





Experimental methods: Learning your own risk preferences under different risk elicitation methods

Rocco Caferra^a, Andrea Morone^{b,c} ,* Donato Pierno^d 

^a Department of Economics and Finance, University of Bari Aldo Moro, Italy

^b Department of Economics, Management and Law, University of Bari Aldo Moro, Italy

^c Corvinus Institute for Advanced Studies (CIAS), Corvinus University of Budapest, Hungary

^d Department of Economics and Finance, Tor Vergata University of Rome, Rome, Italy

ARTICLE INFO

JEL classification:

C91
D81
D83
D91

Keywords:

Risk preferences
Complexity
Experimental methodology
Elicitation methods
Experience

ABSTRACT

This paper examines whether changes in risk preferences reflect genuine behavioural shifts or result from the elicitation method used. In a within-subject experiment, participants made initial risky choices, completed 24 unpaid learning rounds with real dice and recorded outcomes, and then made final choices. We compare four risk elicitation methods: two Ordered Lottery Selection (OLS) and two Multiple Price List (MPL) tasks. Our analysis focuses on how preferences evolve with experience, the role of reported errors during the learning phase, and the effect of method complexity (i.e., the number of potential payoff outcomes in a given lottery) on revealed risk attitudes. Results show that MPL tasks, particularly when more complex, produce stronger changes in behaviour. Participants who made errors during the MPL learning phase were more likely to display increased risk aversion. This suggests that task comprehension and cognitive load influence the stability of preferences. The findings contribute to the broader debate on preference stability and highlight the methodological relevance of comparing elicitation tools within the same experimental framework.

1. Introduction

Are Risk Preferences Stable? Changes in risk preferences could explain the variation in several economic outcomes. The way we choose to take risks defines various aspects of real life, including the labour market, where risk-averse individuals are less likely to be self-employed, and investments, where risk-averse individuals might be less inclined to invest in stock markets (Falk et al., 2018). Therefore, understanding risk preferences is crucial for identifying different economic dynamics (see Schildberg-Hörisch, 2018). Neoclassical theory claims that people have fixed preferences; however, experimental evidence (both laboratory and field) suggests that preferences can change. As highlighted by Plott (1996), and supported by Delaney et al. (2020), preferences are not always known in advance but can emerge and stabilise through learning processes. In particular, Plott's discovered preference hypothesis (DPH) suggests that, when faced with initially unfamiliar tasks, choices shift from a phase of "myopic" instability to more stable configurations thanks to incentives and feedback; consistently, Delaney et al. experimentally document that discovery can remain incomplete when exploration is costly. Several studies have emphasised the lack of stability in risk preferences across elicitation methods (EMs) and over time, suggesting that observed behavioural

changes may reflect task features as much as genuine preference shifts (Bradbury et al., 2015; Brunette & Ngouhou-Poufoun, 2022; Ert & Haruvy, 2017; Holzmeister & Stefan, 2021; Pedroni et al., 2017; Reynaud & Couture, 2012). This study has three main objectives. First, it aims to provide evidence of changes in risk preferences using the Multiple Price List (MPL) EM, based on the hands-on learning methodology presented by Charness et al. (2023) on the "easy to understand" Ordered Lottery Selection (OLS) EM. Second, it compares these changes with those observed under the simpler OLS method. Finally, it explores the determinants of such variation in risk preferences. Based on this, our main research question is the following: *Are the results of Charness et al. (2023) dependent on the « easy to understand » elicitation procedure used?* As noted by Charness et al. (2013), different experimental techniques vary in their complexity, which in the domain of choice under risk is often characterised by the number of payoff outcomes in a particular lottery (Georgalos & Nabil, 2025) for measuring risk; the methods proposed by Eckel and Grossman (2008) and Charness et al. (2023), known as OLS, require individuals to make simple choices to identify their risk preference. In this case, the subject chooses in a one-shot manner which type of lottery they want to play and, therefore, what degree of risk they want

* Corresponding author at: Department of Economics, Management and Law, University of Bari Aldo Moro, Italy.
E-mail address: andrea.morone@uniba.it (A. Morone).

to take. However, subjects' risk preferences may be more sophisticated than a single isolated choice, and it may be more appropriate to ask the subject to make multiple choices in risky contexts to define the shape of their risk preference better. In this case, tools like MPL, made popular by Holt and Laury (2002), ask subjects to make various risky choices, providing multiple confirmations of their preference. Consequently, these tools may more effectively identify subjects' risk preferences. This distinction aligns with the idea of discovery: the MPL incentivises the exploration of options that would otherwise never be tried, while a one-shot OLS choice may not activate the discovery process. Let us clarify with an example: choosing to take a risk occasionally with methods like OLS, does not necessarily mean the subject is always willing to do so. Going to a casino once does not indicate that one subject is consistently risk-loving over time. With the MPL methods, in order to be considered a risk-lover, subjects must repeatedly choose high-risk options within the same framework. Coming back to the casino example, they would need to decide to enter and bet multiple times. This approach, even if complex, can more reliably track risk preferences. As a result, the first issue related to measuring risk preferences is linked to the choice of the tool. We can now consider a second aspect related to risky choices: variations in the measure. In this case, behavioural studies consider a crucial factor related to changes in risk preferences: experience. Two previous studies, closest to our article, have tested how preferences evolve over time and are not static in a computerised laboratory experiment, such as Ert and Haruvy (2017), and in a paper-and-pencil laboratory experiment (Charness et al., 2023). We aim to explain whether the observed changes by Charness et al. (2023) depend on the easy-to-understand elicitation procedure used, i.e., the specific risk EM employed, or reflect a fundamental change in the measure, i.e., individuals' risk preferences. We use a detailed experimental design to examine the stability and consistency of risk preferences using a hands-on learning experience. In our design, the unpaid learning phase works as minimal but systematic feedback: in line with the DPH, and is free of costs for the subjects in line with Delaney et al. (2020). Our main goals are to assess how consistent risk choices are, understand how preferences remain stable or change over time, and evaluate the effectiveness of different EMs like the MPL and OLS. On the one hand, results demonstrate that risk preferences are variable and adaptable, partly due to the role of experience-based learning. Additionally, we find that errors in reporting outcomes and potential payoffs during the learning phase lead to choices disclosing more risk-averse preferences in the MPL method, while this does not hold for the OLS methodology. This effect persists regardless of the number of possible choices (specifically in MPL). Our results highlight how a limited understanding of complex tasks can influence the process of revealing preferences. They also underscore how the choice of elicitation methodology can shape the direction and magnitude of changes in risk preferences.

2. Risk elicitation methods

Numerous experimental methodologies have been developed in economics to elicit and estimate risk preference.¹ These methods vary in complexity: more complex approaches require greater understanding and mathematical reasoning from participants, often involving multiple confirmation choices to classify risk attitudes. Simpler methods, on the other hand, are more accessible and straightforward. This study employs two distinct risk EMs: the MPL and the OLS. Our interest lies in these two methods, which are opposite in terms of complexity as reported by Charness et al. (2013), and in terms of preference measurement procedure (one vs. multiple choices). We examine the factors

¹ For an overview of the different methodologies, we suggest (Charness et al., 2013), which presents a series of prevailing methods to elicit risk preferences and outlines the advantages and disadvantages of each.

that shape risk preference and how hands-on learning experiences can influence this preference.

