Complexity and Willingness to Accept-Willingness to Pay Gap*

Sara Arts† Jianying Qiu‡ Songfa Zhong§

Abstract

This study examines qualitatively and quantitatively the impact of the complexity of the elicitation mechanism and of the valuation of choice objects on preference revelation. Using a large representative sample of the Dutch population, we assess the extent to which complexity can help account for the willingness-to-accept (WTA) and willingness-to-pay (WTP) gap which is widely attributed to non-standard preferences such as loss aversion. We elicited from each participant the WTA-WTP gap for monetary lotteries, and measured the complexity of the price list elicitation mechanism and of the valuation of lotteries. We show that complexity measures systematically relate to the WTA-WTP gap of the lotteries and account for approximately 50% of the gap. Further, complexity explains a significant portion of the WTA-WTP gap across diverse subgroups, while loss aversion shows explanatory potential only among participants who experienced little complexity.

Keywords: willingness to accept-willingness to pay gap, revealed preference, complexity, heuristics, endowment effect, loss aversion

*We thank ODISSEI (Open Data Infrastructure for Social Science and Economic Innovations) in the Netherlands for supporting this work.

†Department of Economics, IMR, Radboud University, Email: sara.arts@ru.nl.

‡Department of Economics, IMR, Radboud University, Email: jianying.qiu@ru.nl.

§Department of Economics, Hong Kong University of Science and Technology; Department of Economics, National University of Singapore, Email: zhongsongfa@gmail.com.

1 Introduction

Revealing unobservable preferences from observable choices is critical in economic analyses. One significant challenge in the revealed preference approach is that decision-makers may find the choice situation too complex to make informed decisions. In the absence of effective ways to simplify the choice situation, complexity could lead to systemic biases and noise in decision-making, thus distorting the inference of underlying preferences.

Two forms of complexity have gained increasing attention. In the first form, decision makers may find some preference elicitation mechanisms too complex (referred to as mechanism complexity hereafter) (Plott and Zeiler, 2005, 2007; Cason and Plott, 2014; Smitizsky et al., 2021). Consequently, they may rely on heuristics to make decisions (Kahneman et al., 1982). Or, they may confuse one mechanism with another, for instance, misperceiving the Becker–DeGroot–Marschak mechanism (Becker et al., 1964) – a form of the second-price auction – as the first-price auction (Cason and Plott, 2014), or the deferred acceptance mechanism for the Boston mechanism (Rees-Jones and Shorrer, 2023). To overcome mechanism complexity, some researchers have sought to develop simpler mechanisms (Li, 2017), while others have emphasized the need to better describe the mechanisms (Gonczarowski et al., 2022).

In the second form of complexity, decision-makers may have difficulty forming a clear valuation for choice objects. We refer to this as valuation complexity. The idea that subjects may not clearly know the value of choice objects is longstanding. Von Neumann and Morgenstern (1944) comment on the completeness axiom that "it is very dubious, whether the idealization of reality which treats this postulate as a valid one, is appropriate or even convenient."[1] In response to valuation complexity, decision makers may behave cautiously by sticking to the default (Masatlioglu and Ok, 2005), postponing their decisions if possible (Danan and Ziegelmeyer, 2006), randomizing deliberately (Agranov and Ortoleva, 2017, 2023; Cettolin and Riedl, 2019; Halevy et al., 2023), or using less costly but error-prone valuation procedures (Oprea, 2022). While these two forms of complexity have been

[1] Closely related notions have been proposed in the literature. These include incomplete preference (Bewley, 2002; Masatlioglu and Ok, 2005; Eliaz and Ok, 2006; Ok et al., 2012; Nishimuray and Ok, 2018); preference imprecision (Dubourg et al., 1994; Butler and Loomes, 2007; Cubitt et al., 2015); and the recently developed notion of cognitive noise, according to which decision-makers' mental representation of the decision situation can be noisy and lead to a range of behavioral anomalies (Khaw et al., 2021; Woodford, 2020; Frydman and Jin, 2021; Vieider, 2021; Enke and Graeber, 2023; Oprea, 2022).

widely discussed, few studies have examined them jointly and quantified their behavioral consequences.

In this study, we qualitatively and quantitatively assess the importance of both forms of complexity in revealing preferences via the willingness-to-accept and willingness-to-pay gap (the WTA-WTP gap, henceforth)(Thaler, 1980; Knetsch, 1989; Kahneman et al., 1990). The WTA-WTP gap holds a central position in the field of behavioral economics and remains a subject of considerable debate (Ericson and Fuster, 2014). Whereas the WTA-WTP gap is often interpreted as a manifestation of reference dependence and loss aversion, it has been proposed and shown that subjects' misconception about the elicitation mechanism may be responsible for the WTA-WTP gap (Plott and Zeiler, 2005; Cason and Plott, 2014). It has also been hypothesized that valuation complexity may contribute to the WTA-WTP gap (Dubourg et al., 1994; Bayrak and Hey, 2020; McGranaghan and Otto, 2022; Cerreia-Vioglio et al., 2022). Concretely, when decision-makers are unsure about the valuation of an object, they act cautiously by bidding a low price when buying it and asking for a high price when selling it (as in, e.g., the buy-low and sell-high heuristic). This heuristic can thus result in the WTA-WTP gap. Building on these studies, we conduct a systematic investigation of the two forms of complexity in the WTA-WTP gap and show that they are critical in this ongoing debate.

We elicit the WTA and WTP of two monetary lotteries in a within-subject setting using the price list elicitation mechanism. To assess the complexity of the price list, we elicit the WTA and WTP of an object with a known preference: a sure payment of 2.5 euro (denoted as 500 points in the experiment). If subjects are loss averse and fully understand the incentives of the price list, they would exhibit a WTA-WTP gap for lotteries but not for the sure payment. On the other hand, if subjects find the price list complex and carry the heuristic of buy-low and sell-high in real life into the lab or misperceive the selected price in the list as the final trading price, they may have a WTA-WTP gap for both lotteries and the sure payment. Furthermore, the WTA-WTP gap for lotteries would correlate significantly with that for the sure payment.

We assess the valuation complexity of monetary lotteries by eliciting a range of prices instead of one precise price. Specifically, as in a price list, subjects face a series of choices between the (fixed) lottery and a (varying) price. Differently, in each choice, subjects have three options: the lottery, the sure payment, or the randomization option, where the computer flips a virtual coin to determine whether subjects receive the lottery or the sure payment. When subjects find the valuation of the lottery complex, they may have a range of prices for which they would prefer to choose the randomization option (Cettolin and Riedl, 2019; Agranov and Ortoleva, 2023; Halevy et al., 2023), and the range of prices can serve as a (probably conservative) measure of valuation complexity. Based on the hypothesis that valuation complexity and caution contribute to the WTA-WTP gap (Cerreia-Vioglio et al., 2022), we test whether subjects with a wider range of prices are more likely to exhibit a larger WTA-WTP gap.

We examine these two forms of complexity in a large-scale experiment with 1856 participants from the LISS (Longitudinal Internet studies for the Social Sciences) panel, a representative sample of the Dutch population. The LISS panel has been widely used for surveys and experiments in both economic policies and academic research (see e.g., Dimmock et al., 2016; Baillon et al., 2017; Cherchye et al., 2017). From each subject, we elicited the WTA and WTP for two monetary lotteries, measures of the two forms of complexity, and a measure of cognitive abilities using matrix reasoning questions from the International Cognitive Ability Resource (ICAR, Condon and Revelle, 2014).

We have four main observations. First, on average, subjects exhibit a WTA-WTP gap for both lotteries and the sure payment: The proportion with a positive (negative) WTA-WTP gap is 58% (27%) for the lotteries and 37% (18%) for the sure payment. More importantly, the WTA-WTP gap for the lotteries and that for the sure payment are significantly correlated. Second, subjects have a substantial range of prices (about 20% of the certainty equivalent of the lottery) and the range is positively linked to the WTA-WTP of the lotteries. Third, in a set of regression analyses, we show that controlling for the WTA-WTP gap of the sure payment reduces the WTA-WTP gap for lotteries by about 40%, and controlling for the ranges of prices reduces it by 15%. Controlling for both forms of complexity reduces it by about 50%. Finally, we find these three patterns are robust with respect to various specifications and in various demographic subgroups, such as age, gender, education, and income. Overall, these observations cannot be accounted for by loss aversion and suggest that the observed WTA-WTP gap for lotteries is partly due to complexity.

We do not preclude the loss aversion explanation. When we merge our data with previous data from Goossens and Knoef (2022), which includes a measure of loss aversion, we find that loss aversion does not correlate with the WTA-WTP gap in the general sample, consistent with findings based on representative samples in Chapman et al. (2021) and Fehr and Kübler (2022). However, the correlation is significant among subjects who do not experience either form of complexity. Furthermore, the role of loss aversion differs substantially across subsamples with different cognitive abilities. Among subjects who performed above the median on the matrix reasoning questions, the WTA-WTP gap is reduced by 53% when accounting for loss aversion and 35% when accounting for the two forms of complexity. In contrast, among subjects who performed below the median in the matrix reasoning questions, the WTA-WTP gap is actually increased by 11% when accounting for loss aversion and reduced by 83% when accounting for the two forms of complexity. While loss aversion is considered to be the most prominent preference-based explanation for the WTA-WTP gap, these results indicate that subjects may exhibit similar WTA-WTP gaps but differ substantially in the underlying reasons.

Our paper builds on two streams of the literature. First, it relates to understanding the elicitation of WTA and WTP from the perspective of mechanism complexity. Loomes et al. (2003) elicited the WTA and WTP of vouchers with a fixed redemption monetary value and showed that, among subjects who made mistakes, the majority overstated the WTA and under-reported the WTP, and such bias persisted after repetition. Plott and Zeiler (2005) systematically demonstrated that subjects' misconceptions of the experimental tasks, such as the BDM mechanism, may be an important source for the WTA-WTP gap and suggest procedures to mitigate such misconceptions (see also Plott and Zeiler, 2007, 2011 and Isoni et al., 2011 for discussions). Further, Cason and Plott (2014) elicited the WTA for an object with a known preference, showing that their subjects confused the second-price auction incentives of the BDM with the first-price auction and asked for a too-high WTA.

Second, this study yields insights into the role of valuation complexity in explaining various anomalies. Focusing on the WTA-WTP gap, Dubourg et al. (1994) showed that subjects' preferences are significantly imprecise, and this imprecision can explain part of the gap. Bayrak and Hey (2020) propose a model in which decision-makers have a set of preferences and showed that this model has the potential to explain the valuation gap.

McGranaghan and Otto (2022) demonstrate that valuation complexity plays a significant role in generating valuation asymmetries. More recently, Cerreia-Vioglio et al. (2022) showed theoretically that the WTA-WTP gap could arise from decision-makers who are unsure about the utility of the choice object and value it with caution.[2]

Our paper combines the two streams of literature mentioned above and makes several contributions. First, we jointly examine both forms of complexity and quantitatively assess their roles in the important anomaly of the WTA-WTP gap. This is possible because our elicitation of WTA and WTP is within-subject, and the WTA-WTP gap at the individual subject level provides a natural benchmark for the severity of deviation from classical theory (Brown et al., 2017, 2021). By examining the reduction of the WTA-WTP gap after controlling for the two forms of complexity separately and jointly, we can show the relative importance of each form in explaining the observed gap.

Second, we use a price list elicitation mechanism and assess its complexity. While the price list elicitation mechanism is widely used and commonly believed to be simpler than the BDM mechanism, its complexity is less well studied. We demonstrate that the mechanism complexity of the price list significantly influences behavior even within this seemingly straightforward elicitation mechanism.

Third, our study considers the WTA-WTP gap in a representative sample of the Dutch population. Responses from representative samples often play a vital role in policy design, so it is important to differentiate responses that reflect preferences from noise or biases. Snowberg and Yariv (2021) compared a wide range of choice behaviors for different subject pools, including a student population, a representative US sample, and subjects from Amazon Mechanical Turk, showing that, whereas correlations between behaviors were similar across samples, non-student samples exhibited higher levels of noise. Adding to two recent studies on the WTA-WTP gap in representative samples (Chapman et al., 2021; Fehr and Kübler, 2022), our study focuses on the use of heuristics driven by complexity and suggests that participants with low cognitive ability in the general population may be influenced

[2]Valuation complexity and caution are also in line with the salience explanation of the WTA-WTP gap (Bordalo et al., 2012). When facing complex choice objects, decision-makers may not be able to pay attention to all relevant attributes. Consequently, they may restrict their attention disproportionately to salient features that could minimize their regret, for example, attractive features when selling and less attractive features when buying. Sheng et al. (2023) examined the attention mechanism in an eye-tracking study and differentiated valuation-related and response-related bias underlying the WTA-WTP gap.

more by complexity.

Finally, our findings have the potential to reconcile some important results regarding the WTA-WTP gap that may otherwise seem surprising or contradictory. For example, the exchange between Plott and Zeiler (2005) and Isoni et al. (2011) showed that, after controlling for subjects' understanding of the elicitation mechanism, the WTA-WTP gap disappeared for mugs but persisted for monetary lotteries. While Plott and Zeiler (2011) did not explicitly identify the underlying mechanism based on their data, they proposed that lotteries may possess a distinct form of complexity. Our study suggests that this complexity could be related to the valuation complexity involved in aggregating multiple payoff components of lotteries (Oprea, 2022). Another example is the mixed evidence on the role of expectation-based reference points in the endowment effect (Ericson and Fuster, 2011; Sprenger, 2015; Cerulli-Harms et al., 2019). Our finding that the role of loss aversion depends on cognitive ability and experiences with complexity suggests that the instability of expectation-based reference effects may be due to complexity, for example, in contingent reasoning, in which expectations require simultaneous consideration of several states and values (Martínez-Marquina et al., 2019; Esponda and Vespa, 2023). A further example is the mixed effects of market experiences on exchange asymmetry, which is closely related to the WTP-WTP gap. Several studies showed that exchange asymmetry can be mitigated or eliminated by extensive market trading experience (List, 2003; Engelmann and Hollard, 2010; List, 2011; Anagol et al., 2018), while Apicella et al. (2014) observed that exchange asymmetry is present for hunter-gatherers with exposure to markets but not for those living in isolated regions. The use of the buy-low and sell-high heuristic when facing complexity found in our study suggests that this may be because some experience with buying and selling is required to develop this heuristic, but extensive experiences may reduce complexity and thus overcome its misuse. In a recent study, Fehr et al. (2022) showed that exchange asymmetry is less pronounced among the poor than the rich, and they attributed this observation to increased attention to decisions that involve higher stakes. Increased attention may help reduce the complexity of the choice environment, thus reducing the exchange asymmetry of the poor.

To sum up, our results shed new light on various explanations for the WTA-WTP gap. Mechanism complexity can help explain the link between the WTA-WTP gap for the lotteries and the sure payment, and valuation complexity provides a natural explanation for the link between the WTA-WTP gap for lotteries and the ranges of possible prices. Our results support the importance of both forms of complexity and show that the widely observed WTA-WTP gap can be in part due to the use of heuristics in the face of complexity. While loss aversion cannot account for these links, it is probably an important underlying factor related to the WTA-WTP gap since loss aversion measured in a different study is still able to account for a substantial portion of the WTA-WTP gap among those who exhibited high performance on matrix reasoning questions.[3] In this regard, our study supports the view that the widely observed WTA-WTP gap "results from many influences," as noted in a comprehensive review (Ericson and Fuster, 2014, p. 571), and these influences differ substantially across subjects.

The paper proceeds as follows. Section 2 explains the experimental design. Section 3 reports the experimental results, and Section 4 concludes.

2 Experimental Design

Each subject was presented with two monetary lotteries and reported their WTA and WTP for each lottery via price lists. We also measured the mechanism complexity of the price list and the valuation complexity of each lottery. All monetary outcomes were denoted by points in the experiment, with an exchange rate of 100 points = 0.50 euro. We explain each of these tasks in detail below.

2.1 Eliciting within-subject WTA and WTP

We elicited, within-subject, the WTA and WTP for two monetary lotteries. The first lottery (L1) offered 900 points or 100 points with equal probability. The second lottery (L2) offered 975 points with 20% chance, 850 points with 30% chance, 150 points with 30% chance, and 25 points with 20% chance.

Following Chapman et al. (2021), we elicited WTA and WTP with a price list with 20 rows (Holt and Laury, 2002). In each row, subjects chose between two options: a lottery and a price. The lottery was fixed in the entire list, while the price changed monotonically

[3] Cerreia-Vioglio et al. (2022) showed that loss aversion itself may arise from complexity and caution.

across rows. The prices were 0, 100, 200, 250, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 600, 700, 800, 900, and 1000 points. Note that the lottery dominated the price of 0 points and was dominated by the price of 1000 points. Similar to Chapman et al. (2021), we pre-selected the dominating option in these two rows. In the price list, we asked subjects to select the row in which they would switch from one option to the other. If a price list was randomly chosen for payment, subjects' payoff was determined by their choice in the randomly chosen row of the price list. To control for the effects of presentation on the elicitation of WTA and WTP, we randomized the option that appeared on the right or the left of the price list as well as the (increasing or decreasing) sequence of prices at the individual level.

