further separated the option tiles as the target and distractors, separating the trials when a given option was the answer versus when it was not. We see that the pattern was viewed earlier than the average target and average distractors ($t_{72}>20.39$, $p < 0.001$, Bonferroni corrected for two comparisons $p_{corrected} < 0.001$).

To further look at individual strategies, we calculated the number of subjects that saw option tiles before pattern tiles. In the easy trials, 30 subjects (53.42%) saw no options before seeing all of the pattern tiles, including the spot where the missing tile would go.

[8] A few actuarily equivalent lotteries have three outcomes which were not analyzed further.

Note, given the easy nature of the trials, subjects need not see each of the pattern tiles to figure out the answer. Considering the number of subjects that did not see any of the options before the first row and the first column, sufficient condition to solve the problem, we find between 72 subjects (93.63%) and 62 subjects (84.93%) do not see any options before these critical tiles. Similarly, 48 subjects (65.75%) saw all the pattern tiles, including the missing tile, before seeing the options for the hard trials. Looking at the first row and column, between 70 subjects (95.89%) and 64 subjects (87.67%) do not see any options before these critical tiles.

The proportion of the trial time that elapsed before any option was fixated on (time spent purely on the pattern) was 0.71 (0.09) for easy trials and 0.72 (0.09) for hard trials; with no difference between the two. The average proportion of time spent on the pattern is 0.85 (0.05) for easy trials and 0.84 (0.06) for hard trials; with no significant difference between them. The proportion of time spent on option is 1 - the proportion of time spent on the pattern. The proportion of trial time spent on the pattern before the first toggle to the options and the average proportion of time spent on the pattern indicates the forward induction strategy (Vigneou & Bors 2006). While they did see a correlation between both measures and difficulty, we did not. This could be because of the binary classification of difficulty rather than item-wise classification for each trial.

Lastly, we look at the number of switches or jumps between the pattern and option tiles. Many alternations point towards response elimination/backward induction strategy, and fewer suggest constructive matching/forward-induction strategy (Bethell-Fox et al., 1984; Vigneou & Bors, 2006). In the easy trial, the median number of switches equals 3.39. That is, subjects, on average, look at the option, go back to the pattern, and come back to the option and so on three times in a given easy trial. As expected, this number is higher in the hard trials, with a median number of switches equal to 7.28 ($t_{72} = 14.77$, $p < 0.001$), given the longer trial time. The correlation of switches in easy and hard trials across subjects is 0.68 (p <0.001). When divided by the trial time, we do not see a significant increase in jumps. Once again, we fail to see differences in strategy with difficulty. While the 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 Ashoka University students, one of the top universities in India, we may have an average IQ higher than that of a typical sample.

5 Cognitive abilities and the ROCL and CI axioms

As with any other axiom of choice, there is some suggestive evidence indicating that violations of the ROCL are due to cognitive limitations (Bernasconi and Bernhofer, 2019; Nebout and Dubois, 2014; Harrison et al., 2015; Prokosheva,2016). Hence, we test if people with higher reasoning abilities violate ROCL less than those with lower reasoning scores. We test if the consistent choices of individuals with high reasoning scores (above median) are higher than those with low reasoning scores (below median).

A significant correlation exists between the number of ROCL consistent choices and reasoning scores for easy questions ($r = 0.21$, $p = 0.036$) and not CI consistent choices. On average, while individuals with above-median scores in easy reasoning trials have made 74% ROCL consistent choices, subjects with below-median scores have made 71%

ROCL consistent choices. A chi-square test shows no significant difference between the distribution of ROCL consistent choices between groups with high and low reasoning scores (p = 0.594).

Interestingly, we find a significant correlation between the number of CI consistent choices and reasoning scores for hard trials (r=0.19, p-value=0.049) and not for ROCL consistent choices. On average, while individuals with above-median scores in hard trials have made 78% CI consistent choices, subjects with below-median scores have made 69% CI consistent choices. Hence, we do observe an increase in the number of the CI consistent choices as performance in the hard trials of the reasoning task increases. The chi-square test shows a significant difference between the distribution of CI consistent choices between groups with high and low reasoning scores (p-value = 0.064).

6 Discussion

Our study looks into the evaluation of compound lotteries, utilising eye-tracking to observe the sequence of gaze patterns when assessing these lotteries. Examining the order of fixations, we test the ROCL and the CI axioms. If subjects use the CI axiom to reduce compound lotteries, they should first evaluate last-stage lotteries. We do not observe this pattern. We find evidence supporting the ROCL axiom through data averaged across subjects. Looking at strategies used by each individual and even each trial, we find evidence of subjects using the CI axiom, but it is limited. The finding holds for all types of compound lotteries considered herein. Similar to the use of backward induction in sequential games with perfect information, we expected to see more support for the CI axiom in more complex compound lotteries where both outcomes of the first-stage were lotteries instead of one.