2.1. The ordered lottery selection method

The OLS method, a simpler design, was initially developed by Binswanger (1980, 1981) and later popularised by Eckel and Grossman (2008). It captures and estimates risk preferences by asking participants to make a single choice among a set of gambles. In our experiment, we use two variants of OLS: a five-gamble and a six-gamble elicitation method. The two OLS methods are presented to our subjects to achieve two different purposes: first, we challenge Charness's results employing the same six-gamble EM. Second, we chose the most common Eckel and Grossman risk elicitation task due to its popularity in the literature. Nonetheless, this EM inspired researchers to present several variations, such as the Charness et al. (2023) one.

- Six-Gamble Method: Based on Charness et al. (2023) (referred to as CH; see Table 6 in Appendix A. The last three columns are not presented to the experimental subjects.)
- Five-Gamble Method: Adapted from Eckel and Grossman (2008) (referred to as EG; see Table 7 in Appendix A. The last three columns are not presented to the experimental subjects.)

Each gamble presents a 50% probability for both the low and high payoffs, respectively presented to our subjects as Event A and Event B (both payoffs are expressed in experimental currency units ECU). The first row represents a sure gamble, while subsequent rows feature gambles with increasing expected payoffs and standard deviations (i.e., higher risk). However, as reported by Eckel and Grossman (2008), the standard deviation is a proxy for the riskiness of the gamble.

2.2. The multiple price list method

The MPL method, a more complex design, was first introduced by Miller et al. (1969) and later popularised by Holt and Laury (2002). This method presents participants with a list of 10 paired lottery choices (MPL10). Each row presents two options: Option A and Option B. Participants must choose their preferred Option for each row.

- Option A: Offers constant payoffs of 12 or 10 ECU, with changing probabilities.
- Option B: Offers constant payoffs of 22 or 0.50 ECU, with changing probabilities.

In the first six rows, Option B has a higher expected payoff than Option A. A rational participant would typically prefer Option B in these rows. The expected value for Option A ranges from 12 to 10.20 ECU, while Option B ranges from 22 to 2.65 ECU. The full list of 10 choices is detailed in Table 8 (see Appendix A. The last four columns are not presented to the experimental subjects.)

We also employ a variation of the MPL, initially introduced by Attanasi et al. (2014) and further developed by Attanasi et al. (2018), which expands the probability range. This version includes 20 decision rows (referred to as MPL20; see Table 9 in Appendix A, the last four columns are not presented to the experimental subjects.). The choice of two MPL methods employed in this study, made possible to observe subjects' behaviour either when probabilities are presented in 10 and 5 percentage point. Nonetheless, the popularity of this EM fosters our choice to employ this methodology, among others. We compare the number of safe choices made by participants before and after the learning periods. This analysis allows us to assess how risk preferences evolve with experience and how different elicitation methods capture these changes.

Table 1
Sample composition.

Sequence	First EM	Second EM	Number of subjects	Number of observations
1	EG	MPL10	26	45
2	MPL10	EG	24	44
3	CH	MPL20	25	42
4	MPL20	CH	22	38
EG			41	41
CH			47	47
MPL10			42	42
MPL20			39	39
Total			97	169

3. Experimental design

The experiment was conducted in May 2024 at the ESSE laboratory at the university of Bari Aldo Moro. We recruited 97 participants, of which 72 correctly performed the two randomly assigned EMs, while 25 only correctly performed one. Our participants are undergraduate economics students aged between 19 and 22, with 51% of females. This allows us to have a total of 169 observations.² We present in Table 1 the number of subjects and the number of valid observations for each sequence. The experiment is carried out paper-and-pencil as presented by Charness et al. (2023), in a controlled laboratory environment.

To avoid order and learning effects across the two EMs (see, Carlsson et al., 2012), each participant is randomly assigned to one of our four sequences described in Table 1. The order in which subjects face the two EMs might affect their risky choice before and after the learning periods; we present a Kolmogorov–Smirnov and a Proportion test in Table 2. The results do not indicate any systematic effects attributable to the task order; this suggests that it does not have any effect whether the subjects start with the *easier to understand* (CH or EG) or *more complex task* (MPL10 or MPL20). However, our within-subjects design may include cross EMs contamination effects and portfolio-building (see, Cubitt et al. (1998), and Harrison and Rutström (2008)) a random lottery incentive system was implemented, i.e., only one risky choice among initial and final of one of the two EMs was randomly chosen for a subject's final payment (Azrieli et al., 2018). Our analysis highlights that changes in preferences over chosen gambles are not influenced by our within-subject experimental design.

Following Charness et al. (2023), experimental subjects were accurately informed by the experimenter about the tables of gambles, with the relative EM scheme for Options A and B. The OLS design, characterised by a 50%–50% probability gamble, has a simple distinction between even(odd) results from a dice rolled, while the MPL design ranges from 5% to 100% probabilities (see supplementary material). We used twenty-face dice for the hands-on learning experience, providing one dice for a subject. During the experiment, the experimenter reads and explains the instructions to the subjects. The experiment is composed of three different phases:

- **Initial Risky Choice.** In the first phase, subjects, after receiving and reading the instructions, are asked to make a choice in order to elicit their first measure of risk preference, according to the assigned EM;

² Each subject in our design provides two observations: one for OLS and one for MPL. Therefore, the difference between twice the number of subjects, and the number of observations, are the incomplete observations that we exclude from our study (equal to 25). All the calculations are performed with an appropriate sample size with type I error $\alpha = 0.10$, effect size $d = 0.5$, and Power $(1 - \beta) = 0.80$. Calculations performed through G*Power (see Faul et al., 2009).

- **Learning phase.** In the second phase consists of 24 unpaid hands-on learning periods. These 24 periods are characterised by 12 periods with pre-selected choices (same for all subjects within EM); here, each person rolled a dice and reports on a record sheet the dice outcome and the would-be payoff, a wrong reported outcome and would-be payoff is considered a mistake (error). However, for the 12 free choices, each subject chooses the preferred gamble(s) and follows all the previous steps (See Appendix A.3).
- **Final Risky Choice.** In the third phase, subjects are asked, as made in the first phase, to make their choice with the same EM.

After these three phases, each subject comes in front of the desk, rolls the dice to determine which of the two EM is valid for payment, then rolls the dice to determine if the first or the third phase will be paid, and finally, rolls the dice to determine the gamble outcome. We paid subjects if they had correctly reported the outcome and would-be payoff, a randomly selected learning period is drawn for the payment of 1 euro.