To elicit WTA, we informed subjects that they were endowed with the lottery. In each row of the price list they chose between selling the lottery for a price or keeping the lottery. To elicit WTP, we endowed subjects with 1000 points. In each row of the price list they chose between buying the lottery for a price or keeping the 1000 points. We define WTA and WTP as the midpoint of the prices in the two switching rows, and the WTA-WTP gap is the difference between WTA and WTP. We randomized the eliciting order of WTA and WTP and separated them by other tasks. Using the language of buying and selling to elicit WTA and WTP is consistent with the literature (see, e.g., Isoni et al., 2011; Chapman et al., 2021). To check whether the framing of the WTA and WTP questions worked as expected, we also elicited the certainty equivalents of the two monetary lotteries in a neutral frame from a separate sample.

2.2 Measuring mechanism complexity

To measure the mechanism complexity of the price list, we elicited subjects' WTA and WTP for a sure payment – receiving 500 points with 100% chance – with the same type of price list as those for the two lotteries. The prices in the list for the sure payment were 450, 475, 499, 501, 525, and 550 points. We replaced the exact value of 500 points with 499 points and 501 points to avoid indifference between the price and the sure payment of 500 points. Similar to the $2 card used by Cason and Plott (2014), our sure payment has no payoff uncertainty and should thus induce a known preference. Should subjects understand the price list correctly, they would switch between 499 and 501 on both WTA

Question type	The choice in each row	
	Option 1	Option 2
WTA	Sell the lottery for y points.	Keep the lottery.
WTP	Buy the lottery for y points and keep the remaining [1000-y] points.	Keep the 1000 points.
	Lotteries	
L1	50%, 900 points; 50%, 100 points.	
L2	20%, 975 points; 30%, 850 points; 30%, 150 points; 20% 25 points.	
	Prices y	
Both lotteries	0, 100, 200, 250, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 600, 700, 800, 900, 1000.	

Table 1: Summary of the elicitation of WTA and WTP and the lotteries in the price list. The options appearing on the left or the right of the price list as well as the (increasing or decreasing) sequence of prices in the price list were randomized on the individual subject level.

and WTP questions. This would result in $\text{WTA}_{sure} = \text{WTP}_{sure} = 500$, and a WTA-WTP gap of zero. We interpret deviations of the WTA and WTP from 500 points as reflecting mechanism complexity.

2.3 Measuring valuation complexity

To measure valuation complexity, we extended the binary choices in the standard price list with a third option: the randomization option of receiving the lottery or the price according to a computerized coin flip. For each lottery, subjects faced an extended price list with 20 rows, referred to as the R-range price list. In each row of the R-range price list, subjects could choose the left option or the right option, which were either the lottery or the price as in the standard price list, or the new middle option of randomization. Similar to the elicitation of WTA and WTP, subjects needed to indicate the row of the R-range price list in which they would switch from the left option to the middle option of randomization, and the row in which they would switch from the middle option to the right option. Subjects could make these two switches in the same row and not select the randomization option at all. Figure 1 gives an example of the R-range price list. We measure valuation complexity by the range of prices in which subjects chose the randomization option (the R-

Price	Choices		
0	Lottery	Randomization	Price
100	Lottery	Randomization	Price
200	Lottery	Randomization	Price
250	Lottery	Randomization	Price
300	Lottery	Randomization	Price
325	Lottery	Randomization	Price
350	Lottery	Randomization	Price
375	Lottery	Randomization	Price
400	Lottery	Randomization	Price
425	**Lottery**	Randomization	Price
450	Lottery	Randomization	Price
475	Lottery	Randomization	Price
500	Lottery	Randomization	Price
525	Lottery	Randomization	**Price**
550	Lottery	Randomization	Price
600	Lottery	Randomization	Price
700	Lottery	Randomization	Price
800	Lottery	Randomization	Price
900	Lottery	Randomization	Price
1000	Lottery	Randomization	Price

Figure 1: An example of the R-range price list. The lower bound of the R-range is 437.5 points, the upper bound is 512.5 points, and the size of the R-range is $512.5 - 437.5 = 75$ points.

range). As shown in Figure 1, the lower bound of the R-range is the midpoint of the prices between which subjects switch from the lottery to the randomization option, and the upper bound of the R-range is the midpoint of the prices between which subjects switch from the randomization option to the price. The size of the R-range is the difference between the two bounds. Whereas choosing the randomization option more than once could be due to multiple reasons (Chew et al., 2022), our interpretation of the R-range as revealing valuation complexity is consistent with the literature (Cettolin and Riedl, 2019; Agranov and Ortoleva, 2023; Halevy et al., 2023), in which randomization is interpreted as evidence of incomplete/imprecise preferences.

The R-range price list is more complicated than the standard price list, and subjects may also perceive mechanism complexity. To control for the mechanism complexity of the R-range price list, we also elicited the R-range for the sure payment of 500 points.

As stipulated by stochastic dominance, subjects who correctly understood the elicitation mechanism of the R-range price list never chose the randomization option in this price list and would switch between 501 points and 499 points. Deviations from the above choices imply mechanism complexity. We used the R-range of the sure payment to capture mechanism complexity of the R-range elicitation mechanism.

2.4 Other measures

One potential source of mechanism and valuation complexity is limited cognitive ability. Given the same elicitation mechanism and choice objects, subjects with low cognitive ability are likely to be more prone to these two forms of complexity. To examine this hypothesis, we included six matrix reasoning questions from the International Cognitive Ability Resource (ICAR, Condon and Revelle, 2014) to assess subjects' cognitive ability. In these questions, subjects had to choose the image that best completed a 3 by 3 matrix (see Appendix G for an example). We incentivized the elicitation of cognitive ability by making subjects' payment dependent on the number of correct answers they provided.

To better understand the forces underlying the R-range, we included a subset of questions from the desirability of control scale, which elicits how subjects perceive the importance or benefits of maintaining control (Burger and Cooper, 1979; Gebhardt and Brosschot, 2002). As Fudenberg et al. (2015) pointed out, a preference for randomization may arise when the decision-maker makes a trade off between the probability of errors and the cost of implementing the desired choice. In this sense, a decision-maker who prefers to maintain control may perceive the implementation cost to be small and be less willing to choose the randomization option. Appendix A.2 explains in more detail how desirability of control may be related to valuation complexity. The scale consisted of statements like "I enjoy making my own decisions" and "Others usually know what is best for me." Subjects reported on a 7-point Likert scale how strongly they agreed or disagreed with these statements, ranging from completely disagree to completely agree. Appendix G provides the complete list of statements.

The questions above also served to separate WTA questions, WTP questions, and R-range questions.

2.5 Sample and procedure

We conducted the experiment using the LISS (Longitudinal Internet studies for the Social Sciences) panel administered by CentERdata (Tilburg University, The Netherlands) in September 2021. The LISS panel is a representative sample of the Dutch population who participate in online surveys monthly.[4]

We collected data from a total of 1856 subjects. Demographic information about the subjects is reported in Table C.1 in Appendix C. For the main sample of 1236 subjects, we used a within-subject design in which subjects reported their WTA, WTP, and R-range for each lottery. To address the concern that subjects may be influenced by their earlier responses, we separated the elicitation of WTA, WTP, and R-range by inserting other questions. We also randomized the order in which subjects responded to the WTA, WTP, and R-range questions. The order of the decisions was either: 1) WTA → matrix reasoning → WTP → desirability of control → R-range; 2) WTP → matrix reasoning → WTA → desirability of control → R-range; 3) R-range → desirability of control → WTA → matrix reasoning → WTP; or 4) R-range → desirability of control → WTP → matrix reasoning → WTA. We further randomized the order of L1 and L2 in each task. For a separate sample of 620 subjects, we replaced the WTA and WTP questions with price lists in a neutral frame to elicit the certainty equivalents for the same two monetary lotteries.

After the online survey was completed, we randomly selected 10% of the subjects and paid them an additional amount based on one of their decisions in the experiment (WTA questions, WTP questions, R-range questions, or their performance on the six matrix reasoning questions). If the decision selected for payment was the randomization option, a random draw determined whether subjects would receive the lottery or the price. If they received the lottery, a further random draw determined the payment of the lottery. On average, the subjects who were selected for the additional payment received 2.97 euros on top of the flat fee of 3 euro for completing the survey. The experiment was approved by the Institutional Review Board of Radboud University.

[4]Households that could not otherwise participate were provided with a computer and Internet connection. The longitudinal survey is fielded in the panel every year, covering a large variety of domains including health, work, education, income, housing, time use, political views, values, and personality.

3 Experimental Results

In reporting the experimental results, we use Wilcoxon signed-rank tests for within-subject statistical differences, Wilcoxon rank-sum tests for between-subject statistical differences, and Obviously-Related Instrumental Variables (ORIV; Gillen et al., 2019) to estimate correlations, unless stated otherwise.[5]

We find that on average subjects exhibited the standard WTA-WTP gap by reporting a higher WTA than WTP. Figure 2 reports violin plots of the average WTA, WTP, and WTA-WTP gap across the two lotteries. Across the two lotteries, the mean WTA is 523.47 and the mean WTP is 411.51, resulting in an average WTA-WTP gap of 111.96. The WTA-WTP gap is about 22% of the expected value of the lotteries, which is comparable to the finding reported by Chapman et al. (2021). In contrast, the mean certainty equivalent from the neutral frame is 470.84, which differs significantly from the WTA and WTP of the lotteries ($p < 0.01$). This suggests that the framing of the WTA and WTP questions in the experiment worked in the expected direction. On the individual level, 58.41% of subjects showed a positive WTA-WTP gap, 15.05% had a WTA-WTP gap of 0, and a non-negligible proportion (26.54%) of subjects exhibited a negative WTA-WTP gap. These values are broadly consistent with the proportions reported in previous studies. For example, in Chapman et al. (2021) these values are 60%, 10%, and 30% respectively.

In the following analyses, we present our results in three steps. First, we report measures of mechanism complexity and valuation complexity and examine their association with the WTA and WTP of the lotteries. Second, we correct the WTA-WTP gap for mechanism complexity and/or valuation complexity. Finally, we demonstrate the robustness of our results by considering different demographic groups, the inclusion of loss aversion, and a comparison across subsamples.

3.1 Mechanism complexity

We find that mechanism complexity, as measured by deviations of WTA_{sure} and WTP_{sure} from 500, is systematic. Figure 3 presents the distribution of WTA_{sure}, WTP_{sure}, and the WTA-WTP gap_{sure}. We find that 64% of WTA_{sure} deviations were positive and 36% were

[5] We use the WTA and WTP of two lotteries for ORIV estimation of correlations. The choice behavior does not significantly differ between the two lotteries, as shown in the Appendix C.

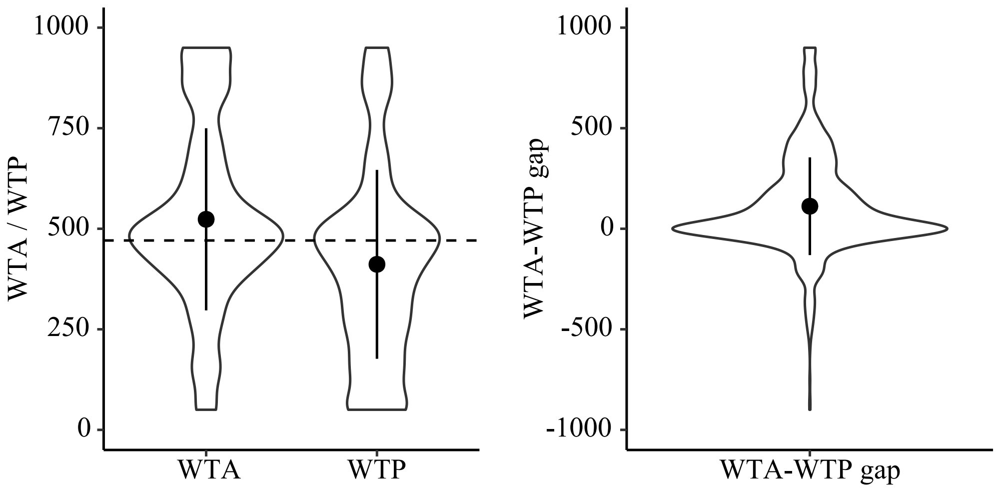


Figure 2: Violin plots of the average WTA, WTP, and WTA-WTP gap of the two lotteries. The dots in the violins denote the means, the lines represent the standard deviations. The dashed line is the average certainty equivalent of the lotteries from the neutral frame (470.84 points). See Table C.2 in Appendix C for details on each lottery separately.

negative. In contrast, 58% of WTP_{sure} deviations were negative and 42% were positive. This leads to more positive WTA-WTP gaps for the sure payment than negative ones. Three logistic regressions (reported in Table C.3 in Appendix C) suggest that the likelihood of reporting $\text{WTA}_{sure} = \text{WTP}_{sure} = 0$ is significantly and positively related to subjects' performance on the matrix reasoning questions, education, self-reported understanding of the experimental tasks, and time spent on the tasks. These factors are commonly associated with subjects' ability or willingness to understand experimental tasks. Thus, deviations of the sure payment's WTA and WTP from 500 are likely to arise from subjects' difficulty with the complexity of the price list elicitation mechanism, which supports the use of these deviations as a proxy for mechanism complexity.

Systematic bias of WTA_{sure} and WTP_{sure} can also be seen when we relate them to the WTA and WTP of the two lotteries. There is a similar asymmetry in the proportions of positive and negative WTA-WTP gaps between the sure payment (36.97% vs. 17.64%) and the lotteries (58.41% vs. 26.54%). Figure 4 depicts violin plots of the average WTA and WTP of the lotteries at each level of WTA_{sure} and WTP_{sure}, respectively. The data pattern suggests a systematic association of WTA and WTP between the sure payment and the lotteries and further shows that the WTA of the lotteries associates more strongly with positive WTA_{sure} deviations than negative ones, while the WTP of the lotteries

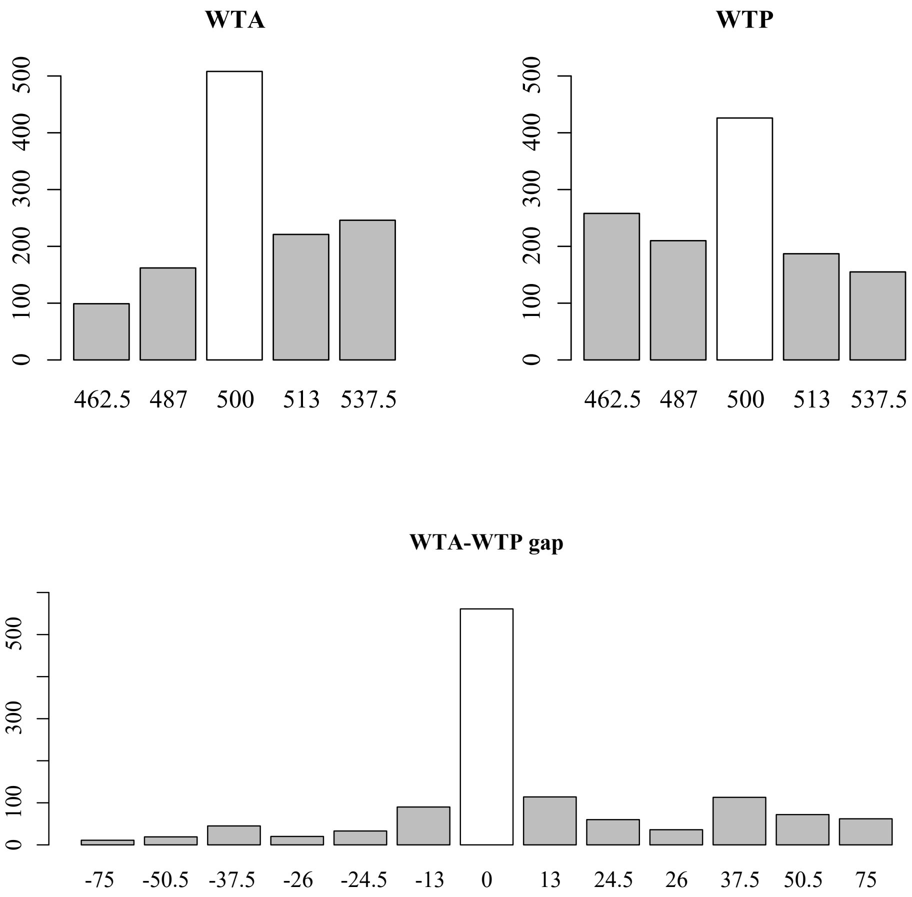


Figure 3: The frequencies of subjects at each level of the WTA, WTP, and WTA-WTP gap for the sure payment. The white bars indicate the optimal WTA and WTP.