Our test, which integrates methods used with choice and valuation data, further supports the ROCL and the CI axioms. When combined with eye-tracking we show more evidence for ROCL than the CI axiom.

In a second task completed by the same subjects, we examined the strategies used in a general problem-solving context, here a pattern completion task. On average, we find that participants examined the pattern and only then looked at the options to select their answers. This suggests that participants computed and answered templates based on the pattern that would fit the missing tile and then looked for that template among the options. This forward induction strategy, i.e. computing the answer before finding it in the options, resonates with the ROCL axiom. On the other hand, if participants tried to fit one of the options in the missing tile without first examining the problem thoroughly, we would conclude with the use of a backward induction strategy.

For robustness checks, we redid the analysis for both tasks with raw data and fixation data without cleaning and found similar results. Therefore, our results are not a product of the cleaning process.

Our lottery task findings are based on hypothetical choices that raise certain concerns regarding the validity of the results. This is because we are not interested in the value of lotteries to estimate preferences; rather, we wanted to use eye movements to differentiate between strategies employed in lottery valuation. Moreover, the validity of the most common method of incentives employed in experiments, the random lottery incentive mechanism itself, relies on either of the two axioms we are investigating. We could have paid subjects for each choice and used small stakes, a choice prone to *income effects (*Schläpfer et al., 2006*)*. Thus, we opted for larger stakes. Further, as we wanted to avoid strategic responses that are more common in valuation tasks than choice tasks, we didn't use real incentives. The valuation nature of the experiment with an incentive-based procedure would have complicated the interpretation of the results and we chose the simpler version with the caveat of non-generalizability. We assume subjects have no special reason to disguise their true preferences.

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. 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 as well we observe the same pattern of forward induction. Furthermore, the findings of Segovia et al. (2022) show that elicited parameters for risk preferences do not vary with presentation formats. Future eye-tracking studies should investigate using the ROCL compared to the CI strategy in different formats.

To investigate problem-solving strategies more generally, we opted for a task from a completely different domain that used patterns instead of numbers. This helps us mitigate the issue of lottery representation and the tendency to process numbers from left to right. Further, we have two levels of difficulty in both tasks. We can consider simple lotteries as easy conditions and compound lotteries as hard conditions. In the reasoning task, we also have easy and hard problems. These are validated by both greater reaction times and lower accuracies in the hard versus the easy conditions in both tasks. While these tasks are not directly comparable, we can compare the intuition behind problem-solving strategies. We find evidence for forward induction in both tasks.

7 Conclusion

In conclusion, our study provides insights into the mechanisms used when evaluating compound lotteries through the utilisation of eye-tracking technology. Our investigation into gaze patterns during lottery evaluation answers an open question in the literature that aims to distinguish between the ROCL and CI axiom through a given test.

Our findings do not align with a backward sequential evaluation pattern, i.e., the CI axiom. Instead, 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 seem to use forward induction strategies in complex and uncertain situations used herein.

References

Abdellaoui, M., Klibanoff, P., & Placido, L. (2015). Experiments on compound risk in relation to simple risk and to ambiguity. *Management Science*, *61*(6), 1306-1322.

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

Bansal, S., & Rosokha, Y. (2018). Impact of compound and reduced specification on valuation of projects with multiple risks. *Decision Analysis*, *15*(1), 27-46.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational behavior and human performance*, *9*(3), 396-406.

Becker, G., DeGroot, M., Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science* 9, 226–236.

Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, 8(3), 205-238.

Bernasconi, M. (1994). Nonlinear preferences and two-stage lotteries: Theories and evidence. *The Economic Journal*, *104*(422), 54-70.

Bernasconi, M., & Bernhofer, J. (2020). Catch me if you can: Testing the reduction of compound lotteries axiom in a tax compliance experiment. *Journal of Behavioral and Experimental Economics*, *84*, 101479.

Bernasconi, M., & Loomes, G. (1992). Failures of the reduction principle in an Ellsberg-type problem. *Theory and Decision*, *32*, 77-100.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological review*, 97(3), 404.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41), 16616-16621.

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in psychology*, 3, 335.

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

Glaholt, M. G., Wu, M. C., & Reingold, E. M. (2009). Predicting preference from fixations. *PsychNology J.*, 7(2), 141-158.

Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71-98.

Gonthier, C., & Roulin, J. L. (2020). Intraindividual strategy shifts in Raven's matrices, and their dependence on working memory capacity and need for cognition. *Journal of Experimental Psychology: General*, 149(3), 564.