For each subject, one of the two EMs was randomly selected for payment, then one between the initial and final risky choices was randomly selected to determine the real payment. This double randomisation avoid the hedging effect, with the average payment being approximately 5 euros.³ The payoff in our elicitation tasks is expressed in ECU. The choice of the payoff schemes aims to replicate the elicitation methods presented by Eckel and Grossman (2008) due to the popularity in the literature (EG), as well as a variation to the previous EM discussed by Charness et al. (2023) CH. Whereas for MPL (Attanasi et al., 2014, MPL10 and MPL20). The choice to keep constant the payoff for the two MPL elicitation tasks makes it possible to observe the subjects' behaviour when the probability is framed in ten and five percentage points. We vary the exchange rate ECU-Euro for each EM in order to avoid higher variation in the lowest and maximum possible payoff.⁴

4. Results

In this section, we present summary statistics and non-parametric tests to analyse changes in risk preferences across the selected EMs. We present results for consistency (or not) for preference towards risk based on the categorisation proposed by Brunette and Ngouhou-Poufoun (2022).

4.1. Individual's risky choice changes

Fig. 1 shows, for each elicitation method, the distribution of initial (black) and final (grey) choices after the 24 periods of learning. We observe evidence for an increasing risk aversion. The direction of the changes in preference over risk is driven substantially by our sample composition, which reports lower initial willingness to take risks compared with the results presented by Charness et al. (2023) (see Table 10 in Appendix A). The learning periods allow subjects to know better their own preference, which, in Charness's result, suggests a final preference towards riskier gambles, while in our case, the results suggest a final preference towards less risky gambles.

We observe significant differences in risky choice between the two experimental phases for the EG task. In line with Charness et al. (2023), we observe a movement to higher numbers, a binomial test gives $z = 1.668$, $p = 0.046$, one-tailed test. However, for both MPL EMs, we do not detect statistically significant differences in risky choices between

³ Those subjects who correctly complete only one task are paid on the randomisation between the initial and final risky choice.

⁴ Here we report the exchange rate for each EM, CH(€0, 15), EG(€0, 20), MPL10 and MPL20(€0, 45).

Table 2
Test order and learning effect by sequence.

Elicitation method	Kolmogorov–Smirnov				Proportion test	
	Initial risky choice		Final risky choice		Lottery changes by sequence	
	Statistic (D)	p-value	Statistic (D)	p-value	Statistic (z)	p-value
CH	0.1175	0.999	0.1000	1.000	-1.156	0.247
EG	0.1099	0.999	0.1575	0.935	0.749	0.453
MPL10	0.2334	0.622	0.2151	0.721	-0.138	0.889
MPL20	0.1203	0.999	0.0936	1.000	0.913	0.361

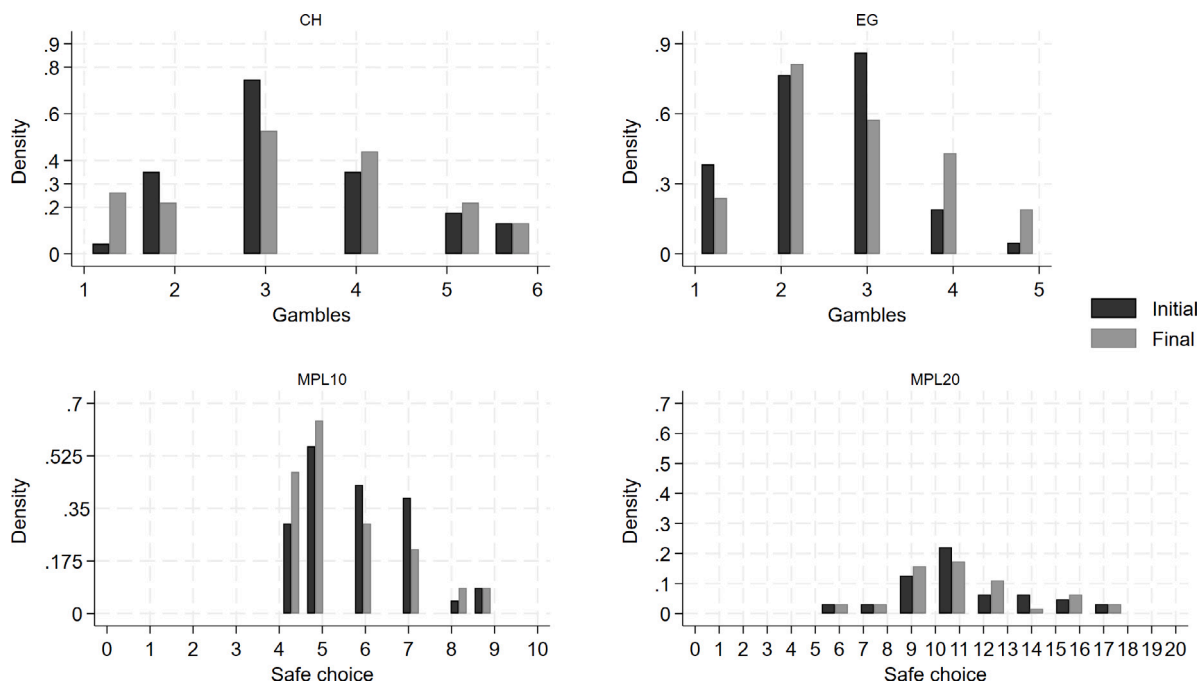


Fig. 1. Density gambles before and after learning.

the initial and final phases. In the two MPLs, the horizontal axis measures the number of safe choices (higher values indicate greater risk aversion). In MPL10, the mass re-centres slightly towards 4–5 at the expense of 6–7, suggesting a slight shift towards less risk aversion; In MPL20, however, a hint of a shift to the right can be seen (more safe choices around 11–13), consistent with greater aversion after the learning periods.

We present non-parametric tests on risk preference changes at the individual level and the correlation between the two risky choices in Table 3. The Spearman correlation coefficient between the initial and final risky choice indicates the effectiveness of the hands-on learning periods. The correlation reported by Charness et al. (2023) was 0.546; in our case, it is lower for both the OLS methods. In our opinion, this highlights a higher effect of the learning periods between the two choices. In contrast, this effect decreases moving from the OLS to the MPL method, which reports a higher correlation, disclosing the weaker effect that the hands-on learning period exerts on the final choice.

In Table 3, we report the results of two test statistics; first, as in Charness et al. (2023), we perform a Kolmogorov–Smirnov test of cumulative distribution for all our EMs. Second, we present the result of a Wilcoxon signed-rank, which is the more appropriate in our case, where we analyse two measurements on the same subject (before and after hands-on learning periods.). Given the experimental design, we treat the Wilcoxon signed-rank test as primary. It shows no systematic change between initial and final choices for CH, MPL10, and MPL20; for EG we find a marginal effect ($z = -1.944$, $p = 0.051$, two-tailed), consistent with Fig. 1’s descriptive shift towards riskier options. A Kolmogorov–Smirnov comparison of marginal distributions suggests no

differences for any EM, as expected, since K-S ignores pairing and has lower power with discrete outcomes.