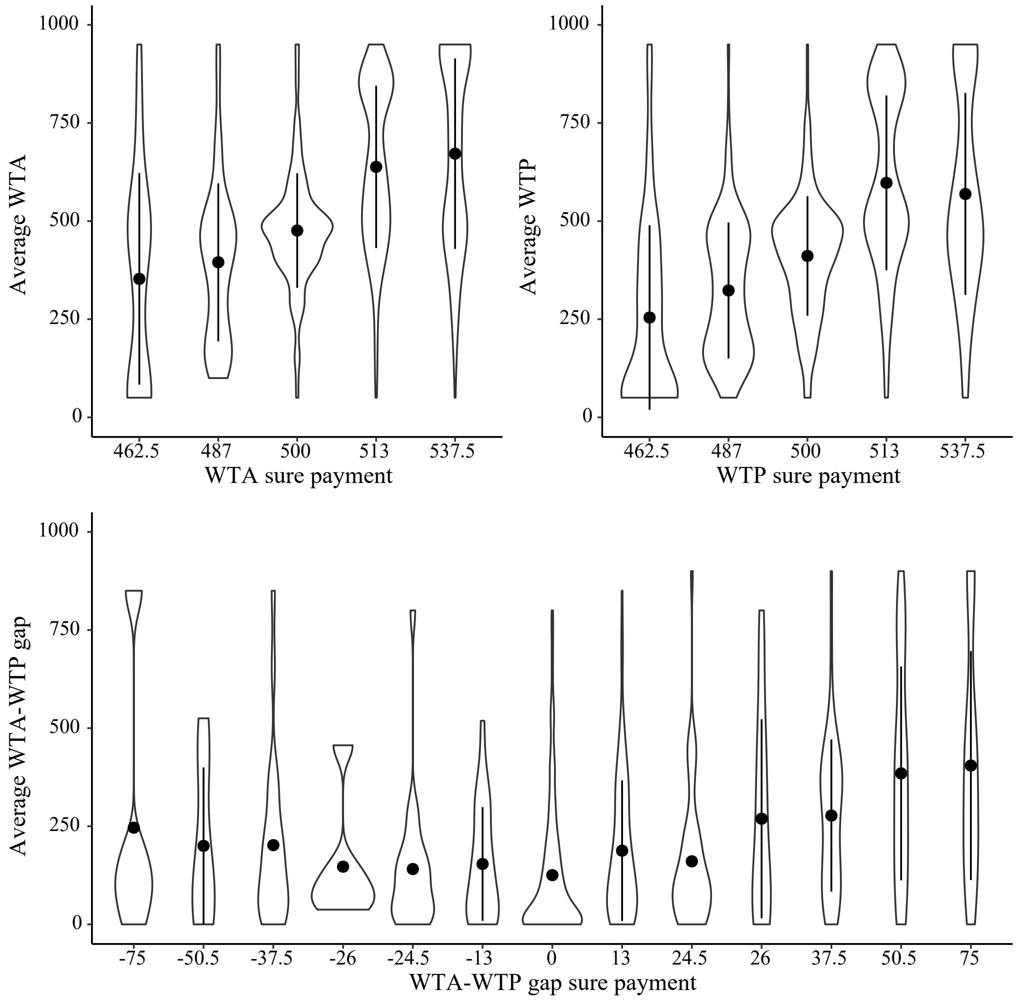


Figure 4: Violin plots of the average WTA, WTP, and WTA-WTP gap of the two lotteries at each level of the sure payment's WTA, WTP, and WTA-WTP gap. The dots in the violins denote the means, the lines represent the standard deviations.

associates more strongly with negative WTP_{sure} deviations than positive ones. Consistent with Figure 4, we find significantly positive correlations between WTA_{sure} and the average WTA of the lotteries ($\rho = 0.599, p < 0.01$), between WTP_{sure} and the average WTP of lotteries ($\rho = 0.612$, $p < 0.01$), and between the WTA-WTP gap of the sure payment and the average WTA-WTP gap of the lotteries ($\rho = 0.578$, $p < 0.01$). Examining positive and negative deviations separately, we find that positive WTA deviations and negative WTP deviations, which are consistent with a positive WTA-WTP gap, correlate significantly with the WTA and WTP of the lotteries ($\rho = 0.107$, $p < 0.05$ for WTA, and $\rho = 0.228$, $p < 0.01$ for WTP), while correlations between negative WTA deviations or positive WTP deviations and those of the lotteries are weaker ($\rho = 0.123$, $p < 0.10$ for WTA and $\rho = -0.084$, $p > 0.10$ for WTP). These results suggest that positive WTA_{sure} deviations and negative WTP_{sure} deviations are systematic biases, while negative WTA_{sure} deviations and positive WTP_{sure} deviations are likely to arise from noise. We summarize our observation below.

Observation 1. *The WTA and WTP of the sure payment deviate systematically from the known preference, and these deviations correlate strongly with the WTA and WTP of the lotteries.*

3.2 Valuation complexity

As we explained in the experimental design, we capture valuation complexity by the range of prices for which subjects chose the randomization option (the R-range). We find that the majority of subjects chose the randomization option at least once for either of the lotteries (67%), and many did so twice or more (56%). This results in a substantial R-range (mean R-range of 103.96 across the two lotteries). Table C.4 in Appendix C provides more details on the upper bound, lower bound, and R-range for each lottery.

The R-range also exhibits some systematic patterns. First, we hypothesized that subjects who have a stronger desire for control are more willing to pick their desired option than the randomization option. Consistent with the hypothesis, we find a significant negative correlation between subjects' desirability of control and their average R-range across the two lotteries ($\rho = -0.103$, $p < 0.05$). We also find a statistically significantly higher level of desirability of control among subjects who do not have an R-range for either lottery than among those who chose the randomization option at least once in one of the two lotteries (5.28 vs. 5.16, $p < 0.01$).

Second, we find a positive and significant correlation between the R-range and the WTA-WTP gap ($\rho = 0.144$, $p < 0.01$). When we restrict to subjects who randomized at least once, we find an even stronger correlation ($\rho = 0.264$, $p < 0.01$). Consistent with the correlation tests, we find that subjects who chose the randomization option more than twice in the two lotteries have a significantly higher WTA-WTP gap than those who chose the randomization option only once or twice (122.07 vs 79.61, $p < 0.01$). This trend can also be seen in Figure 5. The picture for the group of subjects who never chose the randomization option is less clear. They have no R-range, but their WTA-WTP gap is significantly larger

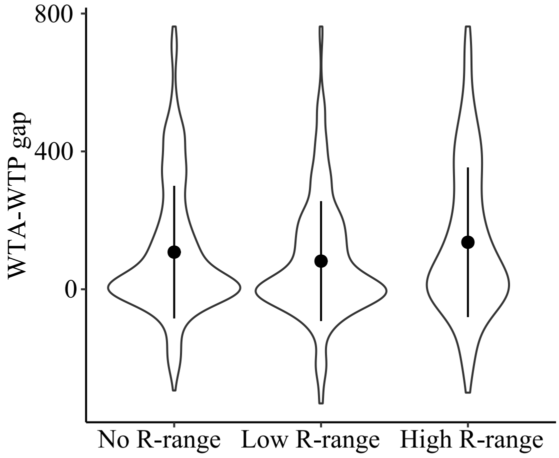


Figure 5: Violin plots of the average WTA-WTP gap at different sizes of the R-range in L1 and L2. Low and high R-range are computed by taking the median split of the R-range (100 points) for subjects with an R-range larger than zero. The dots in the violins denote the means, the vertical lines represent the standard deviations. For better readability, we excluded the top and bottom 2.5% of the WTA-WTP gap (-332.03 and 763.28 points).

compared to those with low valuation complexity (122.04 vs. 78.85, $p < 0.05$) but similar to those with high valuation uncertainty (122.04 vs. 137.88, $p > 0.10$. See Table C.4 in Appendix C for more details). One possible explanation for this is that some subjects did not understand the mechanism used to elicit the R-range. We will examine this point more closely in subsection 3.3 and Appendix B. A further possibility is that the WTA-WTP gap of the group with no R-range relates to factors other than complexity (e.g., loss aversion, to which we return in Subsection 3.4). We offer the following observation regarding valuation complexity.

Observation 2. *The randomization option was chosen frequently. Among subjects who selected the randomization option at least once, the R-range as a measure of valuation complexity is significantly and positively correlated with the WTA-WTP gap.*

To summarize, Table 2 provides the distribution of subjects with/out each form of complexity and their associated average WTA-WTP gap, education, and performance on the matrix reasoning questions. About 53% of subjects exhibited both forms of complexity, and only 9% did not exhibit either form of complexity. Further, subjects with mechanism complexity exhibited significantly higher WTA-WTP gaps than those without, and those with high valuation complexity had higher WTA-WTP gaps than those with low valuation

	Mechanism complexity			No mechanism complexity		
	Valuation complexity			Valuation complexity		
	High	Low	No	High	Low	No
Proportions	23.9%	28.6%	24.3%	6.9 %	7.4%	8.8%
WTA-WTP gap	158.99 (280.38)	91.74 (221.95)	133.23 (282.43)	71.47 (176.85)	31.05 (131.84)	91.23 (153.88)
Education	3.86 (1.48)	3.73 (1.47)	3.73 (1.50)	4.34 (1.29)	4.22 (1.49)	4.81 (1.27)
MR performance	3.28 (1.72)	3.02 (1.80)	2.87 (1.83)	4.20 (1.54)	4.22 (1.79)	4.72 (1.46)

Table 2: The distribution of the subjects with/out each form of complexity, and their associated average WTA-WTP gap, education, and performance in the matrix reasoning questions. The high and low valuation complexity groups are defined by a median split of the R-range (100 points for subjects with mechanism complexity, and 118.75 points for those without). Education level is measured on a scale from 1 to 6. Standard deviations are in parentheses.

complexity ($p < 0.01$ in both tests). Finally, subjects' experience of complexity is related to cognitive ability as measured by education and matrix reasoning performance: Compared with subjects with both forms of complexity, subjects who did not exhibit either form of complexity had more education and performed better on the matrix reasoning questions ($p < 0.01$ in both tests).

3.3 Controlling for mechanism and valuation complexity

The analysis presented above suggests that the majority of subjects exhibited mechanism complexity, which represents a systematic bias. Furthermore, this bias carries over to the WTA and WTP of the lotteries. There is also substantial valuation complexity among subjects, and this relates systematically to the WTA-WTP gap. Since these two forms of complexity do not represent subjects' preferences in the classical sense, we now examine how much the observed WTA-WTP gap can be accounted for by mechanism complexity and/or valuation complexity.

Correcting the WTA-WTP gap for mechanism complexity: We correct the WTA-WTP gap for mechanism complexity in three ways: based on a theoretical analysis (Cason and Plott, 2014), statistically for deviations ($500 - \text{WTP}_{sure}$ and $\text{WTA}_{sure} - 500$), and statistically for deviations while allowing different slopes for positive and negative deviations.

We first correct the WTA-WTP gap for mechanism complexity based on a theoretical analysis that extends from Cason and Plott (2014). In this analysis, we assume that subjects mistake the switching value in the price list for the final trading price, as in the first-price auction (see Appendix A.1 for more details). Given the reported WTP_X and WTA_X for the lottery X, X =L1 and L2, the theoretical analysis suggests that WTP^*_X and WTA^*_X, after correcting for mechanism complexity, are computed as follows:

$$\begin{aligned} WTP^*_X &= \frac{100}{WTP_{sure} - 400} WTP_X \\ WTA^*_X &= \frac{100}{600 - WTA_{sure}} WTA_X - \frac{WTA_{sure} - 500}{600 - WTA_{sure}} 1000. \end{aligned}$$

Note that there is no correction ($WTP^*_X = WTP_X$ and $WTA^*_X = WTA_X$) when $WTP_{sure} = WTA_{sure} = 500$. After the correction, we observe a significantly lower WTA (497.86 vs. 523.47, $p < 0.01$), a higher, although not significantly, WTP (421.70 vs. 411.51, $p > 0.10$), and both are moving closer to the CE of 470.84 from the neutral frame. As a result, the WTA-WTP gap is significantly smaller, and it is reduced by one-third after the correction (76.16 vs. 111.96, $p < 0.01$). The results for the two lotteries separately, presented in Table C.5 in Appendix C, are virtually the same.

A limitation of the theoretical correction is that it assumes that mechanism complexity affects WTA and WTP through the specific misperception of the first-price auction, and that this misperception works to the same extent for L1 and L2. However, neither assumption may hold in the experiment. To relax these assumptions, we proceed to correct WTA and WTP statistically. In each statistical correction model, the dependent variable is the WTA or WTP of the lottery. The independent variables are the mechanism complexity-related variable(s) described above, denoted as **MC**. Explicitly, we first run the following two OLS regressions for each lottery:

$$WTA_X = I_{a,X} + \boldsymbol{\beta}_{a,X}\mathbf{MC_a} + \epsilon; \quad WTP_X = I_{p,X} + \boldsymbol{\beta}_{p,X}\mathbf{MC_p} + \epsilon, \tag{1}$$

where I is the intercept, X =L1 or L2, or the average of the two lotteries, the vector of coefficient(s) is denoted by $\boldsymbol{\beta}$, and the subscript a or p denote WTA or WTP. These regressions capture the idea that how subjects reported their WTA and WTP may have been systematically affected by mechanism complexity, and the term $\boldsymbol{\beta}\mathbf{MC}$ captures the component of the WTA or WTP that does not derive from classical preferences. Similar regression forms have been extensively used in financial research to account for the influence of various risk factors on excess asset returns (see e.g., Fama and French, 1992). With the estimated $\hat{\boldsymbol{\beta}}$, we can correct the WTA and WTP for mechanism complexity as follows:

$$WTA^*_X = WTA_X - \hat{\boldsymbol{\beta}}_{a,X}\mathbf{MC_a}; \quad WTP^*_X = WTP_X - \hat{\boldsymbol{\beta}}_{p,X}\mathbf{MC_p}.$$

The variable(s) $\mathbf{MC}$ differs across the correction method. In the second correction, $\mathbf{MC_a} = (\text{WTA}_{sure} - 500)$ and $\mathbf{MC_p} = (500 - \text{WTP}_{sure})$. Similar to the results of the theoretical correction, after the statistical correction we observe a significantly lower WTA (497.95 vs. 523.47, $p < 0.01$), a significantly higher WTP (428.13 vs. 411.51, $p < 0.05$), and both moving closer to the neutral frame. Consequently, the WTA-WTP gap is significantly smaller, reduced by about 38% after the correction (69.82 vs. 111.96, $p < 0.01$).

In the third correction, we further allow for different slopes for positive and negative deviations. As Figure 4 suggests, the WTA of the lotteries is more strongly associated with positive WTA deviations of the sure payment than with negative ones, and the WTP of the lotteries is more strongly associated with negative WTP deviations of the sure payment than with positive ones. To account for this, we also include an interaction term with deviations multiplied by a dummy variable that indicates negative deviations ($\mathbf{D_{neg}}$). According to statistical model 1, $\mathbf{MC_a} = (WTA_{sure} - 500, \mathbf{D_{neg}}(WTA_{sure} - 500))$ and $\mathbf{MC_p} = (500 - WTP_{sure}, \ \mathbf{D_{neg}}(500 - WTP_{sure}))$. Consistent with the previous two corrections, in the third correction we observe a significantly lower WTA (489.66 vs. 523.47, $p < 0.01$), a significantly higher WTP (428.31 vs 411.51, $p < 0.01$), and a significantly smaller WTA-WTP gap, which is about 55% of the original level (61.35 vs. 111.96, $p < 0.01$).

Figure 6 provides an overview of the three corrections. As we can see, after using each correction method the WTA-WTP gap is reduced significantly, from about one-third to 45%. More details can be found in Table C.5 in Appendix C. In Appendix D, we perform an additional correction, where we use dummy variables for each deviation of the sure payment's WTA or WTP from 500 to correct for the WTA, WTP, and WTA-WTP gap of the lotteries. The results are broadly similar, with a 40% reduction of the lotteries' WTA-WTP gap after the correction.

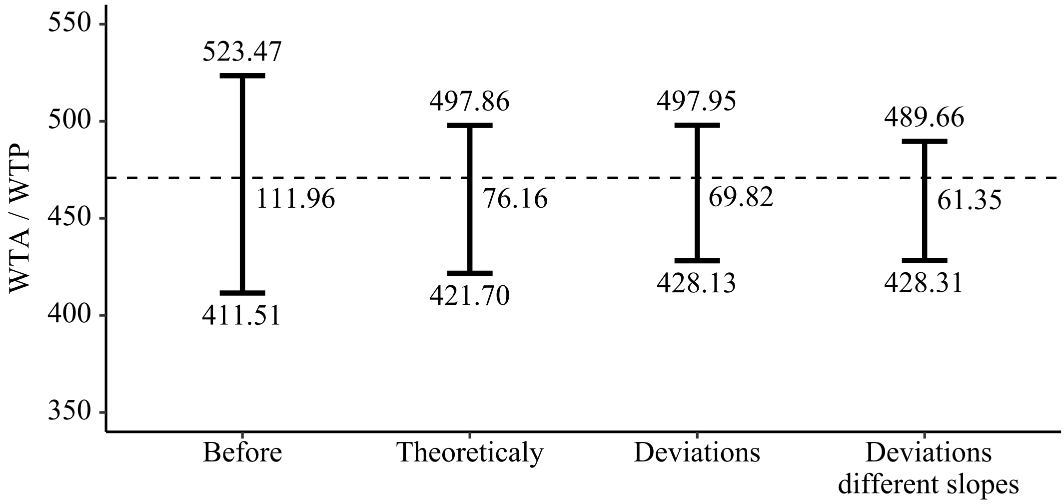


Figure 6: The WTA, WTP, and WTA-WTP gap for the lotteries before and after each of the three mechanism complexity correction methods. The dashed line is the average certainty equivalent of the lotteries from the neutral frame (470.84 points).