Gonthier, C., & Thomassin, N. (2015). Strategy use fully mediates the relationship between working memory capacity and performance on Raven's matrices. *Journal of Experimental Psychology: General*, 144(5), 916.

Hajimoladarvish, N. (2018). How do people reduce compound lotteries? *Journal of Behavioral and Experimental Economics* 75, 126–133.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica* 75(2), 503–536.

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2010). The fourfold pattern of risk attitudes in choice and pricing tasks. *The Economic Journal*, *120*(545), 595-611.

Harrison, G. W., Martinez-Correa, J. and Swarthout, J. T. (2015). Reduction of Compound Lotteries with Objective Probabilities: Theory and Evidence. *Journal of Economic Behavior & Organization*, 119, 32-55.

Harrison, G. W., & Swarthout, J. T. (2019). Eye-tracking and economic theories of choice under risk. *Journal of the Economic Science Association*, 5, 26-37.

Hoffman JE. (1998) Visual Attention and Eye Movements. In H. Pashler (Ed.), *Attention (1st ed.)*. Psychology Press.

Janowski, V. (2012). Computational Biases in Decision-Making, Ph.D Thesis Dissertation, California Institute of Technology.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of economic theory*, 104(1), 16-47.

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

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13(10), 1292-1298.

Lichtenstein, S. and Slovic, P. (2006). The construction of preference. Cambridge: Cambridge University Press.

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

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods*, 42(1), 188-204.

Raven, J., & Raven, J. (2003). Raven Progressive Matrices. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 223–237). Kluwer Academic/Plenum Publishers.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 372.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly journal of experimental psychology,* 62(8), 1457-1506.

Rosen, L. D., & Rosenkoetter, P. (1976). An eye fixation analysis of choice and judgment with multiattribute stimuli. *Memory & Cognition*, 4(6), 747-752.

Segal, U. (1990). Two-Stage Lotteries without the Reduction Axiom. *Econometrica* 58(2), 349-377.

Segovia, M., Palma, M., Lusk, J. L., & Drichoutis, A. (2022). Visual formats in risk preference elicitation: What catches the eye? *MPRA Paper No. 115572.*

Schläpfer, F. (2006). Survey protocol and income effects in the contingent valuation of public goods: A meta-analysis. *Ecological Economics*, 57(3), 415-429.

Schmidt, U., & Trautmann, S. T. (2014). Common consequence effects in pricing and choice. *Theory and decision*, *76*, 1-7.

Shashidhara, S., Mitchell, D. J., Erez, Y., & Duncan, J. (2019). Progressive recruitment of the frontoparietal multiple-demand system with increased task complexity, time pressure, and reward. *Journal of cognitive neuroscience*, 31(11), 1617-1630.

Starmer, C. and Sugden, R. (1991). Does the Random-Lottery Incentive System Elicit True Preferences? An Experimental Investigation. *American Economic Review* 81(4), 971-978.

Stewart, N., Hermens, F., & Matthews, W. J. (2016). Eye movements in risky choice. *Journal of behavioral decision making*, 29(2-3), 116-136.

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

Su, Y., Rao, L. L., Sun, H. Y., Du, X. L., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(6), 1765.

Vigneau, F., & Bors, D. A. (2008). The quest for item types based on information processing: An analysis of Raven's Advanced Progressive Matrices, with a consideration of gender differences. *Intelligence*, 36(6), 702-710.

Vigneau, F., Caissie, A. F., & Bors, D. A. (2006). Eye-movement analysis demonstrates strategic influences on intelligence. *Intelligence*, 34(3), 261-272.

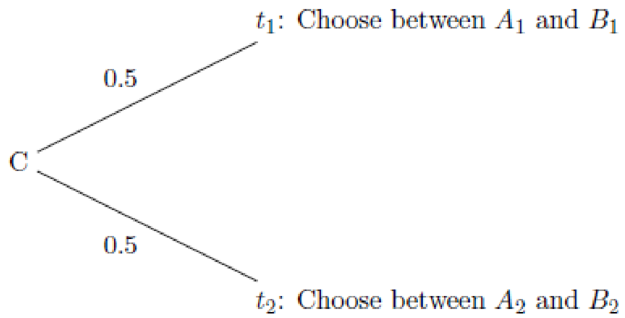
Von Neumann, J., & Morgenstern, O. (1944). *The Theory of Games and Economic Behavior*. Princeton university press.

Zhang, X., Aimone, J. A., Alsharawy, A., Li, F., Ball, S. B., & Smith, A. (2024). The effects of task difficulty and presentation format on eye movements in risky choice. *Frontiers in Behavioral Economics*, 3, 1321301.