In Fig. 2, we show the proportion of subjects in each EM that change risky choice after the learning experience. We analyse the consistency of preferences towards risk, following three out of four categories among those presented by Brunette and Ngouhou-Poufoun (2022). The first category refers to those subjects disclosing “consistent preferences” towards risk. Subjects in this category do not change their risky choice (gamble for the OLS method or a number of safe choices for MPL), which leads us to observe subjects with a consistent preference towards risk. The second category is called “strong inconsistency”; here, we include those subjects that switched from risk aversion to risk loving after the learning period. In other words, this category collects individuals who completely reversed their preference (disclosing contradictory shape of their utility function over money). The third category, called “weak inconsistency”, refers to those subjects that change preference from risk aversion to risk neutrality or from risk loving to risk neutrality, and vice versa. The black colour in Fig. 2 represents the percentage of subjects for each EM that falls into the category “Consistent preferences”. Our second category looks at those subjects that switch from a choice indicating risk aversion to a risk-loving preference. In our whole sample, we observe such strong changes in preference over risk for a few subjects in the CH EM (only 3 of 21 changes, equal to 14.28%). In contrast, all the other subjects disclosing changes in choice after the learning periods fall into the third category. Subjects’ behaviour in this category suggests that although after the hands-on learning periods, there is a change in choice over risk, the subject’s preference represented by the shape of the utility

Table 3
Initial and Final risky choice after observing the realisations.

Elicitation method	Wilcoxon signed-rank		Kolmogorov–Smirnov		Spearman	
	Statistic (z)	p-value	Statistic (D)	p-value	Statistic (ρ)	p-value
CH	0.187	0.851	0.1220	0.921	0.3403	0.0298
EG	-1.944	0.051	0.1702	0.504	0.3697	0.0109
MPL10	1.468	0.142	0.1429	0.785	0.6932	0.0000
MPL20	0.180	0.857	0.0513	1.000	0.7271	0.0000

Note: The reported p-values are two-tailed test.

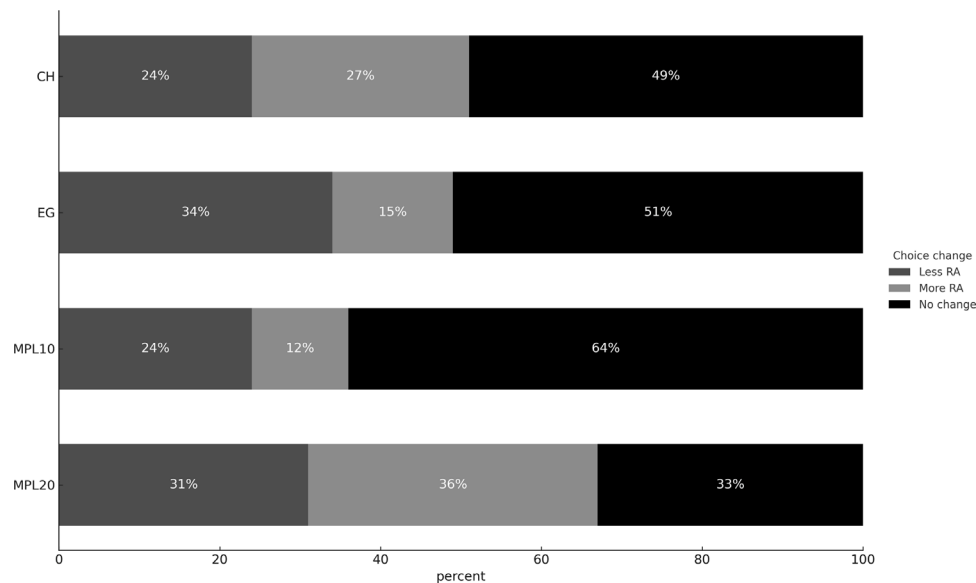


Fig. 2. Choice changes.

function remains pretty much the same or, at most, changes by one step. In other words, a subject in the CH first selects Gamble 3 and then Gamble 1, where both choices denote risk-averse preference, while a subject who initially chooses Gamble 5 and then Gamble 4 discloses a change passing from risk-neutral to risk-averse preference. While we observe no significant differences in the proportion of subjects that change preference for the OLS tasks (in line with Charness et al., 2023), the same does not hold when looking at the MPL tasks. A proportion test for MPL10 gives $z = 2.61$ $p = 0.0044$ and MPL20 gives $z = -2.94$ $p = 0.0016$, one-tailed tests. However, our interest lies in the direction of the change if the subject discloses more or less willingness to take risks. Table 4 displays the findings from proportion Z-tests and binomial tests, which evaluate changes in risk aversion (RA) across four experimental techniques (CH, EG, MPL10, MPL20), categorised by whether they signify increased or decreased risk aversion. The analyses were performed using a significance threshold $\alpha = 0.10$, an 80% target power ($\beta = 0.20$), and a medium effect size (Cohen’s $h = 0.5$). The analysis reveals that several conditions yield statistically significant results, particularly in the “Less RA” direction for CH, MPL10, and MPL20, as well as in the “More RA” direction for CH, EG, and MPL10. Notably, the p-values for CH and MPL10 in both directions are consistently below 0.005, accompanied by high observed power levels (> 0.92), suggesting robust evidence for behavioural shifts in risk preferences under these decision elicitation methods. However, the power analysis reveals some critical limitations due to sample size. For example, the condition EG - Less RA shows a p-value of 0.0287, which is statistically significant at the 10% level, yet the power (0.7198) falls short of the conventional threshold, implying an elevated risk of Type II error. More pronounced is the case of MPL20 - More RA, where the p-value (0.0782) is marginal, and the observed power is particularly low (0.5563), indicating insufficient sensitivity to detect an effect of the expected magnitude with the current sample. The closely matching outcomes of Z-tests and binomial

tests contribute to the internal validity of the analysis. However, given that the sample sizes are between 39 and 47 for each condition, these results must be approached with caution. Several effects could lack sufficient power, and marginal significance should not be overemphasised unless supported by adequate statistical power. In summary, while the data provide initial support for method-induced shifts in risk aversion, consistent with behavioural economic theories on framing and elicitation effects, future studies should replicate these findings with larger samples to ensure the reliability and generalisability of the observed effects.

This analysis supports the hypothesis that, in each experimental condition, there are systematic tendencies towards more risk-averse choices, with sufficient statistical power ensured by the sample sizes and expected effect size. However, we observe a greater reduction in risk aversion in those EMs with a lower complexity. Within the domain of decision-making under risk, complexity is characterised by the number of potential payoff outcomes in a given lottery (see Georgalos & Nabil, 2025).