Controlling the WTA-WTP gap for valuation complexity: To control the WTA-WTP gap for valuation complexity, we follow the theoretical analysis in Appendix A.2 and control the WTA and WTP of the lottery separately for the corresponding R-range with the following regressions:

$$WTA_X = I_{a,X} + \beta_a \text{R-range}_X + \epsilon; \quad WTP_X = I_{p,X} + \beta_p \text{R-range}_X + \epsilon,$$

where I is the intercept, X is L1, L2, or the average of the two lotteries. With the estimated $\hat{\beta}_i, i = a, p$, the WTA* and WTP*, after controlling for valuation complexity, are calculated as follows:

$$WTA^*_X = WTA_X - \hat{\beta}_a \text{R-range}_X; \quad WTP^*_X = WTP_X - \hat{\beta}_p \text{R-range}_X.$$

By controlling for valuation complexity, we observe a change in the WTA, WTP, and WTA-WTP gap for the lotteries in the same direction as for the mechanism complexity corrections, although with a smaller magnitude. The WTA is significantly lower (514.92 vs. 523.47, $p < 0.01$), and the WTP is significantly higher (419.62 vs. 411.51, $p < 0.01$). Consequently, the WTA-WTP gap is reduced significantly by about 15% (95.30 vs. 111.96,

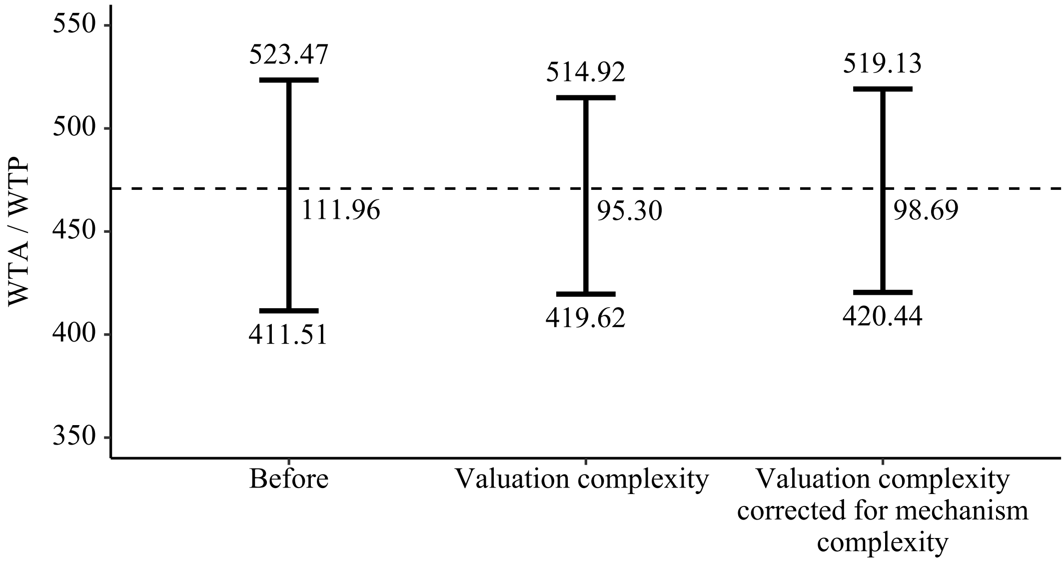


Figure 7: The WTA, WTP, and WTA-WTP gap for the lotteries before and after controlling for (mechanism complexity corrected) valuation complexity. The dashed line represents the average certainty equivalent of the lotteries in the neutral frame (470.84 points).

$p < 0.01$).

As discussed in section 2.3, there could be complexity with respect to the R-range price list, which could distort our measure of valuation complexity. Therefore, we perform an additional analysis by first correcting the upper bound and the lower bound of the R-range for mechanism complexity, and then using the mechanism complexity corrected R-range to control the WTA-WTP gap for valuation complexity. More details on these procedures can be found in Appendix B. Controlling for the mechanism complexity corrected R-range produces results comparable to those using the elicited R-range directly: The WTA decreases significantly (519.13 vs. 523.47, $p < 0.01$), and the WTP increases significantly (420.44 vs. 411.51, $p < 0.01$), with the average WTA-WTP gap reduced significantly by 12% (98.69 vs. 111.96, $p < 0.01$). Figure 7 provides an overview of the WTA, WTP, and WTA-WTP gap for the lotteries before and after controlling for (mechanism complexity corrected) valuation complexity. More details can be found in Tables B.1 and C.6 in Appendix C. Overall, our results suggest that valuation complexity plays an important role in the WTA-WTP gap.

Considering two forms of complexity simultaneously: We now examine the WTA, WTP, and resulting WTA-WTP gap for the lotteries accounting for mechanism complexity and valuation complexity simultaneously. We conduct four analyses. In Models 1 and 3, mechanism complexity is the deviation of the sure payment's WTA or WTP from 500, and in Models 2 and 4, mechanism complexity also allows for different slopes for positive and negative deviations. In Models 1 and 2 valuation complexity is the elicited R-range, and in Models 3 and 4 valuation complexity is the R-range corrected for mechanism complexity.

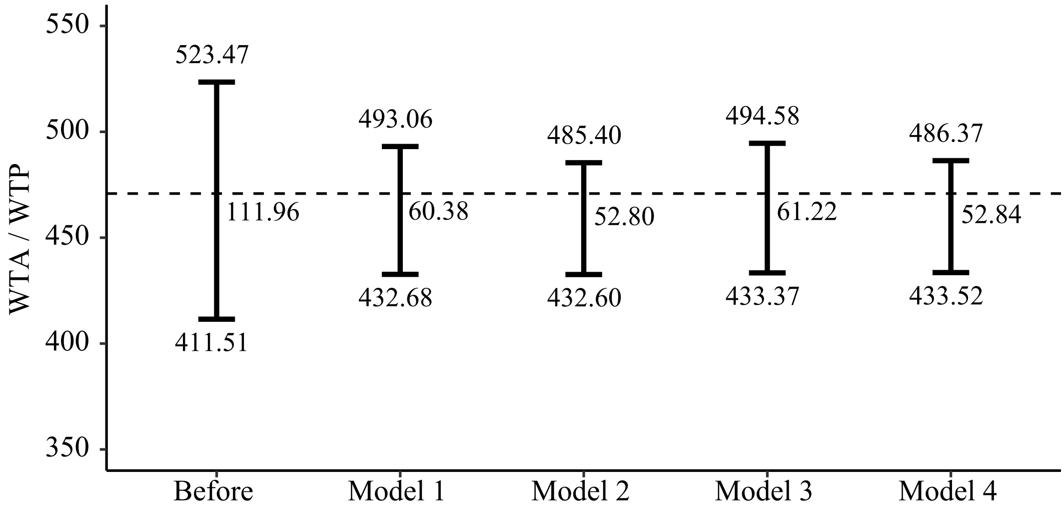


Figure 8: The WTA, WTP, and WTA-WTP gap for the lotteries before and after controlling for both mechanism complexity and valuation complexity. In Model 1 and 3 mechanism complexity is the deviation of the sure payment's WTA or WTP from 500, and in Model 2 and 4 mechanism complexity additionally allows for different slopes for positive and negative deviations. In Model 1 and 2 valuation complexity is the elicited R-range, and in Model 3 and 4 valuation complexity is the R-range corrected for mechanism complexity. The dashed line represents the average certainty equivalent of the lotteries in the neutral frame (470.84 points).

As we can see from Figure 8, the WTA-WTP gap decreases substantially in all four models after accounting for both forms of complexity, ranging from 45% to 53%, with an average residual WTA-WTP gap around 50% of the original level. More details can be found in Table C.7 in Appendix C. These results suggest that classical preference-based explanations of the WTA-WTP gap like loss aversion can not explain these results. For a more comprehensive understanding of the WTA-WTP gap, it is important to consider the

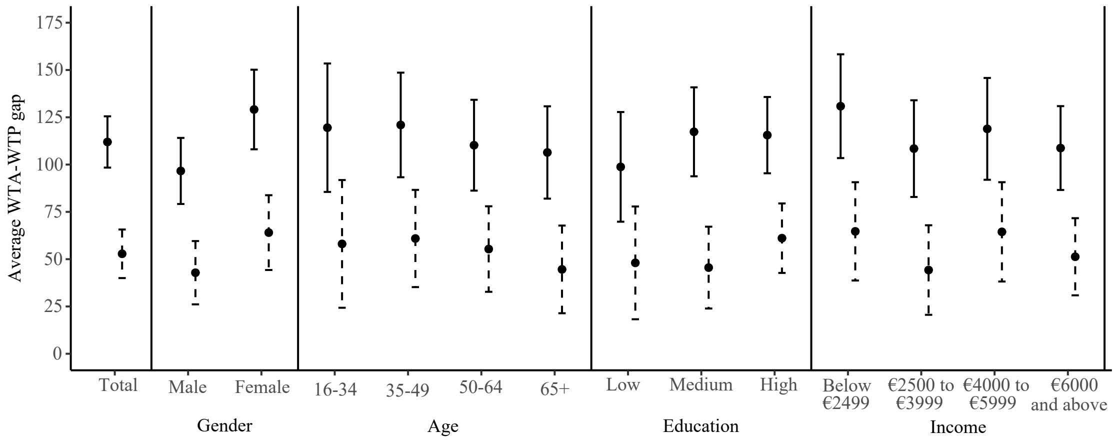


Figure 9: The WTA-WTP gap in various demographic groups. The solid lines are the elicited WTA-WTP gap. The dashed lines are after controlling for mechanism complexity and valuation complexity (Model 4 as defined in Figure 8). Error bars show the 95% confidence intervals.

two forms of complexity. We summarize our observation in the following statement.

Observation 3. *Accounting for mechanism complexity alone reduces the WTA-WTP gap by about 40%. Controlling for valuation complexity alone leads to a 15% reduction in the WTA-WTP gap. When considering both forms of complexity simultaneously, the WTA-WTP gap decreases by over 50%.*

3.4 Robustness checks

Demographic groups: One concern is that the significant impact of mechanism complexity and valuation complexity on the WTA-WTP gap applies primarily to certain demographic groups, e.g., those with low education, and thus the insights cannot be generalized to all groups. Our representative sample allows us to examine this concern. We consider demographic groups based on gender (male, female), age (16-34, 35-49, 50-64, and 65 or above), education (based on the categorization of Statistics Netherlands, with low for primary or pre-vocational education, medium for pre-university or vocational education, and high for higher vocational or university education), and household monthly income (< 2499, $2500 - 3999$, $2500 - 3999$, and 6000 or above). Table C.1 in Appendix C reports the percentages of subjects in each demographic group. Figure 9 displays means and 95% confidence intervals of the WTA-WTP gap in various demographic groups, both before and after controlling for mechanism complexity and valuation complexity. We make two main observations. First, the WTA-WTP gap exhibits some heterogeneity but exists in all demographic groups. Second, controlling the WTA-WTP gap for mechanism complexity and valuation complexity has a significant and substantial effect in all demographic groups.

Loss aversion: The main purpose of our study is to demonstrate the role of mechanism complexity and valuation complexity in the WTA-WTP gap. Since "[l]oss aversion has been the leading theory " for explaining the WTA-WTP gap (Ericson and Fuster, 2014, p. 571), as a further robustness check, we examine whether the above results hold when we also include loss aversion as a control.

We use the loss aversion measure from an independent study based on the LISS panel (Goossens and Knoef, 2022).[6] A unified notion of loss aversion would predict a significant and positive relationship between the WTA-WTP gap and loss aversion for risky bets involving gains and losses. Our sample has an overlap of 533 subjects with the sample of Goossens and Knoef (2022). The analyses below consider these subjects. The correlation between the WTA-WTP gap and the loss aversion coefficient is positive and thus consistent with the loss aversion explanation, but not significant ($\rho = 0.067$, $p > 0.10$).[7] This result is consistent with the findings of Chapman et al. (2021) and Fehr and Kübler (2022), who found no significant correlation between loss aversion and the endowment effect in representative samples from the United States and Germany, respectively. Remarkably, when categorizing subjects into groups with/out each form of complexity as in Table 2, we find a statistically significant correlation between loss aversion and the WTA-WTP gap in the group who did not experience mechanism or valuation complexity ($\rho = 0.507$, $p < 0.01$), while there are correlations of nearly zero and that are not significant in the other three groups. Table C.8 further shows that the group that did not experience either form of

[6] In this measure, subjects faced a price list with five rows. In each row subjects faced two lotteries A and B. The lottery A paid $-X$ and $+30$ with equal likelihood. The lottery B paid -2 and $+10$ with equal likelihood. Lottery B stayed the same across rows, while x in lottery A increased (-22,-16,-11,-8,-6). Subjects were asked to indicate the row they would switch from A to B. The loss aversion coefficient is calculated as $\lambda = \frac{30-10}{-2-X}$. For example, $\lambda = 1$ when $X = -22$.

[7] We also checked the correlation between the WTA/WTP ratio or its square root and loss aversion, and the result is similar ($\rho = 0.002$, $p > 0.10$). Further, we do not find statistically significant correlations between loss aversion and mechanism complexity or valuation complexity (ρ close to zero, $p > 0.10$ for all tests).

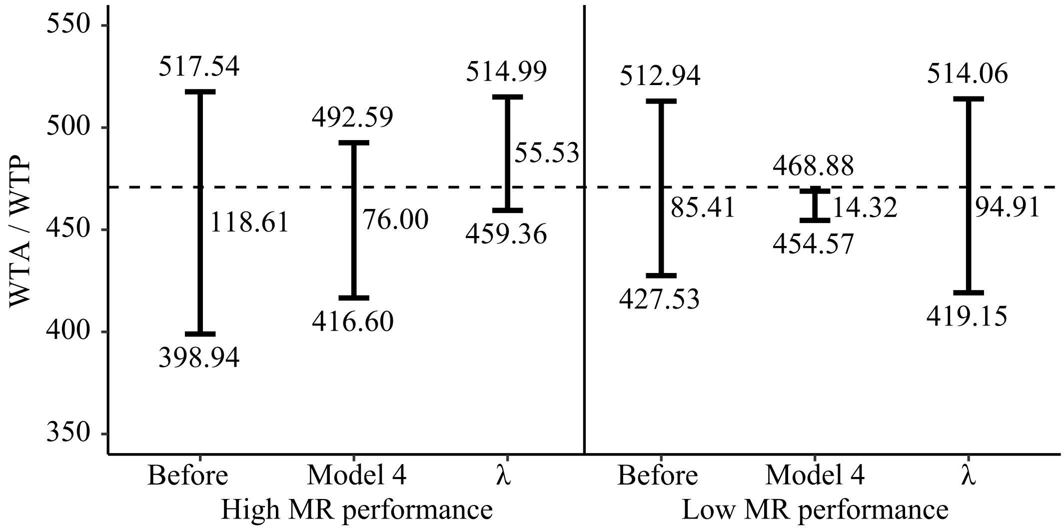


Figure 10: The WTA, WTP, and WTA-WTP gap for the lotteries in two subsamples based on subjects' performance on the matrix reasoning questions before and after controlling for mechanism complexity and the corrected valuation complexity as well as loss aversion (λ) in the overlapping subsample. The regression controlling for λ only considers loss aversion. Model 4, as defined in Figure 8, controls for both forms of complexity. The dashed line represents the average certainty equivalent of the lotteries in the neutral frame (470.84 points).

complexity had a substantial WTA-WTP gap (122.78), comparable to the gap in the overall sample. This corroborates our earlier speculation that the substantial WTA-WTP gap in the group with no R-range could be partly due to loss aversion.

Cognitive ability: The experience of complexity depends on cognitive ability. As a further analysis, we split the sample into two subsamples based on subjects' performance (above or below the median) on the matrix reasoning questions. We find that the correlation between loss aversion and the WTA-WTP gap is significant for the high matrix reasoning performance groups but not for the low performance group ($\rho = 0.182$, $p < 0.01$ vs. $\rho = -0.038$, $p > 0.10$).

We then performed the same analysis in Subsection 3.3 for each subsample separately. Figure 10 illustrates the results. Accounting for the two forms of complexity significantly reduces the WTA-WTP gap in both subsamples, with the impact of complexity being greater in the subsample with low matrix reasoning performance: about 35% for those who performed equal to or above the median vs. 83% for those below the median on the matrix reasoning questions. This suggests that complexity plays a more important role for those with lower cognitive ability.