Appendix

Random lottery incentive mechanism

Experiments employing the random lottery incentive mechanism inform subjects at the beginning that one of the tasks will be picked randomly, and their choice of that given task will determine their pay-off from the experiment. This procedure is more compatible with the CI axiom as subjects treat each problem in isolation, whereas under the ROCL, the probability of incentive is part of the evaluation process (Harrison et al., 2015). If there are n tasks, the probability that task i is selected for actual payment is $\frac{1}{n}$. For simplicity, let's assume there are only two tasks; t_1 and t_2; hence, the probability that each task being chosen is 0.5. Assume in each task, subjects are asked to choose between two options A_i and B_i, where i denotes the task number. This choice problem can be represented by the decision tree below.

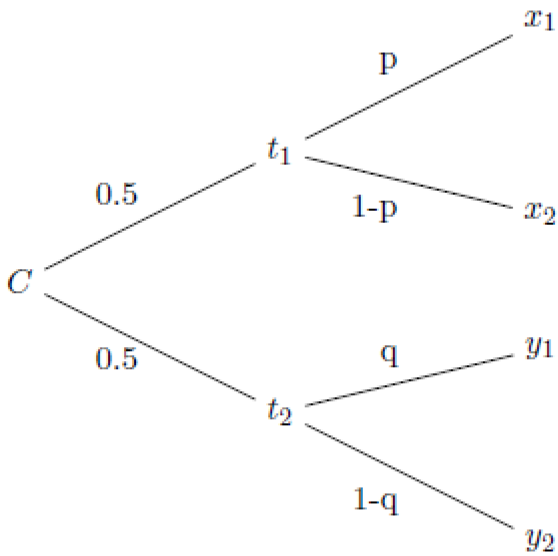


Under the ROCL axiom of EU, if a subject prefers A_1 to B_1 in a single choice problem, they will also choose A_1 to B_1 in the above problem under the random lottery incentive mechanism as

$$u(A_1) > u(B_1) \Leftrightarrow 0.5u(A_1) + 0.5u(A_2/B_2) > 0.5u(B_1) + 0.5u(A_2/B_2).$$

Under the CI axiom, which is compatible by non-EU theories as well, subjects first evaluate the last stage lotteries of a compound lottery. Hence, they first evaluate A_1 and B_1 and replace them with their certainty equivalents that can be calculated according to any decision theory. If they put more value on A_1 as compared to B_1, then they will choose A_1 under the random lottery incentive mechanism as well.

Now, consider a case where t_1 and t_2 involve the valuation of lotteries. This choice problem can be represented by a decision tree below.



Under the CI axiom, to evaluate the above compound lottery, subjects first need to evaluate t_1 and t_2 and replace them with their certainty equivalents. Hence, their evaluation is not influenced by the random lottery incentive mechanism.

Under the ROCL axiom of expected utility, subjects reduce the above compound lottery as

$$0.5\left[pu\left(x_1\right) + (1 - p)u\left(x_2\right)\right] + 0.5\left[qu\left(y_1\right) + (1 - q)u\left(y_2\right)\right]$$

Given the existence of a utility function, subjects' evaluation of lotteries in t_1 and t_2 in isolation is the same as their evaluation under the random lottery incentive mechanism.

Tables

Table A1: The B-D procedure on linked certainty equivalents of B-SB lottery pairs and AE(B)-SB lottery pairs

Pairs	Proportion indicating higher CE for B as compared to SB (π_1)	Proportion indicating higher CE for $AE(B)$ as compared to $SB(\pi_2)$	$\|\pi_1 - \pi_2\|$
1	0.19	0.12	0.07
2	0.53	0.27	0.26
3	0.95	0.96	0.01
4	0.85	0.74	0.11
5	0.93	0.95	0.01
6	0.93	0.88	0.05
7	0.30	0.19	0.11
8	0.75	0.64	0.11
9	0.95	0.95	0.00

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 0.24.

Table A2: The B-D procedure on linked certainty equivalents of 18 pairs (Both A and B lotteries)

Pairs	Proportion indicating higher CE for compound lottery as compared to Simple ones (π_1)	Proportion indicating higher CE for the actuarily equivalent lottery as compared to Simple ones (π_2)	$\|\pi_1 - \pi_2\|$
1	0.22	0.16	0.05
2	0.26	0.32	0.05
3	0.93	0.96	0.03
4	0.79	0.88	0.08
5	0.32	0.27	0.04
6	0.92	0.85	0.07
7	0.44	0.22	0.22
8	0.62	0.58	0.04
9	0.84	0.90	0.07
10	0.19	0.12	0.07
11	0.53	0.27	0.26
12	0.95	0.96	0.01
13	0.85	0.74	0.11
14	0.93	0.95	0.01
15	0.93	0.88	0.05
16	0.30	0.19	0.11
17	0.75	0.64	0.11
18	0.95	0.95	0.00

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 0.30.