4.2. Learning periods and final choice

A common belief in the “hot hand” and the “law of averages” (see Bennett et al., 2010; Gallistel, 2012; Miller & Sanjurjo, 2018) might bring expectations of positive or negative serial correlation in dice rolls. However, even if these effects are present, as reported by Charness et al. (2023), there is no evidence that the change in risk preference is driven by the results observed during the unpaid learning periods. Fig. 3 and Table 11 (see Appendix A) shows a Kolmogorov–Smirnov test of the difference between the final and initial risky choice depending on whether more or less than half of the practice rolls were favourable. There is no evidence that the outcomes of the practice rolls drive change in risk preference.

Table 4

Power calculations are executed assuming $\alpha = 0.10$, $\beta = 0.20$, Effect size (Cohen's $h = 0.5$).

EM	Sample	Direction	Count	Z-test proportion			Binomial test		
				p-value	Z stat	Power	p-value	Z stat	Power
CH	41	Less RA	10	0.0010	-3.2796	0.9639	0.0015	-3.2796	0.9639
EG	47	Less RA	16	0.0287	-2.1880	0.7198	0.0400	-2.1880	0.7198
MPL10	42	Less RA	10	0.0007	-3.3947	0.9731	0.0009	-3.3947	0.9731
MPL20	39	Less RA	12	0.0163	-2.4019	0.7941	0.0237	-2.4019	0.7941
CH	41	More RA	11	0.0030	-2.9673	0.9251	0.0043	-2.9673	0.9251
EG	47	More RA	7	0.0000	-4.8135	0.9999	0.0000	-4.8135	0.9999
MPL10	42	More RA	5	0.0000	-4.9377	1.0000	0.0000	-4.9377	1.0000
MPL20	39	More RA	14	0.0782	-1.7614	0.5563	0.1081	-1.7614	0.5563

Results of Z-tests and binomial tests with observed powers for both increases and decreases in risk aversion.

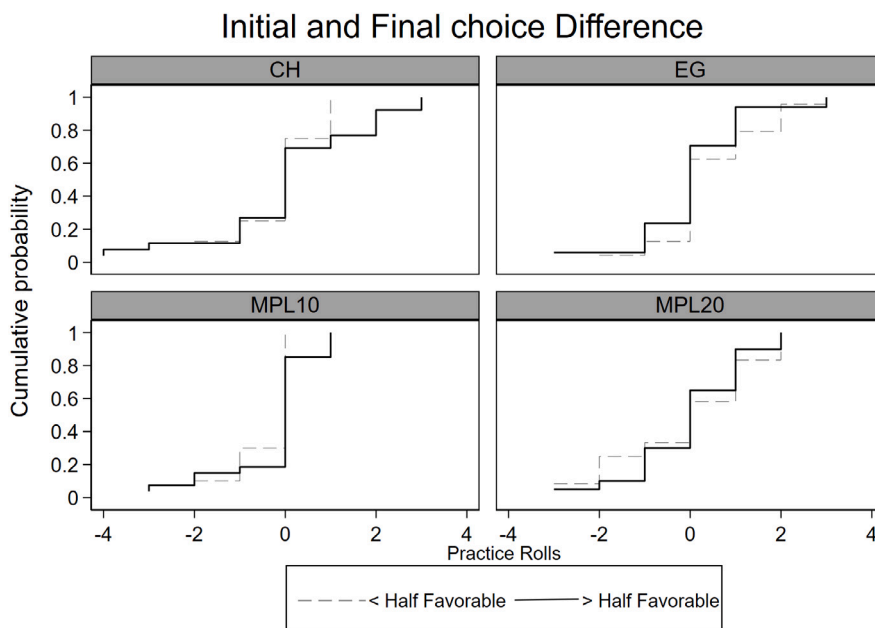


Fig. 3. Difference in risk preference before and after practice, by favourable realisations. We exclude observations with exactly 12 realisations.

4.3. Method complexity and final choice

The methods presented are characterised, as reported by Charness et al. (2013), by different complexity: on the one hand, we have a one-choice risk EM (OLS) where subjects are asked to choose the preferred gamble over several gambles. However, on the other hand, we have a series of choices between paired gambles (MPL) where subjects are asked to make one choice for each pair, and then only one row is drawn and played to determine the payment.

Charness et al. (2013) suggests that one of the main disadvantages of the complex method is the fact that subjects do not understand the procedure, Dave et al. (2010) in their results show that participants with difficulties in understanding MPL task with varying probabilities are those with a low level of numeracy. Furthermore, their recent contribution (Charness et al., 2023; Willadsen et al., 2024) highlight the interplay between the results of the Cognitive Reflection Test and willingness to take risks with their positive relation, Nielsen and Rehbeck (2022) in addition provide evidence of how the likelihood of mistakes in incentivised choice experiments might be affected by the heterogeneity in cognitive abilities biasing the estimated subjects preference. Fig. 4 shows CDFs of the initial–final choice difference (positive value = shift to riskier options in OLS, while the opposite holds for MPL). Distributions cluster near zero in all tasks. In MPL20, the subjects who make mistakes⁵ lies to the right of non-reporting mistakes,

⁵ We refer to the number of wrong reported dice outcomes and would-be payoff among the 24 periods.

indicating greater aversion towards risk; a Kolmogorov–Smirnov test rejects equality only in MPL20⁶ ($D = 0.443$; $p = 0.030$ one-tailed test). In EG, the visual shift is not statistically significant. In CH and MPL10, the CDFs largely overlap.

Table 5 presents linear regressions exploring how errors, elicitation mechanisms, and individual characteristics affect final risky choices and the standardised difference from initial preferences. The analysis compares two elicitation methods—OLS and MPL—and captures behavioural adjustments following a learning phase. A key insight concerns the *Errors* variable. In the MPL condition, higher error rates significantly predict a higher number of safe choices at the end of the experiment ($p < 0.01$). This suggests that participants who struggled to understand the task during the learning phase became more risk-averse. This relationship does not emerge under the OLS condition, highlighting how task comprehension plays a more critical role when the elicitation method is cognitively demanding. Gender effects also differ by elicitation method. In the MPL regression, the *Female* coefficient is positive and significant, indicating that women made more risk-averse choices. This result is consistent with established findings in the behavioural economics literature, which show that females tend to exhibit greater risk aversion than males (Charness & Gneezy, 2012; Croson & Gneezy, 2009). Notably, this gender effect is absent under

⁶ Table 12 in Appendix A, reports the test for all the EMs on a two-tailed p -value.