As a comparison, when we use loss aversion instead of the two forms of complexity to account for the WTA-WTP gap, controlling for loss aversion alone reduces the gap by 53% (from 118.61 to 55.53) for subjects with high matrix reasoning performance, but it even increases the gap by 11% among subjects with low matrix reasoning performance. More details are presented in Table E.4, Table E.5, and other related tables in Appendix E. These results suggest that not only is the WTA-WTP gap influenced by a variety of factors but also that these factors differ substantially among different subjects.[8] We summarize the results of this subsection as follows:

Observation 4. *The observed roles of mechanism and valuation complexity hold for different subgroups based on age, gender, education, and income. Complexity plays a more important role in the subgroup with lower cognitive ability, and loss aversion plays a more important role in the subgroup with higher cognitive ability.*

4 Concluding Remarks

Using a large representative sample of the Dutch population, we show that mechanism and valuation complexity contributed substantially to the WTA-WTP gap of the monetary lotteries elicited using a price list mechanism. More specifically, we find that mechanism complexity, as measured by deviations from a known preference, correlates significantly with the WTA and WTP of monetary lotteries. Meanwhile, valuation complexity, as measured by the R-range, relates systematically to the WTA-WTP gap of the monetary lotteries. A series of quantitative analyses demonstrate that mechanism complexity alone has the potential to account for about 40% of the observed WTA-WTP gap for the lotteries,

[8] A perhaps surprising finding in Figure 10 (also tables E.4 and E.5) is that, after accounting for the two forms of complexity, the group with low matrix reasoning performance exhibited a lower WTA-WTP gap in comparison to the group with high matrix reasoning performance. However, this reduction should not be misinterpreted as evidence that these subjects would have naturally displayed a lower WTA-WTP gap in the absence of these two forms of complexity. If these subjects had a complete understanding of the elicitation mechanism and the valuation of the lotteries, they might have been able to make more informed decisions to reveal their loss aversion-driven WTA-WTP gap like those with high matrix reasoning performance. Chew et al. (2018) measured subjects' comprehension of the experiment and showed that high-comprehension subjects exhibited ambiguity aversion while low-comprehension subjects appeared to be ambiguity neutral by choosing randomly.

valuation complexity alone for about 15%, and the two forms of complexity together for about 50%.

Our observations can be accommodated by linking complexity and heuristics. Complexity naturally arises in studies that involve experiments or surveys. When decision-makers face complexity, they often rely on heuristics to make decisions (Simon, 1955; Kahneman and Tversky, 1974; Gigerenzer et al., 1999; Ortoleva, 2010; Iyengar and Kamenica, 2010). Moreover, most heuristics tend to bias decisions systematically in certain directions.[9] In our experiment, the buy-low and sell-high heuristic could reduce the WTP and increase the WTA, and thus result in the WTA-WTP gap of the lotteries and the sure payment. In addition, recent studies have increasingly recognized that subjects may find choice objects difficult to value, and the complexity of valuing choice objects may be responsible for some behavioral biases that were previously thought to reflect underlying preferences (Oprea, 2022; Enke et al., 2023). When subjects are uncertain about the precise valuation of a choice object and have to decide, they may act cautiously, which can also be viewed as a heuristic (as with bad-deal aversion; e.g., Weaver and Frederick, 2012). Cerreia-Vioglio et al. (2022) demonstrated that valuation uncertainty and caution together can explain the endowment effect, loss aversion, and the certainty effect. Consistent with this idea, a part of the WTA-WTP gap in our setting arises because valuation uncertainty and caution reduce the WTP and increase the WTA, as demonstrated by the link between the R-range and the WTA-WTP gap.

Although the framework that links complexity and heuristics has gained increasing attention in the behavioral literature, it faces several significant challenges that are also present in the current study. First, there is no well-established approach to reveal and quantify the complexity that subjects perceive. Consistent with the method in Cettolin and Riedl (2019), Agranov and Ortoleva (2023), and Halevy et al. (2023), this study employed randomization to measure valuation complexity. However, there are alternative methods. For example, some studies have focused on subjects' decision confidence (Dubourg et al.,

[9]For example, the heuristic of choosing in the middle could bias choices in the price list at the center of the list (Bosch-Domènech and Silvestre, 2013). The heuristic of maintaining the status quo unless offered a strictly better alternative could generate the endowment effect and the default bias (Masatlioglu and Ok, 2005; Sautua, 2017). The heuristic of anchor-and-adjustment could induce insufficient responses to information, resulting in anchoring effects and the over/underreaction of prices in financial markets (Hogarth and Einhorn, 1992).

1994; Butler and Loomes, 2007; Cubitt et al., 2015; Enke and Graeber, 2021) or cognitive uncertainty (Enke and Graeber, 2023). Others measure valuation complexity with respect to lotteries and temporal payments based on subjects' calculation errors in algebraic choices, which serve as no-risk-no-time mirrors of risky or temporal choices (Oprea, 2022; Enke et al., 2023). These elicitation methods reflect different conceptualizations of complexity, such as incomplete preferences, preference imprecision, preference uncertainty, cognitive noise, and cognitive uncertainty. While there is evidence that different measures of valuation complexity are significantly correlated with each other and thus may have similar behavioral contents (Arts et al., 2020), it is desirable to have a unified framework to precisely delineate the conceptual disparities and behavioral consequences of these distinct methods (Gabaix and Graeber, 2023).

It can also be difficult to measure mechanism complexity. A popular approach for eliciting mechanism complexity, which we adopted here, is to use objects with known preferences and measure complexity with deviations from these known preferences (Cason and Plott, 2014). However, the objects with known preferences differ from actual choice objects. This implies that we do not directly measure mechanism complexity for the actual choice objects. This difference is important when mechanism complexity could interact with valuation complexity, e.g., mechanism complexity may be more severe when objects are more complex in our setting. The potential interaction between the two forms of complexity can also complicate the elicitation of valuation complexity, as it takes place in an elicitation mechanism. In our investigation, we find that our subjects faced substantial mechanism complexity when we elicited their R-ranges, although it appears that mechanism complexity did not systematically bias the impact of R-range on the WTA-WTP gap. Further, there is evidence that our mechanism for eliciting valuation complexity underestimates it, likely because a large group of subjects found the elicitation mechanism too complex and opted for simplistic choice rules. Research in this area has been rare, and future investigation is needed.

Second, it is difficult to identify the heuristics subjects use when facing complexity. There are many heuristics, and their application may depend on the decision situation and individual characteristics. Price list elicitation can result in several heuristics including the central tendency (Bosch-Domènech and Silvestre, 2013), relying on a randomly generated anchor (Ariely et al., 2003), emphasizing the first choice in the choice list (Hermann and Musshoff, 2016), the buy-low and sell-high heuristic (Kahneman et al., 1982), or caution in the face of valuation complexity (Cerreia-Vioglio et al., 2022). We controlled for the central tendency and anchoring heuristics by using the same ranges of sure payments on the WTA and WTP questions and randomizing the order (ascending or descending) of the sure payments in the price list. Additionally, the WTA-WTP gap of the sure payment is consistent with the buy-low and sell-high heuristic but not with caution. While these results suggest that the buy-low and sell-high heuristic may be one important driving force behind the WTA-WTP gap of the lotteries, we leave it for future studies to understand the effects of multiple heuristics on the elicitation of WTA and WTP.

A further challenge relates to separating heuristics from noisy responses. For example, while the tendency to switch in the middle of the choice lists and to choose the middle of the budget lines is in line with heuristic rules, it can also be viewed as noisy responses (Choi et al., 2006; Enke and Graeber, 2023; Halevy and Mayraz, 2022). In our experiment, we find that many subjects had a negative WTA-WTP gap for the sure payment. Further analysis indicates that these negative WTA-WTP gaps may result from noisy responses. However, without additional information, we are unable to conclusively differentiate noisy responses from heuristics. Consequently, our control of complexity on preference inference is likely to be too conservative.

Our study adds to the emerging literature on the effect of complexity on choice behavior and points to the need to separate preferences from heuristics and noise. Earlier works in behavioral economics have provided substantial insights into the identification of heuristics in various situations (Simon, 1955; Kahneman and Tversky, 1974). Recent studies have provided more concrete evidence and formal analysis with respect to differentiating preferences from noise and biases (Cerreia-Vioglio et al., 2022; Oprea, 2022; Enke and Graeber, 2023; Enke et al., 2023). These studies shed light on the importance of unraveling complexity as a fundamental aspect of decision-making processes and call for more research to provide comprehensive answers to these challenges and deepen our understanding of complexity and decision-making.

References

Agranov, M. and Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68.

Agranov, M. and Ortoleva, P. (2023). Ranges of preferences and randomization. *Review of Economics and Statistics*, Forthcoming.

Anagol, S., Balasubramaniam, V., and Ramadorai, T. (2018). Endowment Effects in the Field: Evidence from India's IPO Lotteries. *The Review of Economic Studies*, 85(4):1971–2004.

Apicella, C. L., Azevedo, E. M., Christakis, N. A., and Fowler, J. H. (2014). Evolutionary origins of the endowment effect: Evidence from hunter-gatherers. *American Economic Review*, 104(6):1793–1805.

Ariely, D., Loewenstein, G., and Prelec, D. (2003). "Coherent Arbitrariness": Stable Demand Curves Without Stable Preferences*. *The Quarterly Journal of Economics*, 118(1):73–106.

Arts, S., Ong, Q., and Qiu, J. (2020). Measuring subjective decision confidence. MPRA Paper 106811, University Library of Munich, Germany.

Baillon, A., Bleichrodt, H., Huang, Z., and van Loon, R. P. (2017). Measuring ambiguity attitude: (Extended) multiplier preferences for the American and the Dutch population. *Journal of Risk and Uncertainty*, 54(3):269–281.

Bayrak, O. K. and Hey, J. D. (2020). Decisions under risk: Dispersion and skewness. *Journal of Risk and Uncertainty*, 61(1):1–24.

Becker, G., DeGroot, M., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9:226–232.

Bewley, T. F. (2002). Knightian decision theory: Part i. *Decisions in Economics and Finance*, 25:79–110.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience in experimental tests of the endowment effect. *American Economic Review*, 102(3):47–52.

Bosch-Domènech, A. and Silvestre, J. (2013). Measuring risk aversion with lists: a new bias. *Theory and Decision*, 75(4):465–496.

Brown, J. R., Kapteyn, A., Luttmer, E. F. P., and Mitchell, O. S. (2017). Cognitive Constraints on Valuing Annuities. *Journal of the European Economic Association*, 15(2):429–

462.

Brown, J. R., Kapteyn, A., Luttmer, E. F. P., Mitchell, O. S., and Samek, A. (2021). Behavioral Impediments to Valuing Annuities: Complexity and Choice Bracketing. *The Review of Economics and Statistics*, 103(3):533–546.

Burger, J. M. and Cooper, H. M. (1979). The desirability of control. *Motivation and emotion*, 3(4):381–393.

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

Cason, T. and Plott, C. (2014). Misconceptions and game form recognition: Challenges to theories of revealed preference and framing. *Journal of Political Economy*, 122(6):1235 – 1270.

Cerreia-Vioglio, S., Dillenberger, D., and Ortoleva, P. (2022). Caution and Reference Effects. PIER Working Paper Archive 22-015, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Cerulli-Harms, A., Goette, L., and Sprenger, C. (2019). Randomizing endowments: An experimental study of rational expectations and reference-dependent preferences. *American Economic Journal: Microeconomics*, 11(1):185–207.

Cettolin, E. and Riedl, A. (2019). Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization. *Journal of Economic Theory*, 181:547 – 585.

Chapman, J., Dean, M., Ortoleva, P., Snowberg, E., and Camerer, C. (2021). On the relation between willingness to accept and willingness to pay.

Cherchye, L., Demuynck, T., De Rock, B., and Vermeulen, F. (2017). Household consumption when the marriage is stable. *American Economic Review*, 107(6):1507–34.

Chew, S. H., Miao, B., Shen, Q., and Zhong, S. (2022). Multiple-switching behavior in choice-list elicitation of risk preference. *Journal of Economic Theory*, 204:105510.

Chew, S. H., Ratchford, M., and Sagi, J. S. (2018). You need to recognise ambiguity to avoid it. *The Economic Journal*, 128(614):2480–2506.

Choi, S., Fisman, R., Gale, D., and Kariv, S. (2006). Substantive and procedural rationality in decisions under uncertainty. *ESRC Centre for Economic Learning and Social Evolution*.

Condon, D. M. and Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence*, 43:52–64.

Cubitt, R., Navarro-Martinez, D., and Starmer, C. (2015). On preference imprecision. *Journal of Risk and Uncertainty*, 50(1):1–34.

Danan, E. and Ziegelmeyer, A. (2006). Are preferences complete? an experimental measurement of indecisiveness under risk. Papers on strategic interaction, Max Planck Institute of Economics, Strategic Interaction Group.

Dimmock, S., Kouwenberg, R., and Wakker, P. (2016). Ambiguity attitudes in a large representative sample. *Management Science*, 62(5):1363–1380.

Dubourg, W. R., Jones-Lee, M. W., and Loomes, G. (1994). Imprecise preferences and the wtp-wta disparity. *Journal of Risk and Uncertainty*, 9(2):115–33.

Eliaz, K. and Ok, E. A. (2006). Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior*, 56(1):61–86.

Engelmann, D. and Hollard, G. (2010). Reconsidering the effect of market experience on the "endowment effect". *Econometrica*, 78(6):2005–2019.

Enke, B. and Graeber, T. (2021). Cognitive uncertainty in intertemporal choice. Working paper, Harvard University.

Enke, B. and Graeber, T. (2023). Cognitive uncertainty. *Quarterly Journal of Economics*, forthcoming.

Enke, B., Graeber, T., and Oprea, R. (2023). Complexity and time. NBER Working Papers 31047, National Bureau of Economic Research, Inc.

Ericson, K. and Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *The Quarterly Journal of Economics*, 126(4):1879–1907.

Ericson, K. and Fuster, A. (2014). The endowment effect. *Annual Review of Economics*, 6(1):555–579.

Esponda, I. and Vespa, E. (2023). Contingent Thinking and the Sure-Thing Principle: Revisiting Classic Anomalies in the Laboratory. *The Review of Economic Studies.*

Fama, E. and French, K. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fehr, D., Fink, G., and Jack, B. K. (2022). Poor and Rational: Decision-Making under

Scarcity. *Journal of Political Economy*, 130(11):2862–2897.

Fehr, D. and Kübler, D. (2022). The endowment effect in the general population. *CESifo Working Paper*, (10150).

Frydman, C. and Jin, L. J. (2021). Efficient coding and risky choice. *The Quarterly Journal of Economics*, 137(1):161–213.

Fudenberg, D., Iijima, R., and Strzalecki, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica*, 83(6):2371–2409.

Gabaix, X. and Graeber, T. (2023). The complexity of economic decisions. Technical report, Harvard University.

Gebhardt, W. A. and Brosschot, J. F. (2002). Desirability of control: Psychometric properties and relationships with locus of control, personality, coping, and mental and somatic complaints in three dutch samples. *European journal of personality*, 16(6):423–438.

Gigerenzer, G., Todd, P. M., and the ABC Research Group (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Gillen, B., Snowberg, E., and Yariv, L. (2019). Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study. *Journal of Political Economy*, 127(4):1826–1863.

Gonczarowski, Y. A., Heffetz, O., and Thomas, C. (2022). Strategyproofness-Exposing Mechanism Descriptions. Papers 2209.13148, arXiv.org.

Goossens, J. T. and Knoef, M. (2022). Covid-19 crisis: Do extreme events affect preferences and trading behavior? Technical report, Tilburg University.

Halevy, Y. and Mayraz, G. (2022). Identifying Rule-Based Rationality. *The Review of Economics and Statistics*, pages 1–44.

Halevy, Y., Walker-Jones, D., and Zrill, L. (2023). Difficult decisions. Working papers, University of Toronto, Department of Economics.

Hermann, D. and Musshoff, O. (2016). Measuring time preferences: Comparing methods and evaluating the magnitude effect. *Journal of Behavioral and Experimental Economics*, 65:16–26.

Hogarth, R. M. and Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24(1):1–55.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic* *review*, 92(5):1644–1655.

Isoni, A., Loomes, G., and Sugden, R. (2011). The willingness to pay—willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations: Comment. *American Economic Review*, 101(2):991–1011.

Iyengar, S. S. and Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, 94(7-8):530–539.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, 98(6):1325–48.

Kahneman, D., Slovic, P., and Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.

Kahneman, D. and Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131.

Khaw, M. W., Li, Z., and Woodford, M. (2021). Cognitive imprecision and small stakes risk aversion. *The review of economic studies*, 88(4):1979–2013.

Knetsch, J. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, 79(5):1277–84.

Li, S. (2017). Obviously strategy-proof mechanisms. *The American Economic Review*, 107(11):3257–3287.

List, J. A. (2003). Does Market Experience Eliminate Market Anomalies? *The Quarterly Journal of Economics*, 118(1):41–71.

List, J. A. (2011). Does market experience eliminate market anomalies? the case of exogenous market experience. *American Economic Review*, 101(3):313–17.

Loomes, G., Starmer, C., and Sugden, R. (2003). Do anomalies disappear in repeated markets? *The Economic Journal*, 113(486):C153–C166.

Luce, R. and Suppes, P. (1965). *Preference, Utility, and Subjective Probability*, chapter 3. in Handbook of Mathematical Psychology, Wiley, New York.

Martínez-Marquina, A., Niederle, M., and Vespa, E. (2019). Failures in contingent reasoning: The role of uncertainty. *American Economic Review*, 109(10):3437–3474.