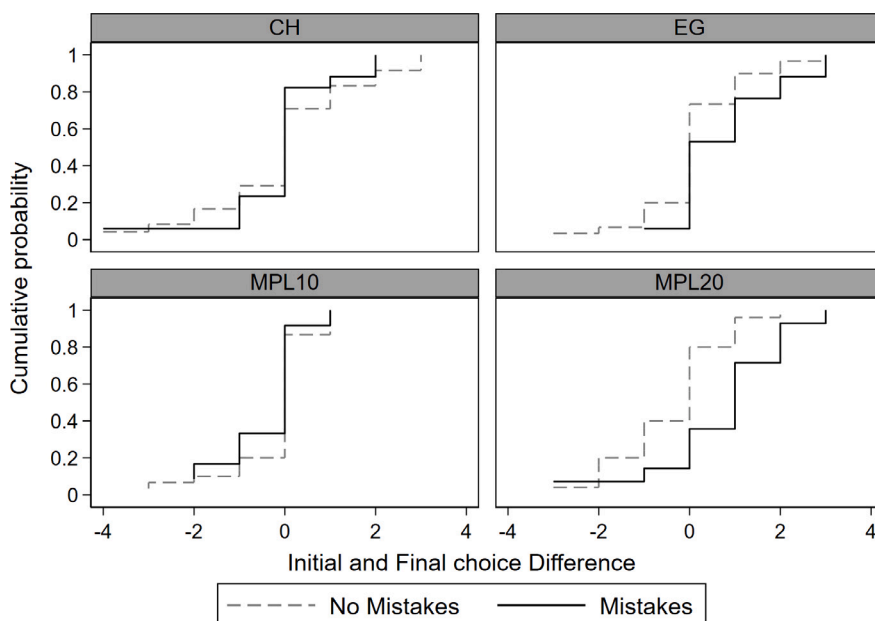


Fig. 4. Cumulative distribution function difference in choice by mistake.

Table 5
Linear regression comparing OLS and MPL.

	<i>Dependent variables:</i>			
	Final choice OLS	Final choice MPL	Std Diff OLS	Std Diff MPL
Female	0.119 (0.267)	0.0110** (0.00414)	0.0154 (0.0594)	0.000537 (0.000273)
Errors	-0.00400 (0.0241)	0.0668** (0.0244)		
Unfavourable outcomes	-0.230 (0.270)	0.0719 (0.256)	-0.0588 (0.0611)	0.00652 (0.0193)
Initial choice	0.414** (0.141)	0.875*** (0.0435)	-0.129*** (0.0307)	-0.00874** (0.00320)
EM= EG	-0.181 (0.289)		-0.0344 (0.0655)	
EM= MPL20		0.850* (0.358)		0.0731* (0.0289)
Standardised errors			-0.0106 (0.138)	0.0851* (0.0326)
Constant	1.984*** (0.544)	0.316 (0.298)	0.450*** (0.122)	0.0105 (0.0222)
Observations	88	81	88	81
R ²	0.167	0.905	0.236	0.113
Adjusted R ²	0.116	0.899	0.190	0.053

Robust standard errors in parentheses are clustered at individual level.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

OLS, suggesting that the structure of the elicitation method together with the learning experience, may moderate its expression. Furthermore, this methodological contrast aligns with a 12-week panel using both MPL and OLS that finds no aggregate gender gap and shows that experience and stakes damp observable differences (Bandyopadhyay et al., 2021). Nonetheless, the new learning methodology employed in this paper and the use of different EMs sheds light on a new research avenue to investigate the role of the different individual-specific traits (i.e, cognitive ability and gender) with risk preference.

5. Conclusion

This paper addresses a central question in experimental economics: are observed changes in risk preferences genuine, or do they reflect the elicitation method used? This question directly connects to the Discovered Preference Hypothesis, which posits that feedback and experience help subjects discover stable tastes rather than transform them (Plott, 1996). Building on the within-subject design by Charness et al. (2023), we compare two risk elicitation tools Ordered Lottery Selection

(OLS) and Multiple Price List (MPL) to assess how task complexity and hands-on learning influence risk preferences. Our findings support the view that risk preferences, as elicited in laboratory settings, are sensitive to experience and the structure of the elicitation method. While changes after the learning phase are modest under OLS, they are more pronounced under MPL, particularly when complexity increases. Participants who reported errors during MPL learning were significantly more likely to become risk-averse in their final choice, indicating that task comprehension plays a critical role in shaping decisions. These patterns are consistent with evidence on numeracy and decision-making under cognitive load (Dave et al., 2010; Holzmeister & Stefan, 2021; Nielsen & Rehbeck, 2022). Yet we acknowledge, in line with Becker and Stigler’s critique (Stigler & Becker, 1977), that changes in choice behaviour may not imply changes in underlying preferences. In the spirit of the DPH (Plott, 1996), the learning process may reveal latent preferences rather than transform them. In our case, what we label as “preference shifts” could be better understood as changes in the expression of preferences due to increased familiarity with the task. This interpretation leaves room for a richer, more stable preference framework in which learning influences choices without altering the fundamental utility function. Moreover, we recognise that repeated exposure to the same task, independent of learning, may itself lead to behavioural variation. Known effects such as regression to the mean, decision fatigue, or a preference for change can induce shifts even without new information. While our design does not include a “pure repetition” control condition, the strong correlation between initial and final choices (particularly under MPL, $\rho = 0.69$ for MPL10 and $\rho = 0.72$ for MPL20) suggests that initial rankings remain informative. Future research should include a baseline condition without learning to isolate changes due purely to task repetition. The contribution of this paper lies in its comparative use of different EMs under a unified experimental structure. We show that complexity and comprehension are not neutral features of design but active forces in shaping elicited behaviour. These findings invite caution in using elicited risk preferences as explanatory variables in other domains, particularly when cognitive demands differ across methods. Further research should examine how robust these findings are in real-world contexts, across diverse populations, and with longer time horizons. Exploring how stable individual rankings remain across multiple tasks, and how methodological framing influences observed preferences, will help refine how we interpret laboratory measures of risk attitude.

CRedit authorship contribution statement

Rocco Caferra: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Andrea Morone:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Donato Pierno:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Ethical statement

This manuscript complies with the Ethical Rules of this journal.

Appendix A. Tables

A.1. Elicitation methods

See Tables 6–9.

A.2. Results

See Tables 10–12.

A.3. Learning

Here we present some tables drawn from phase two of the experiment, the learning phase.

Subsection	Sequence of actions in each period
Pre-selected elections (12 periods)	1. Review the selected gamble; 2. Roll the dice to determine if the output is an even or odd number; 3. Record the corresponding payment;
Free elections (12 periods)	1. Decide between the five gambles; 2. Roll the dice to determine if the output is an even or odd number; 3. Record the corresponding payment;

Pre-selected elections

In the next table, the lotteries are listed in each period, and it is shown which of the 5 lotteries has been pre-selected for you by the experimenter. Please note which event occurred (even or odd sum) and note the payment. You can refer to the table in front of the room where the payoffs are displayed for each bet. Remember that the result of this section will not be used for payment:

Period	Pre-selected gamble					Realised event		Payoff
1	1	2	3	4	5	Odd	Even	
2	1	2	3	4	5	Odd	Even	
3	1	2	3	4	5	Odd	Even	
4	1	2	3	4	5	Odd	Even	
5	1	2	3	4	5	Odd	Even	
6	1	2	3	4	5	Odd	Even	
7	1	2	3	4	5	Odd	Even	
8	1	2	3	4	5	Odd	Even	
9	1	2	3	4	5	Odd	Even	
10	1	2	3	4	5	Odd	Even	
11	1	2	3	4	5	Odd	Even	
12	1	2	3	4	5	Odd	Even	

Free elections

In the next table, write down the decision you make each period, the event realised (odd or even sum), and write the payment. The table in front of the room displays the payoffs corresponding to each gamble and event. Remember that the result of this section will not be used for payment:

Period	Pre-selected gamble					Realised event		Payoff
1	1	2	3	4	5	Odd	Even	
2	1	2	3	4	5	Odd	Even	
3	1	2	3	4	5	Odd	Even	
4	1	2	3	4	5	Odd	Even	
5	1	2	3	4	5	Odd	Even	
6	1	2	3	4	5	Odd	Even	
7	1	2	3	4	5	Odd	Even	
8	1	2	3	4	5	Odd	Even	
9	1	2	3	4	5	Odd	Even	
10	1	2	3	4	5	Odd	Even	
11	1	2	3	4	5	Odd	Even	
12	1	2	3	4	5	Odd	Even	

Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.socec.2025.102450>.

Data availability

The author confirms that all data generated or analysed during this study are included in this published article. Furthermore, primary and secondary sources and data supporting the findings of this study were all available at the time of submission.

Table 6
CH risk EM.

Choice (50/50 Gamble)	Event B	Event A	Expected return	Standard deviation	Implied CRRA range	Fraction of Subjects (%)
Gamble 1	28	28	28	0	$r > 3.46$	6,06
Gamble 2	24	36	30	6	$1.16 < r < 3.46$	11,11
Gamble 3	20	44	32	12	$0.71 < r < 1.16$	33,33
Gamble 4	16	52	34	18	$0.50 < r < 0.71$	18,18
Gamble 5	12	60	36	24	$-0.22 < r < 0.50$	17,17
Gamble 6	2	66	34	32	$r < -0.22$	14,14

Elicitation method CH, [Charness et al. \(2023\)](#) gamble choices.

Table 7
EG risk EM.

Choice (50/50 Gamble)	Event B	Event A	Expected return	Standard deviation	Implied CRRA range	Fraction of Subjects (%)
Gamble 1	16	16	16	0	$r > 2$	4,30
Gamble 2	12	24	18	6	$0.67 < r < 2$	16,40
Gamble 3	8	32	20	12	$0.38 < r < 0.67$	32,38
Gamble 4	4	40	22	18	$0.20 < r < 0.38$	22,70
Gamble 5	0	48	24	24	$r < 0.20$	23,80

Elicitation method EG, [Eckel and Grossman \(2008\)](#), gamble choices.

Table 8
MPL10 risk EM.

Decision	Option A	Option B	Expected payoff difference	Implied CRRA range	Number of safe choices	Proportion of choices (%)
1	20/20 of 12; 0/20 of 10	20/20 of 22; 0/20 of 0,50	10,00	$-\infty < r \leq -2.010$	0-1	1,00
2	18/20 of 12; 2/20 of 10	18/20 of 22; 2/20 of 0,50	8,05	$-0.878 \leq r \leq -0.648$	2	1,00
3	16/20 of 12; 4/20 of 10	16/20 of 22; 4/20 of 0,50	6,10	$-0.450 \leq r \leq -0.273$	3	6,00
4	14/20 of 12; 6/20 of 10	14/20 of 22; 6/20 of 0,50	4,15	$-0.112 \leq r \leq 0.038$	4	26,00
5	12/20 of 12; 8/20 of 10	12/20 of 22; 8/20 of 0,50	2,20	$0.180 \leq r \leq 0.317$	5	26,00
6	10/20 of 12; 10/20 of 10	10/20 of 22; 10/20 of 0,50	0,25	$0.454 \leq r \leq 0.592$	6	23,00
7	8/20 of 12; 12/20 of 10	8/20 of 22; 12/20 of 0,50	-1,70	$0.736 \leq r \leq 0.891$	7	13,00
8	6/20 of 12; 14/20 of 10	6/20 of 22; 14/20 of 0,50	-3,65	$1.068 \leq r \leq 1.287$	8	3,00
9	4/20 of 12; 16/20 of 10	4/20 of 22; 16/20 of 0,50	-5,60	$1.613 \leq r < +\infty$	9-10	1,00
10	2/20 of 12; 18/20 of 10	2/20 of 22; 18/20 of 0,50	-7,55			

Elicitation method MPL10, [Holt and Laury \(2002\)](#) gamble choices.

Table 9
MPL20 risk EM.

Decision	Option A	Option B	Expected payoff difference	Implied CRRA range	Number of safe choices	Proportion of safe choices
1	20/20 of 12; 0/20 of 10	20/20 of 22; 0/20 of 0.50	10.00	$-\infty < r \leq -2.010$	0-1	1.00
2	19/20 of 12; 1/20 of 10	19/20 of 22; 1/20 of 0.50	9.03	$-2.010 \leq r \leq -1.512$	2	0.00
3	18/20 of 12; 2/20 of 10	18/20 of 22; 2/20 of 0.50	8.05	$-1.512 \leq r \leq -1.156$	3	1.00
4	17/20 of 12; 3/20 of 10	17/20 of 22; 3/20 of 0.50	7.08	$-1.156 \leq r \leq -0.878$	4	0.00
5	16/20 of 12; 4/20 of 10	16/20 of 22; 4/20 of 0.50	6.10	$-0.878 \leq r \leq -0.648$	5	1.50
6	15/20 of 12; 5/20 of 10	15/20 of 22; 5/20 of 0.50	5.13	$-0.648 \leq r \leq -0.450$	6	0.05
7	14/20 of 12; 6/20 of 10	14/20 of 22; 6/20 of 0.50	4.15	$-0.450 \leq r \leq -0.273$	7	2.50
8	13/20 of 12; 7/20 of 10	13/20 of 22; 7/20 of 0.50	3.18	$-0.273 \leq r \leq -0.112$	8	3.00
9	12/20 of 12; 8/20 of 10	12/20 of 22; 8/20 of 0.50	2.20	$-0.112 \leq r \leq 0.038$	9	8.00
10	11/20 of 12; 9/20 of 10	11/20 of 22; 9/20 of 0.50	1.23	$0.038 \leq r \leq 0.180$	10	7.50
11	10/20 of 12; 10/20 of 10	10/20 of 22; 10/20 of 0.50	0.25	$0.180 \leq r \leq 0.317$	11	1.00
12	9/20 of 12; 11/20 of 10	9/20 of 22; 11/20 of 0.50	-0.73	$0.317 \leq r \leq 0.454$	12	5.50
13	8/20 of 12; 12/20 of 10	8/20 of 22; 12/20 of 0.50	-1.70	$0.454 \leq r \leq 0.592$	13	4.00
14	7/20 of 12; 13/20 of 10	7/20 of 22; 13/20 of 0.50	-2.68	$0.592 \leq r \leq 0.736$	14	12.00
15	6/20 of 12; 14/20 of 10	6/20 of 22; 14/20 of 0.50	-3.65	$0.736 \leq r \leq 0.891$	15	20.50
16	5/20 of 12; 15/20 of 10	5/20 of 22; 15/20 of 0.50	-4.63	$0.891 \leq r \leq 1.068$	16	13.00
17	4/20 of 12; 16/20 of 10	4/20 of 22; 16/20 of 0.50	-5.60	$1.068 \leq r \leq 1.287$	17	4.50
18	3/20 of 12; 17/20 of 10	3/20 of 22; 17/20 of 0.50	-6.58	$1.287 \leq r \leq 1.613$	18	8.50
19	2/20 of 12; 18/20 of 10	2/20 of 22; 18/20 of 0.50	-7.55	$1.613 \leq r < +\infty$	19-20	6.00
20	1/20 of 12; 19/20 of 10	1/20 of 22; 19/20 of 0.50	-8.53			

Elicitation method MPL20, [Attanasi et al. \(2014\)](#) gamble choices.