Masatlioglu, Y. and Ok, E. (2005). Rational choice with status quo bias. *Journal of Economic Theory*, 121(1):1–29.

McGranaghan, C. and Otto, S. (2022). Choice uncertainty and the endowment effect. *Journal of Risk and Uncertain*, 65:83–104.

Nishimuray, H. and Ok, E. (2018). Preference structures. Working paper, New York University.

Ok, E. A., Ortoleva, P., and Riella, G. (2012). Incomplete Preferences Under Uncertainty: Indecisiveness in Beliefs versus Tastes. *Econometrica*, 80(4):1791–1808.

Oprea, R. (2022). simplicity equivalents. Working papers, University of California.

Ortoleva, P. (2010). Status quo bias, multiple priors and uncertainty aversion. *Games and Economic Behavior*, 69(2):411–424.

Plott, C. R. and Zeiler, K. (2005). The willingness to pay-willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. *The American Economic Review*, 95(3):pp. 530–545.

Plott, C. R. and Zeiler, K. (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? *The American Economic Review*, 97(4):1449–1466.

Plott, C. R. and Zeiler, K. (2011). The willingness to pay—willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations: Reply. *American Economic Review*, 101(2):1012–28.

Rees-Jones, A. and Shorrer, R. (2023). Behavioral economics in education market design: A forward-looking review. *Journal of Political Economy Microeconomics*, 1(3):557–613.

Sautua, S. I. (2017). Does uncertainty cause inertia in decision making? an experimental study of the role of regret aversion and indecisiveness. *Journal of Economic Behavior and Organization*, 136:1 – 14.

Sheng, F., Wang, R., Liang, Z., Wang, X., and Platt, M. L. (2023). The art of the deal: Deciphering the endowment effect from traders' eyes. *Science Advances*, 9(34):eadf2115.

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

Smitizsky, G., Liu, W., and Gneezy, U. (2021). The endowment effect: Loss aversion or a buy-sell discrepancy? *Journal of Experimental Psychology: General*, pages 1–11.

Snowberg, E. and Yariv, L. (2021). Testing the waters: Behavior across participant pools. *American Economic Review*, 111(2):687–719.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy*, 123(6):1456–1499.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of economic behavior & organization*, 1(1):39–60.

Vieider, F. M. (2021). Decisions under uncertainty as bayesian inference on choice options. Technical report, Ghent University.

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

Weaver, R. and Frederick, S. (2012). A reference price theory of the endowment effect. *Journal of Marketing Research*, 49:696 – 707.

Woodford, M. (2020). Modeling imprecision in perception, valuation, and choice. *Annual Review of Economics*, 12(1):579–601.

Appendices

A Appendix: Theoretical analysis of mechanism complexity and the R-range

As explained in the experimental design, we elicited a measure of mechanism complexity and valuation complexity for each subject. Below we provide an illustration of how these two factors relate to the WTA-WTP gap and how to control the WTA-WTP gap for these two measures.

A.1 Correcting theoretically the WTA-WTP gap for mechanism complexity

For analytical convenience, we assume a linear utility function. The decision maker perceives it as a loss when giving up the lottery and as a gain when receiving the lottery. It can then be shown that $WTP = EV_X/\lambda$ and $WTA = \lambda EV_X$, where EV_X denote the expected value of the lottery and λ the loss aversion coefficient. In line with Cason and Plott (2014), we assume the decision maker is subject to mechanism complexity of the first price auction when reporting WTA and WTP. Let $0 \leq \alpha_i \leq 1$ denote the degree of mechanism complexity, with a small α indicating a weaker mechanism complexity, and i=a for WTA and p for WTP. Let p denote a price, $\underline{p}$ and $\bar{p}$ the minimum and the maximum prices in the price list. The expected payoff (π) of stating WTP or WTA in the price list is, respectively:

$$\begin{aligned}
E(\pi_{WTP}) &= \frac{1}{\bar{p}-\underline{p}}\int_{\underline{p}}^{WTP}\left[EV_X - \lambda\alpha_p WTP - (1-\alpha_p)\lambda p\right]dp \\
&= \frac{1}{\bar{p}-\underline{p}}\left\{(WTP-\underline{p})\left[EV_X - \lambda\alpha_p WTP\right] - (1-\alpha_p)\lambda(1/2WTP^2 - 1/2\underline{p}^2))\right\} \\
E(\pi_{WTA}) &= \frac{1}{\bar{p}-\underline{p}}\int_{WTA}^{\bar{p}}\left[(\alpha_a WTA + (1-\alpha_a)p - \lambda EV_X\right]dp \\
&= \frac{1}{\bar{p}-\underline{p}}\left\{(\bar{p}-WTA)\left[\alpha_a WTA - \lambda EV_X\right] - (1-\alpha_a)(\bar{p}^2/2 - WTA^2/2)\right\}
\end{aligned}$$

The optimal WTA and WTP that the decision maker reports are then:

$$\begin{aligned} WTP &= \frac{1}{(1+\alpha_p)} EV_X/\lambda + \frac{\alpha_p}{(1+\alpha_p)}\underline{p} \\ WTA &= \frac{1}{1+\alpha_a}\lambda EV_X + \frac{\alpha_a}{1+\alpha_a}\bar{p} \end{aligned}$$

When X is a sure payment, the decision maker exchanges between sure payments and there should be no loss aversion. Thus, if the decision maker correctly understands the incentives in the price list, she should report WTA=WTP=500 for the sure payment. Deviating from 500 implies mechanism complexity. We can infer the decision maker's degree of mechanism complexity about the price list (α_i) from the reported WTA_{sure} and WTP_{sure} for the sure payment as:

$$\begin{aligned} \alpha_p &= \frac{500 - WTP_{sure}}{WTP_{sure} - 400} \\ \alpha_a &= \frac{WTA_{sure} - 500}{600 - WTA_{sure}}, \end{aligned}$$

where 400 is $\underline{p}$ and 600 $\bar{p}$ of the sure payment. Correcting for mechanism complexity, the preference-based WTP^* and WTA^* for the lottery are respectively:

$$\begin{aligned} WTP^* &= EV_X/\lambda = WTP + \alpha_p(WTP - \underline{p}) \\ WTA^* &= \lambda EV_X = WTA - \alpha_a(\bar{p} - WTA), \end{aligned}$$

where $\underline{p} = 0$ and $\bar{p} = 1000$ for the two lotteries in our experiment. Inserting these values in the above equations, the underlying WTP^* and WTA^* corrected for mechanism complexity are:

$$\begin{aligned} WTP^* &= \frac{100}{WTP_{sure} - 400} WTP \\ WTA^* &= \frac{100}{600 - WTA_{sure}} WTA - \frac{WTA_{sure} - 500}{600 - WTA_{sure}} 1000. \end{aligned}$$

Note that $WTP^* = WTP$ and $WTA^* = WTA$ when $WTP_{sure} = 500$ and $WTA_{sure} = 500$.

A.2 What the R-range reveals

There are a number of ways to relate the R-range to valuation complexity. For example, the literature of incomplete preferences assumes that decision makers may have incomplete preferences over some choices, and they choose to randomize between the options when indecisive (Bewley, 2002; Eliaz and Ok, 2006; Cettolin and Riedl, 2019; Halevy et al., 2023). Below we discuss how the R-range captures valuation complexity in the framework of Fudenberg et al. (2015). We then motivate our statistical analysis about how to control the WTA-WTP gap for valuation complexity.

When facing a lottery X and a price y, a decision maker behaving according to Fudenberg et al.'s (2015) functional form chooses the optimal randomization probability as

$$p^* = \text{argmax}_p \quad pU(X) + (1-p)U(y) - c(p) - c(1-p),$$

where $U(\cdot)$ is the Von Neumann–Morgenstern expected utility, $u(\cdot)$ is the Bernoulli utility, and $c(p)$ is a convex and continuously differentiable cost function. Fudenberg et al. (2015) show that their representation corresponds to a form of an uncertainty averse decision maker who is unsure about her true utility and uses randomization to hedge her preference uncertainty. When the decision maker faces the lottery X, a price y, and a randomization option where the decision maker receives either option with probability 0.5, she may strictly prefer the randomization option over X or y. To see this, notice that the first order condition of the above equation is simply $c'(p) - c'(1-p) = U(X) - u(y)$, which predicts that the optimal p is close to 0.5 when $U(X) - U(y)$ is small.

We can infer the level of valuation complexity that the decision maker perceives over the choice problem between X and y from the range of prices y that the decision maker chooses the randomization option. Define the smallest ($\underline{y}$) and largest ($\bar{y}$) prices of y that such that the decision maker prefers the randomization option as

$$\begin{aligned}
u(\underline{y}) + 0.5[U(X) - u(\underline{y})] - 2c(0.5) &\geq U(X) - c(1.0) - c(0) \\
\Rightarrow u(\underline{y}) &= U(X) - 2[c(1.0) + c(0) - 2c(0.5)], \\
u(\bar{y}) + 0.5[U(X) - u(\bar{y})] - 2c(0.5) &\geq u(\bar{y}) - c(1.0) - c(0), \\
\Rightarrow u(\bar{y}) &= U(X) + 2[c(1.0) + c(0) - 2c(0.5)].
\end{aligned}$$

The term $2[c(1.0) + c(0) - 2c(0.5)] = \frac{c(1.0)-c(0.5)}{1-0.5} - \frac{[c(0.5)-c(0)]}{0.5-0}$ measures the convexity of the cost function, which, according to Fudenberg et al. (2015), is positively related to valuation complexity. This is most clear when the cost function takes the form of $c(p) = p\log(p)/\eta$. Simple calculation shows that the optimal randomization probability is the familiar logit/logistic choice rule: $p^* = \frac{e^{\eta U(X)}}{e^{\eta U(X)} + e^{\eta u(y)}}$. As shown by Holman and

Marley, the parameter $1/\eta$ can be linked to the variance of the i.i.d. Gumbel preference shocks in a random utility representation (Luce and Suppes, 1965, p.338), with a larger η corresponding to a smaller valuation complexity. In light of Fudenberg et al. (2015), the cost function may also relate to a measure of control desirability because, as Fudenberg et al. (2015) point out, a preference for randomization arises because the decision maker trades off the probability of errors against the cost of making the desired choice. A decision maker who prefers to maintain controls may perceive the implementation cost to be small and is less willing to randomize.

Let $\Delta = 2[c(1.0)+c(0)-2c(0.5)]$. It follows that we can estimate a measure of valuation complexity Δ from the R-range as $\Delta = \frac{u(\bar{y})-u(\underline{y})}{2}$. Note that the "true" valuation of the lottery is $U(X) = [u(\bar{y})+u(\underline{y})]/2$. When the decision maker perceives valuation complexity about the lottery and is asked for WTP and WTA, she behaves cautiously and understates her WTP and overstates her WTA, respectively (Cerreia-Vioglio et al., 2022), resulting a positive WTA-WTP gap. This could arise even when the decision maker is not loss averse. One way to capture the above mechanism is:

$$\begin{aligned} u(WTP) &= U(X) - \delta_p \Delta \\ u(WTA) &= U(X) + \delta_a \Delta, \end{aligned}$$

where $\delta_i, i = a, b$ denotes the decision maker's cautious attitude toward valuation complexity when stating WTP and WTA. This motivates the following analysis in the main text:

$$\begin{aligned} WTA_X &= I + \beta_1 \text{R-range}_X + \epsilon \\ WTP_X &= I + \beta_2 \text{R-range}_X + \epsilon, \end{aligned}$$

where I is the intercept, X is L1, L2, or the average of the two lotteries.

B Appendix: Correcting the R-range for mechanism complexity

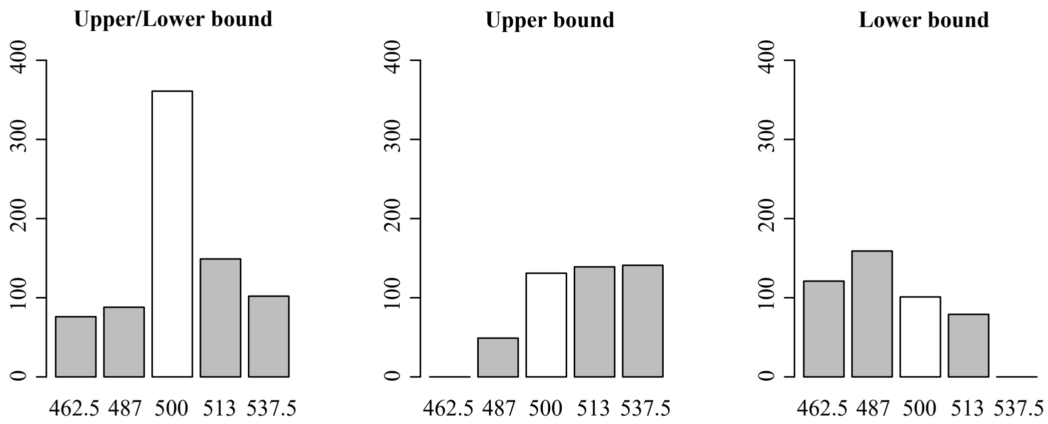


Figure B.1: Distribution of the upper bound and lower bound of the sure payment. The left panel shows the distribution of the bounds for subjects who do not randomize for the sure payment. The right two panels show the distribution of the bounds for subjects who randomize at least once for the sure payment. The white bars show the optimal response of 500 points for the upper and lower bound.

To correct the R-range for mechanism complexity, we use subjects' two bounds of the R-range for the sure payment. The optimal upper bound and lower bound for the sure payment are both 500, and subjects should never choose the randomization option. Deviating from these optimal responses implies mistakes or biases. When subjects never chose the randomization option but had non-optimal bounds for the sure payment, it is unclear how these non-optimal responses affect the R-range systematically. The left histogram in Figure B.1 suggests that these non-optimal responses have no systematic pattern. We, therefore, make no correction of the R-range for these deviations from the optimal response. On the other hand, when subjects chose the randomization option at least once and thus have an R-range for the sure payment, this could represent a systematic bias and may increase the R-range for the lotteries if subjects would exhibit the same tendency in the R-range task for the lotteries. Indeed, the right two histograms in Figure B.1 suggest that these non-optimal responses are systematically biased: The upper bound is systematically biased upward and the lower bound is systematically biased downward.

We correct the R-range for mechanism complexity by using deviations from the optimal

	Lottery	Before correction	Correcting for deviations
Upper Bound	L1	551.62	527.69**
		(233.61)	(226.85)
	L2	538.55	513.96**
		(229.99)	(223.89)
	Average	545.09	519.83**
		(209.22)	(202.08)
Lower Bound	L1	449.14	467.84***
		(239.10)	(233.02)
	L2	433.10	451.94***
		(237.23)	(230.98)
	Average	441.12	459.89***
		(213.65)	(206.76)
R-range	L1	102.48	57.85***
		(172.42)	(160.64)
	L2	105.45	62.02***
		(174.91)	(160.66)
	Average	103.96	59.94***
		(149.59)	(134.24)

Table B.1: The upper bound, the lower bound, and the R-range before and after correction for mechanism complexity. We only correct the upper bound and lower bound for subjects who chose the randomization option at least once for the sure payment. For subjects who did not choose the randomization option the uncorrected upper and lower bound is included in the computation of the mean and standard deviation after correction. Wilcoxon signed-rank tests were performed to test the significance of the difference between the upper bound, the lower bound, and the R-range before and after the correction for mechanism complexity, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

responses. In these statistical correction regressions, we exclude subjects who never chose the randomization option but reported non-optimal bounds for the sure payment. As discussed above, their non-optimal responses do not systematically bias the R-range, and it is unclear how to correct these R-ranges. We use their uncorrected R-range in further analysis.

Table B.1 reports the corrected bounds and R-range of the two lotteries separately and the averages. We see that mechanism complexity correction significantly lowers the upper bound (average of the two lotteries: 519.83 vs 545.09, $p < 0.05$) and increases the lower bound (average of the two lotteries: 459.89 vs 441.12, $p < 0.01$). This leads to significantly smaller R-ranges (average of the two lotteries: 59.94 vs 103.96, $p < 0.01$).

C Appendix: Additional tables

	WTA-WTP sample	Neutral frame sample
Sample size	1236	620
Female	47.17	48.06
Age		
16-34	10.76	11.13
35-49	19.74	15.97
50-64	33.25	24.52
65 and above	36.25	48.39
Education level		
Low	24.19	24.35
Medium	33.82	29.35
High	41.75	46.13
Household income		
0-2499	22.25	21.29
2500-3999	24.68	26.45
4000-5999	22.57	24.68
6000 and above	25.16	24.52
Occupation		
Paid work	50.97	42.26
House work	7.36	7.58
Retired	30.26	40.97
Others	11.41	9.19
Household composition		
Partner	66.91	63.06
# of children	0.49	0.37

Table C.1: Sample composition of the LISS panel members who participated in the experiment. The numbers represent the percentage of subjects in each demographic group, except for the number of children, which is the average. Column 1 shows subjects from the within-subject design in which they reported their WTA, WTP, and R-range. Column 2 shows the subject who completed the neutrally framed questions.