Table 10
Summary statistics initial and final risky choice.

Elicitation method	Initial choice				Final choice			
	Mean	Median	Mode	Standard deviation	Mean	Median	Mode	Standard deviation
CH	3.36	3	3	1.19	3.29	3	3	1.43
EG	2.44	2	3	0.95	2.78	3	2	1.14
MPL10	5.76	6	5	1.30	5.47	5	5	1.38
MPL20	11	10	10	2.90	11	10	9	2.86

Summary statistics before and after observing the realisations^a.

^a The sample composition of Charness et al. (2023) in footnote 7 disclose higher initial and final average gamble value. The reported Standard deviations are not statistically different between initial and final risky choice.

Table 11
Cumulative distribution function initial and final risky choice by favourable realisations.

Kolmogorov–Smirnov test		
Elicitation method	Statistic (D)	<i>p</i> -value
CH	0.230	0.900
EG	0.1495	0.979
MPL10	0.148	0.997
MPL20	0.150	0.996

Test difference in initial and final risky choice by favourable realisations.

Table 12
Cumulative distribution function initial and final risky choice by errors.

Kolmogorov–Smirnov test		
Elicitation method	Statistic (D)	<i>p</i> -value
CH	0.1152	1.000
EG	0.2039	0.758
MPL10	0.1333	0.998
MPL20	0.4429	0.059

Tests difference in initial and final risky choice by error.

References

- Attanasi, G., Corazzini, L., Georgantzis, N., & Passarelli, F. (2014). Risk aversion, overconfidence and private information as determinants of majority thresholds. *Pacific Economic Review*, 19(3), 355–386.
- Attanasi, G., Georgantzis, N., Rotondi, V., & Vigani, D. (2018). Lottery-and survey-based risk attitudes linked through a multichoice elicitation task. *Theory and Decision*, 84, 341–372.
- Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, 126(4), 1472–1503.
- Bandyopadhyay, A., Begum, L., & Grossman, P. J. (2021). Gender differences in the stability of risk attitudes. *Journal of Risk and Uncertainty*, 63(2), 169–201.
- Bennett, C., Baird, A., & Miller, M. (2010). L, WG (2010). Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: an argument for proper multiple comparisons correction. *Journal of Serendipitous Unexpected Results*, 1, 1–5.
- Binswanger, H. P. (1980). Attitudes toward risk: Experimental measurement in rural India. *American Journal of Agricultural Economics*, 62(3), 395–407.
- Binswanger, H. P. (1981). Attitudes toward risk: Theoretical implications of an experiment in rural India. *The Economic Journal*, 91(364), 867–890.
- Bradbury, M. A. S., Hens, T., & Zeisberger, S. (2015). Improving investment decisions with simulated experience. *Review of Finance*, 19(3), 1019–1052.
- Brunette, M., & Ngouhou-Poufoun, J. (2022). Are risk preferences consistent across elicitation procedures? A field experiment in Congo basin countries. *The Geneva Risk and Insurance Review*, 47(1), 122–140.
- Carlsson, F., Mørkbak, M. R., & Olsen, S. B. (2012). The first time is the hardest: A test of ordering effects in choice experiments. *Journal of Choice Modelling*, 5(2), 19–37.

- Charness, G., Chemaya, N., & Trujano-Ochoa, D. (2023). Learning your own risk preferences. *Journal of Risk and Uncertainty*, 67(1), 1–19.
- Charness, G., & Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior and Organization*, 83(1), 50–58.
- Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization*, 87, 43–51.
- Crosno, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2), 448–474.
- Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics*, 1, 115–131.
- Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty*, 41, 219–243.
- Delaney, J., Jacobson, S., & Moenig, T. (2020). Preference discovery. *Experimental Economics*, 23(3), 694–715.
- Eckel, C. C., & Grossman, P. J. (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior and Organization*, 68(1), 1–17.
- Ert, E., & Haruvy, E. (2017). Revisiting risk aversion: Can risk preferences change with experience? *Economics Letters*, 151, 91–95.
- Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, 133(4), 1645–1692.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- Gallistel, C. (2012). On rationalism and optimality: Responses to the Miller and Nevin commentaries. *Behavioural Processes*, 90(1), 87–88.
- Georgalos, K., & Nabil, N. (2025). Testing models of complexity aversion. *Journal of Behavioral and Experimental Economics*, 116, Article 102354.
- Harrison, G. W., & Rutström, E. E. (2008). Risk aversion in the laboratory. In *Research in experimental economics, Risk aversion in experiments* (pp. 41–196). Emerald Group Publishing Limited.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.
- Holzmeister, F., & Stefan, M. (2021). The risk elicitation puzzle revisited: Across-methods (in) consistency? *Experimental Economics*, 24, 593–616.
- Miller, L., Meyer, D. E., & Lanzetta, J. T. (1969). Choice among equal expected value alternatives: Sequential effects of winning probability level on risk preferences. *Journal of Experimental Psychology*, 79(3p1), 419.
- Miller, J. B., & Sanjurjo, A. (2018). Surprised by the hot hand fallacy? A truth in the law of small numbers. *Econometrica*, 86(6), 2019–2047.
- Nielsen, K., & Rehbeck, J. (2022). When choices are mistakes. *American Economic Review*, 112(7), 2237–2268.
- Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., & Rieskamp, J. (2017). The risk elicitation puzzle. *Nature Human Behaviour*, 1(11), 803–809.
- Plott, C. R. (1996). Rational individual behavior in markets and social choice processes: the discovered preference hypothesis. In *The rational foundations of economic behaviour* (pp. 225–250). McMillian.
- Reynaud, A., & Couture, S. (2012). Stability of risk preference measures: results from a field experiment on French farmers. *Theory and Decision*, 73, 203–221.
- Schildberg-Hörisch, H. (2018). Are risk preferences stable? *Journal of Economic Perspectives*, 32(2), 135–154.
- Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2), 76–90.
- Willadsen, H., Zaccagni, S., Piovesan, M., & Wengström, E. (2024). Measures of cognitive ability and choice inconsistency. *Journal of Economic Behavior and Organization*, 220, 495–506.