	L1	L2	Average	Sure payment
WTA	527.54	519.40	523.47	505.08
	(243.92)	(247.26)	(226.40)	(20.48)
WTP	415.32	407.70	411.51	496.63
	(252.29)	(253.78)	(234.84)	(22.66)
The WTA-WTP gap	112.22***	111.70***	111.96***	8.45***
	(275.34)	(275.10)	(242.99)	(27.07)
CE neutral	476.55	465.12	470.84	
	(215.96)	(214.47)	(199.35)	

Table C.2: Means of WTA, WTP, WTA-WTP gaps, and CE in the neutral frame for L1, L2 and the average across the two lotteries, as well as the means of WTA, WTP, and WTA-WTP gap of the sure payment. Wilcoxon signed-rank tests were performed to test the significance of the WTA-WTP gap. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dependent variable	$\text{WTA}_{sure} = 500$	$\text{WTP}_{sure} = 500$	WTA-WTP $\text{gap}_{sure} = 0$
Independent variables		Log odds	
Matrix reasoning score	0.26***	0.26***	0.10***
	(0.04)	(0.04)	(0.03)
Education	0.24***	0.14***	0.03
	(0.04)	(0.05)	(0.04)
Self-reported understanding	0.20***	0.21***	0.16***
	(0.05)	(0.05)	(0.04)
Response time	0.04**	0.04**	−0.02*
	(0.02)	(0.02)	(0.01)
Intercept	−2.97***	−2.92***	−1.01***
	(0.26)	(0.27)	(0.23)
Efron's R^2	0.12	0.10	0.02

Table C.3: Logistic regressions with the optimal response as the dummy dependent variable (1 indicates the optimal response: $\text{WTA}_{sure} = 500$, $\text{WTP}_{sure} = 500$, and WTA-WTP $\text{gap}_{sure} = 0$). The independent variables include subjects' IQ (their score on the matrix reasoning questions), education level (categorical, from 1 to 6), self-reported understanding of the experimental questions (likert scale from 1 to 5), and time spent on the corresponding task (time on the WTA questions, time on the WTP questions, and total time on both the WTA and WTP questions for the sure payment). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Lottery	Upper bound	Lower bound	R-range
L1	551.62	449.14	102.48
	(233.61	(239.10)	(172.42)
L2	538.55	433.10	105.45
	(229.99)	(237.23)	(174.91)
Average	545.09	441.12	103.96
	(209.22)	(213.65	(149.59)

Lottery	Randomizing frequency:		
	0	1	2 or more
L1	121.05	67.00	127.19
	(295.71)	(237.75)	(265.57)
L2	118.13	83.65	121.76
	(280.48)	(247.70)	(284.37)
	Randomizing frequency:		
	0	1 or 2	3 or more
Total in L1 and L2	122.04	79.61	122.07
	(255.09)	(226.17)	(241.32)

Table C.4: The left panel reports the mean upper bounds, lower bound, and the R-range for L1, L2, and the averages across lotteries. The right panel reports the mean WTA-WTP gap for subjects who chose the randomization option 0 times, 1 time, and 2 times or more for each of the corresponding lottery (and 0, 1 or 2, and 3 times or more, respectively, for the total times randomizing in L1 and L2). Standard deviations are presented in parentheses.

	Lottery	Before correction	Correcting theoretically	Correcting for deviations	Correcting for deviations different slopes
WTA	L1	527.54	504.70***	500.09***	488.82***
		(243.92)	(257.55)	(217.37)	(217.02)
	L2	519.40	491.02***	495.82	490.50***
		(247.26)	(274.85)	(228.26)	(228.18)
	Average	523.47	497.86***	497.95***	489.66***
		(226.40)	(240.35)	(201.69)	(201.49)
WTP	L1	415.32	425.50	431.97**	438.02***
		(252.29)	(259.21)	(226.07)	(226.00)
	L2	407.70	417.90	424.30**	418.60**
		(253.78)	(264.37)	(227.86)	(227.80)
	Average	411.51	421.70	428.13**	428.31***
		(234.84)	(238.00)	(206.49)	(206.49)
WTA-WTP gap	L1	112.22	79.20***	68.12***	50.81***
		(275.34)	(312.22)	(264.27)	(264.18)
	L2	111.70	73.12***	71.52***	71.90***
		(275.10)	(324.00)	(264.77)	(265.03)
	Average	111.96	76.16***	69.82***	61.35***
		(242.99)	(279.24)	(230.79)	(231.06)

Table C.5: WTA, WTP, and WTA-WTP gap before and after each of the three correction methods. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before and after the correction, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Lottery	Before controlling	After controlling for the R-range	After controlling for the corrected R-range
WTA	L1	527.54	519.84***	523.27***
		(243.92)	(243.58)	(243.63)
	L2	519.40	514.30***	517.64***
		(247.26)	(247.11)	(247.21)
	Average	523.47	514.92***	519.13***
		(226.40)	(226.07)	(226.20)
WTP	L1	415.32	421.55***	421.84***
		(252.29)	(252.07)	(251.64)
	L2	407.70	413.37***	413.54***
		(253.78)	(253.60)	(253.33)
	Average	411.51	419.62***	420.44***
		(234.84)	(234.55)	(233.99)
The WTA-WTP gap	L1	112.22	98.29***	101.43***
		(275.34)	(274.34)	(273.71)
	L2	111.70	100.93***	104.10***
		(275.10)	(274.52)	(274.40)
	Average	111.96	95.30***	98.69***
		(242.99)	(241.81)	(241.17)

Table C.6: WTA, WTP, and WTA-WTP gap before and after controlling for the (MC corrected) R-range. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before and after controlling for the (MC corrected) R-range, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Lottery	Before controlling	After controlling for both mechanism complexity and the R-range			
			Model 1	Model 2	Model 3	Model 4
WTA	L1	527.54	496.95***	486.45***	497.57***	486.62***
		(243.92)	(217.30)	(216.97)	(217.25)	(216.92)
	L2	519.40	491.81***	486.69***	493.75***	488.30***
		(247.26)	(228.16)	(228.09)	(228.19)	(228.12)
	Average	523.47	493.06***	485.40***	494.58***	486.37***
		(226.40)	(201.56)	(201.38)	(201.55)	(201.35)
WTP	L1	415.32	437.00***	442.88***	436.89***	443.26***
		(252.29)	(225.91)	(225.84)	(225.64)	(225.57)
	L2	407.70	425.90***	420.21***	426.52***	420.73***
		(253.78)	(227.85)	(227.78)	(227.79)	(227.72)
	Average	411.51	432.68***	432.60***	433.37***	433.52***
		(234.84)	(206.39)	(206.39)	(206.14)	(206.14)
WTA-WTP gap	L1	112.22	59.95***	43.57***	60.68***	43.36***
		(275.34)	(263.83)	(263.79)	(263.33)	(263.24)
	L2	111.70	62.90***	66.48***	67.23***	67.57***
		(275.10)	(264.57)	(264.83)	(264.49)	(264.74)
	Average	111.96	60.38***	52.80***	61.22***	52.84***
		(242.99)	(230.29)	(230.59)	(229.84)	(230.11)

Table C.7: WTA, WTP, and WTA-WTP gap before and after controlling for both mechanism complexity and the R-range. In Model 1 and 3 mechanism complexity is deviations from the optimal response, while in Model 2 and 4 mechanism complexity is deviations from the optimal response while allowing for different slopes for positive and negative deviations. In Model 1 and 2 the R-range is uncorrected, and in Model 3 and 4 the R-range is mechanism complexity corrected. Wilcoxon signed-rank tests were performed to test the significance of the difference between WTA, WTP, and WTA-WTP gap before the correction and after the correction, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Mechanism complexity			No mechanism complexity		
	Valuation complexity			Valuation complexity		
	High	Low	No	High	Low	No
Proportions	23.8%	29.5%	22.5%	7.7 %	8.1%	8.4%
WTA-WTP gap	151.62	103.54	95.78	76.22	0.58	122.78
	(285.83)	(234.38)	(267.05)	(170.66)	(128.25)	(195.53)
Education	3.84	3.58	3.74	4.27	4.16	4.71
	(1.40)	(1.45)	(1.36)	(1.25)	(1.51)	(1.16)
MR performance	3.78	3.22	3.23	4.39	4.33	4.76
	(1.49)	(1.73)	(1.78)	(1.39)	(1.64)	(1.26)
Correlation between λ and the WTA-WTP gap	0.155	0.083	-0.044	-0.068	-0.146	0.507^{***}

Table C.8: The distribution of the subjects with/out each form of complexity, their associated average WTA-WTP gap, education, performance in the matrix reasoning questions, and loss aversion (λ), and the correlation between the WTA-WTP gap and λ. The high and low valuation complexity groups are defined by a median split of the R-range (100 points for subjects with mechanism complexity, and 131.25 points for those without). Education level is measured on a scale from 1 to 6. Standard deviations are in parentheses.

D Appendix: Mechanism complexity and valuation complexity correction with dummy variables for each deviation

In subsection 3.3 we describe three mechanism complexity correction methods. Here, we discuss one additional method based on the same statistical correction model 1. For this correction, we use dummy variables for each deviation of the sure payment's WTA or WTP from 500 to correct for the WTA, WTP, and WTA-WTP gap of the lotteries. Specifically, in statistical model 1, $\mathbf{MC_a} = [537.5, 513, 487, 462.5]^T$ denotes a vector of dummy variables for each possible value of the WTA_{sure} other than the correct response of 500 points, and $\mathbf{MC_p} = [537.5, 513, 487, 462.5]^T$ denotes a vector of dummy variables for each possible value of the WTP_{sure} other than the correct response of 500 points. This correction relaxes the assumption that a higher deviation from 500 for the sure payment corresponds to a higher mechanism complexity. In Table D.1 we present the results of the correction for mechanism complexity only and the joint correction of mechanism complexity and valuation complexity. When controlling for valuation complexity we look at both the uncorrected R-range and the R-range corrected for mechanism complexity. In the latter case, we correct the R-range using dummy variables for each deviation as described above.

Consistent with the corrections reported in the main paper, we observe a significantly smaller WTA-WTP gap after correcting for mechanism complexity with dummy variables for each deviation (64.72 vs 111.96, $p < 0.01$). This is a reduction of about 40%, comparable to the two statistical correction methods we reported in the main text. Also consistent with these earlier corrections, we see that WTA is significantly lower after correction (475.98 vs. 523.47, $p < 0.01$). WTP gives a less straightforward result, as we observe contrasting results for the two lotteries: for L1 we find a higher but insignificant WTP after correction (425.76 vs 415.32, $p > 0.10$), and for L2 we find a significantly lower WTP after correction (396.77 vs 407.70, $p < 0.01$). On average this results in a statistically significant, but only very slightly lower WTP after correction for mechanism complexity with dummy variables (411.27 vs 411.51, $p < 0.01$).

Additionally, including the R-range gives similar results as for the earlier analyses. We find a significant reduction when considering mechanism complexity together with both the uncorrected R-range (56.07 vs 111.96, $p < 0.01$) and the mechanism complexity corrected R-range (55.01 vs 111.96, $p < 0.01$), which corresponds to a reduction of about 50%.

	Lottery	Before correction	Correcting for each deviation: mechanism complexity only	with uncorrected R-range	with corrected R-range
WTA	L1	527.54	478.27***	474.56***	474.45***
		(243.92)	(214.06)	(213.95)	(213.88)
	L2	519.40	473.70***	469.42***	470.10***
		(247.26)	(224.31)	(224.19)	(224.19)
	Average	523.47	475.98***	470.46***	470.43***
		(226.40)	(197.74)	(197.57)	(197.49)
WTP	L1	415.32	425.76	429.99***	431.35***
		(252.29)	(222.41)	(222.28)	(222.11)
	L2	407.70	396.77***	397.04	396.79
		(253.78)	(220.69)	(220.69)	(220.69)
	Average	411.51	411.27***	414.39***	415.42***
		(234.84)	(200.78)	(200.72)	(200.65)
WTA-WTP gap	L1	112.22	52.51***	44.56***	43.10***
		(275.34)	(264.23)	(263.73)	(263.29)
	L2	111.70	76.92***	72.38***	73.30***
		(275.10)	(264.12)	(263.99)	(263.02)
	Average	111.96	64.72***	56.07***	55.01***
		(242.99)	(230.66)	(230.18)	(229.90)

Table D.1: WTA, WTP, and WTA-WTP gap before and after correcting for mechanism complexity by including each deviation as a dummy variable. The first model includes mechanism complexity only, the second and third model additionally control for valuation complexity by including the uncorrected R-range and the corrected R-range, respectively. The corrected R-range controls for mechanism complexity with dummy variables for each deviation. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before and after the correction, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Appendix: Robustness check across performance in matrix reasoning

We performed a robustness check by taking a median split of our sample based on the performance on the matrix reasoning questions (median = 3.5) and performed the analysis separately for those subjects who performed above median (N = 618) and those who performed below median (N = 618).

Above median in matrix reasoning performance

	L1	L2	Average	Sure
WTA	524.19	514.95	519.57	504.30
	(210.59)	(214.93)	(195.33)	(18.37)
WTP	401.27	393.22	397.25	496.05
	(217.04)	(219.37)	(201.21)	(20.85)
The WTA-WTP gap	122.92***	121.72***	122.32***	8.25***
	(252.77)	(261.23)	(230.37)	(24.83)

Below median in matrix reasoning performance

	L1	L2	Average	Sure
WTA	530.89	523.85	527.37	505.86
	(273.35)	(275.93)	(253.80)	(22.38)
WTP	429.37	422.17	425.77	497.21
	(282.67)	(283.49)	(263.63)	(24.33)
The WTA-WTP gap	101.52***	101.68***	101.60***	8.65***
	(296.02)	(288.17)	(254.76)	(29.15)

Table E.1: Means of the WTA, the WTP, and the WTA-WTP gap for L1, L2, the average across the two lotteries, and the sure payment. Wilcoxon signed-rank tests were performed to test the significance of the difference between WTA and WTP. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Above median in matrix reasoning performance

	Lottery	Before correction	Correcting mechanism complexity: theoretically	for deviations	for deviations different slopes
WTA	L1	524.19	502.68***	502.95***	495.83
		(210.59)	(223.98)	(190.02)	(189.77)
	L2	514.95	489.42***	496.48*	492.10
		(214.93)	(239.42)	(199.92)	(199.83)
	Average	519.57	496.05***	499.72*	493.96
		(195.33)	(209.61)	(175.94)	(175.77)
WTP	L1	401.27	417.13***	418.42	416.10
		(217.04)	(230.91)	(197.24)	(197.22)
	L2	393.22	409.87**	409.59	399.58***
		(219.37)	(236.48)	(201.61)	(201.28)
	Average	397.25	413.50***	414.01	407.84
		(201.21)	(209.43)	(180.70)	(180.56)
WTA-WTP gap	L1	122.92	85.55***	84.53***	79.73***
		(252.77)	(289.90)	(242.00)	(242.25)
	L2	121.72	79.55***	86.89***	92.51***
		(261.23)	(310.20)	(251.21)	(251.53)
	Average	122.37	82.55***	85.71***	86.12***
		(230.37)	(267.50)	(218.72)	(219.00)

Below median in matrix reasoning performance

	Lottery	Before correction	Correcting mechanism complexity: theoretically	for deviations	for deviations different slopes
WTA	L1	530.89	506.72***	497.37***	478.93***
		(273.35)	(287.40)	(241.55)	(241.00)
	L2	523.85	492.63***	495.29***	488.43***
		(275.93)	(306.40)	(253.48)	(253.40)
	Average	527.37	499.67***	496.33	483.68***
		(253.80)	(267.75)	(224.45)	(224.16)
WTP	L1	429.37	433.87	444.31*	469.96***
		(282.67)	(284.66)	(250.77)	(249.96)
	L2	422.17	425.94*	437.45***	446.16***
		(283.49)	(289.56)	(250.12)	(250.03)
	Average	425.77	429.90	440.88***	458.06***
		(263.63)	(263.41)	(228.23)	(227.83)
WTA-WTP gap	L1	101.52	72.85	53.06***	8.97***
		(296.02)	(333.17)	(284.87)	(283.46)
	L2	101.68	66.69*	57.83***	42.27
		(288.17)	(337.36)	(278.28)	(278.78)
	Average	101.60	69.77*	55.45***	25.62***
		(254.76)	(290.58)	(242.59)	(242.58)

Table E.2: WTA, WTP, and WTA-WTP gap before and after each of the three correction methods. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before the correction and after the correction, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Above median in matrix reasoning performance

	Lottery	Before controlling for R-range	After controlling for R-range	After controlling for mechanism complexity corrected R-range
WTA	L1	524.19	515.01***	519.59***
		(210.59)	(210.08)	(210.33
	L2	514.95	516.00***	518.03***
		(214.93)	(214.92)	(214.80)
	Average	519.57	512.68***	517.36***
		(195.33)	(195.10)	(195.28)
WTP	L1	401.27	401.85***	405.62***
		(217.04)	(217.04)	(216.81)
	L2	393.22	403.55***	404.60***
		(219.37)	(218.74)	(217.66)
	Average	397.25	404.74***	408.44***
		(201.21)	(200.95)	(199.98)
WTA-WTP gap	L1	122.92	113.16***	113.97***
		(252.77)	(252.28)	(251.94)
	L2	121.72	112.45***	113.42***
		(261.23)	(260.80)	(260.46)
	Average	122.37	107.94***	108.91***
		(230.37)	(229.50)	(228.82)

Below median in matrix reasoning performance

	Lottery	Before controlling for R-range	After controlling for R-range	After controlling for mechanism complexity corrected R-range
WTA	L1	530.89	524.43***	526.39
		(273.35)	(273.11)	(272.88)
	L2	523.85	513.12***	517.01
		(275.93)	(275.31)	(275.26)
	Average	527.37	517.22***	520.20***
		(253.80)	(253.35)	(253.13)
WTP	L1	429.37	440.16***	435.71***
		(282.67)	(282.02)	(281.78)
	L2	422.17	423.44***	422.48***
		(283.49)	(283.48)	(283.49)
	Average	425.77	434.06***	431.19***
		(263.63)	(263.33)	(263.26)
WTA-WTP gap	L1	101.52	84.27***	90.68
		(296.02)	(294.43)	(293.52)
	L2	101.68	89.68***	94.53***
		(288.17)	(287.44)	(287.48)
	Average	101.60	83.17***	89.01***
		(254.76)	(253.27)	(252.69)

Table E.3: The WTA, the WTP, and the WTA-WTP gap before and after controlling for the (corrected) R-range. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before and after controlling for R-range, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Above median in matrix reasoning performance

	Lottery	Before controlling	After controlling for both forms of complexity			
			Model 1	Model 2	Model 3	Model 4
WTA	L1	524.19	496.83***	491.39***	498.26***	492.04***
		(210.59)	(189.76)	(189.58)	(189.72)	(189.51)
	L2	514.95	499.66*	495.26	499.15	494.81
		(214.93)	(199.85)	(199.75)	(199.82)	(199.73)
	Average	519.57	496.28***	491.46***	497.10***	491.60***
		(195.33)	(175.88)	(175.73)	(175.87)	(175.70)
WTP	L1	401.27	417.97	415.62	420.72***	418.63***
		(217.04)	(197.24)	(197.22)	(197.17)	(197.15)
	L2	393.22	418.4***	408.23***	418.52***	408.78***
		(219.37)	(201.10)	(200.74)	(200.40)	(200.09)
	Average	397.25	419.83***	413.63***	422.07***	416.45***
		(201.21)	(180.51)	(180.37)	(179.95)	(179.84)
WTA-WTP gap	L1	122.92	78.87***	75.76***	77.54***	73.41***
		(252.77)	(241.75)	(242.01)	(241.46)	(241.72)
	L2	121.72	81.21***	87.03***	80.62***	86.03***
		(261.23)	(251.07)	(251.41)	(250.68)	(250.98)
	Average	122.37	76.45***	77.83***	75.03***	75.15***
		(230.37)	(218.28)	(218.63)	(217.63)	(217.92)

Below median in matrix reasoning performance

	Lottery	Before controlling	After controlling for both forms of complexity			
			Model 1	Model 2	Model 3	Model 4
WTA	L1	530.89	496.92***	478.58***	496.19***	477.88***
		(273.35)	(241.55)	(240.99)	(241.51)	(240.96)
	L2	523.85	484.48***	477.48***	488.39***	481.46***
		(275.93)	(252.80)	(252.72)	(252.74)	(252.67)
	Average	527.37	490.04***	477.44***	491.75***	479.22***
		(253.80)	(224.24)	(223.96)	(224.11)	(223.84)
WTP	L1	429.37	453.88***	478.10***	449.86***	474.21***
		(282.67)	(250.19)	(249.44)	(249.98)	(249.24)
	L2	422.17	431.90*	441.19***	433.18	442.82***
		(283.49)	(249.94)	(249.83)	(249.82)	(249.71)
	Average	425.77	443.72***	460.04***	442.58***	459.10***
		(263.63)	(228.19)	(227.81)	(228.18)	(227.80)
WTA-WTP gap	L1	101.52	43.05***	0.49***	46.33***	3.67***
		(296.02)	(284.27)	(282.95)	(283.72)	(282.43)
	L2	101.68	52.58***	36.29***	55.21***	38.64***
		(288.17)	(278.28)	(278.83)	(278.39)	(278.94)
	Average	101.60	46.32***	17.40***	49.17***	20.12***
		(254.76)	(242.09)	(242.13)	(241.88)	(241.93)

Table E.4: The WTA, WTP, and WTA-WTP gap before and after controlling for both forms of complexity in two groups based on their performance in matrix reasoning questions. In Model 1 and 3 mechanism complexity is deviations from the optimal response, while in Model 2 and 4 mechanism complexity is deviations from the optimal response while allowing for different slopes for positive and negative deviations. In Model 1 and 2 the R-range is uncorrected, and in Model 3 and 4 the R-range is mechanism complexity corrected. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before the correction and after the correction, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Above median in matrix reasoning performance (N = 306)

	Lottery	Before controlling	Controlling λ	Combined (Model 3)	Combined (Model 3)+λ	Combined (Model 4)	Combined (Model 4)+λ
WTA	L1	518.91	525.39***	494.15	495.34***	490.16	491.30***
		(213.48)	(213.43)	(189.36)	(189.35)	(189.28)	(189.28)
	L2	516.18	504.59***	507.93	491.44***	501.92	485.37***
		(216.70)	(216.55)	(200.84)	(200.52)	(200.71)	(200.38)
	Average	517.54	514.99***	498.42	491.59***	492.59	485.66***
		(198.57)	(198.56)	(177.70)	(177.64)	(177.54)	(177.48)
WTP	L1	402.04	462.96***	413.57	450.85***	416.48	454.98***
		(209.94)	(208.13)	(185.89)	(185.12)	(185.86)	(185.08)
	L2	395.83	455.95***	423.42	452.51***	410.61	439.05***
		(222.43)	(220.76)	(197.70)	(197.25)	(197.22)	(196.81)
	Average	398.94	459.46***	421.26	456.02***	416.60	451.60***
		(200.63)	(198.75)	(174.95)	(174.23)	(174.88)	(174.17)
WTA -WTP gap	L1	116.87	62.43***	80.59	44.48***	73.67	36.32***
		(247.44)	(247.14)	(233.24)	(233.06)	(233.28)	(233.01)
	L2	120.34	48.64***	84.51	38.93***	91.31	46.32***
		(257.05)	(254.78)	(250.08)	(248.72)	(250.75)	(249.35)
	Average	118.61	55.53***	77.16	35.57***	76.00	34.06***
		(224.42)	(223.01)	(212.53)	(211.66)	(212.96)	(212.01)

Below median in matrix reasoning performance (N = 227)

	Lottery	Before controlling	Controlling λ	Combined (Model 3)	Combined (Model 3)+λ	Combined (Model 4)	Combined (Model 4)+λ
WTA	L1	513.99	507.81***	487.38	479.48***	467.87	461.12***
		(266.69)	(266.65)	(238.10)	(238.03)	(237.42)	(237.36)
	L2	511.89	520.30***	478.13	483.01***	468.66	473.84***
		(274.61)	(274.54)	(250.40)	(250.38)	(250.25)	(250.22)
	Average	512.94	514.06***	482.17	481.00***	468.88	468.42***
		(255.54)	(255.54)	(227.88)	(227.88)	(227.55)	(227.55)
WTP	L1	434.86	425.88***	447.53	424.66***	461.46	439.03***
		(282.08)	(282.05)	(251.76)	(251.48)	(251.49)	(251.25)
	L2	420.21	412.41***	425.16	401.06***	446.46	423.56***
		(279.50)	(279.47)	(251.86)	(251.56)	(251.25)	(251.00)
	Average	427.53	419.15***	435.32	411.37***	454.57	431.56***
		(265.76)	(265.73)	(235.35)	(235.03)	(234.79)	(234.52)
WTA -WTP gap	L1	79.13	81.94***	39.85	54.82***	6.41	22.09***
		(295.06)	(295.16)	(285.48)	(285.67)	(285.64)	(285.80)
	L2	91.69	107.88***	52.98	81.94***	22.20	50.28***
		(295.68)	(295.57)	(279.10)	(279.03)	(279.79)	(279.70)
	Average	85.41	94.91***	46.85	69.63***	14.32	36.86***
		(267.62)	(267.67)	(253.13)	(253.24)	(253.52)	(253.59)

Table E.5: WTA, WTP, and WTA-WTP gap before and after controlling for mechanism complexity and the (MC corrected) R-range as well as the loss aversion coefficient λ. In Controlling for λ we consider only the loss aversion λ in the regression. Combined models (Model 3 and Model 4) are those defined in Table D.4, which control for both mechanism complexity and the (MC corrected) R-range. In Combined (Model 3 or 4) $+\lambda$, we include additionally λ. The sample used for this part consist of 306 and 227 subjects, above and below median performance respectively, who overlap with the study from which we obtained the loss aversion measure. Wilcoxon signed-rank tests were performed to test the significance of the difference between the WTA, WTP, and WTA-WTP gap before and after the inclusion of λ, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

F Appendix: Regressions to control WTA and WTP for mechanism complexity or/and valuation complexity

		WTA	
Explanatory variables	L1	L2	Average
($\text{WTA}_{sure} - 500$)	5.40^{***}	4.64^{***}	5.02^{***}
	(0.30)	(0.32)	(0.28)
Intercept	500.09^{***}	495.82^{***}	497.95^{***}
	(6.37)	(6.69)	(5.91)
Adjusted R^2	0.21	0.15	0.21
		WTP	
Explanatory variables	L1	L2	Average
($500 - \text{WTP}_{sure}$)	-4.94^{***}	-4.93^{***}	-4.94^{***}
	(0.28)	(0.29)	(0.26)
Intercept	431.97^{***}	424.30^{***}	428.13^{***}
	(6.50)	(6.56)	(5.94)
Adjusted R^2	0.20	0.19	0.23

Table F.1: Regression results for WTA and WTP correction for mechanism complexity using the deviation from the optimal responses for the WTA and WTP of the sure payment as a continuous variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		WTA	
Explanatory variables	L1	L2	Average
($\text{WTA}_{sure} - 500$)	6.05^{***}	4.94^{***}	5.50^{***}
	(0.44)	(0.46)	(0.41)
$\mathbf{D_{neg}}(\text{WTA}_{sure} - 500)$	-1.70^{**}	-0.80	-1.25
	(0.85)	(0.90)	(0.79)
Intercept	488.82^{***}	490.50^{***}	489.66^{***}
	(8.51)	(8.95)	(7.90)
Adjusted R^2	0.21	0.15	0.21
		WTP	
Explanatory variables	L1	L2	Average
($500 - \text{WTP}_{sure}$)	-5.27^{***}	-4.63^{***}	-4.95^{***}
	(0.46)	(0.47)	(0.50)
$\mathbf{D_{neg}}(500 - \text{WTP}_{sure})$	0.74	-0.70	0.02
	(0.84)	(0.84)	(0.76)
Intercept	438.02^{***}	418.60^{***}	428.31^{***}
	(9.41)	(9.49)	(8.60)
Adjusted R^2	0.20	0.19	0.23

Table F.2: Regression results for WTA and WTP correction for mechanism complexity using the deviation from the optimal responses for the WTA and WTP of the sure payment as a continuous variable while allowing for different slopes for positive and negative deviations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	WTA		
Explanatory variables	L1	L2	Average
R-range	0.08^{*}	0.05	0.08^{*}
	(0.04)	(0.04)	(0.04)
Intercept	519.84^{***}	514.30^{***}	514.92^{***}
	(8.06)	(8.21)	(7.84)
Adjusted R^2	0.002	0.000	0.002
	WTP		
Explanatory variables	L1	L2	Average
R-range	-0.06	-0.05	-0.08^{*}
	(0.04)	(0.04)	(0.04)
Intercept	421.55^{***}	413.37^{***}	419.62^{***}
	(8.35)	(8.43)	(8.13)
Adjusted R^2	0.001	0.001	0.002

Table F.3: Regression results for controlling WTA and WTP for the R-Range. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Upper bound		
Explanatory variables	L1	L2	Average
$(\text{Upper}_{sure} - 500)$	4.96^{***}	4.71^{***}	4.84^{***}
	(0.46)	(0.45)	(0.40)
Intercept	510.30^{***}	498.61^{***}	504.45^{***}
	(7.74)	(7.59)	(6.75)
Adjusted R^2	0.12	0.12	0.15
	Lower bound		
Explanatory variables	L1	L2	Average
$(500 - \text{Lower}_{sure})$	-4.14^{***}	-4.18^{***}	-4.16^{***}
	(0.48)	(0.47)	(0.42)
Intercept	459.92^{***}	443.78^{***}	451.85^{***}
	(7.76)	(7.58)	(6.72)
Adjusted R^2	0.08	0.09	0.11

Table F.4: Regression results for R-range correction for mechanism complexity using the deviation from the optimal responses for the randomization task of the sure payment as a continuous variable. For this regression subjects who did not choose the randomization option for the sure lottery and reported an upper bound and lower bound other than 501 and 499, respectively, are excluded. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	WTA		
Explanatory variables	L1	L2	Average
R-range after MC correction	0.07^{*}	0.03	0.07
	(0.04)	(0.04)	(0.05)
Intercept	523.27^{***}	517.64^{***}	519.13^{***}
	(7.37)	(7.54)	(7.04)
Adjusted R^2	0.002	0.000	0.001
	WTP		
Explanatory variables	L1	L2	Average
R-range after MC correction	-0.11^{**}	-0.09^{**}	-0.15^{***}
	(0.04)	(0.04)	(0.05)
Intercept	421.84^{***}	413.54^{***}	420.44^{***}
	(7.61)	(7.73)	(7.29)
Adjusted R^2	0.004	0.003	0.006

Table F.5: Regression results for controlling WTA and WTP for the R-Range after the correction for mechanism complexity using deviations from the optimal upper and lower bound of the sure payment as a continuous variable. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	WTA		
Explanatory variables	L1	L2	Average
$(\text{WTA}_{sure} - 500)$	5.39^{***}	4.64^{***}	5.00^{***}
	(0.30)	(0.32)	(0.28)
R-range (uncorrected)	0.03	0.04	0.05
	(0.04)	(0.04)	(0.04)
Intercept	496.95^{***}	491.81^{***}	493.06^{***}
	(7.31)	(7.74)	(7.10)
Adjusted R^2	0.21	0.15	0.21
	WTP		
Explanatory variables	L1	L2	Average
$(500-\text{WTP}_{sure})$	-4.94^{***}	-4.92^{***}	-4.92^{***}
	(0.28)	(0.29)	(0.26)
R-range (uncorrected)	-0.05	-0.02	-0.04
	(0.04)	(0.04)	(0.04)
Intercept	437.00^{***}	425.90^{***}	432.68^{***}
	(7.53)	(7.61)	(7.19)
Adjusted R^2	0.20	0.19	0.23

Table F.6: Regression results for WTA and WTP correction for mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment) and the uncorrected R-Range combined. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		WTA	
Explanatory variables	L1	L2	Average
$(\text{WTA}_{sure} - 500)$	6.01^{***}	4.93^{***}	5.46^{***}
	(0.44)	(0.46)	(0.41)
$\mathbf{D_{neg}}(\text{WTA}_{sure} - 500)$	-1.65^{*}	-0.78	-1.20
	(0.85)	(0.90)	(0.79)
R-range (uncorrected)	0.03	0.04	0.04
	(0.04)	(0.04)	(0.04)
Intercept	486.45^{***}	486.69^{***}	485.40^{***}
	(9.10)	(9.71)	(8.71)
Adjusted R^2	0.21	0.15	0.21
		WTP	
Explanatory variables	L1	L2	Average
$(500-\text{WTP}_{sure})$	-5.26^{***}	-4.61^{***}	-4.92^{***}
	(0.46)	(0.47)	(0.42)
$\mathbf{D_{neg}}(500-\text{WTP}_{sure})$	0.73	-0.72	-0.01
	(0.84)	(0.84)	(0.76)
R-range (uncorrected)	-0.05	-0.02	-0.04
	(0.04)	(0.04)	(0.04)
Intercept	442.88^{***}	420.21^{***}	432.60^{***}
	(10.12)	(10.13)	(9.41)
Adjusted R^2	0.20	0.19	0.23

Table F.7: Regression results for WTA and WTP correction for mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment while allowing for different slopes for positive and negative deviations) and the uncorrected R-Range combined. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		WTA	
Explanatory variables	L1	L2	Average
$\text{WTA}_{sure} - 500$	5.39***	4.64***	5.01***
	(0.30)	(0.32)	(0.28)
R-range after mechanism	0.04	0.03	0.06
complexity correction	(0.04)	(0.04)	(0.04)
Intercept	497.57***	493.75***	494.58***
	(6.73)	(7.15)	(6.43)
Adjusted R^2	0.21	0.15	0.21
		WTP	
Explanatory variables	L1	L2	Average
$(500-\text{WTP}_{sure})$	−4.92***	−4.91***	−4.90***
	(0.28)	(0.29)	(0.26)
R-range after mechanism	−0.09**	−0.04	−0.09**
complexity correction	(0.04)	(0.04)	(0.04)
Intercept	436.89***	426.52***	433.37***
	(6.88)	(6.99)	(6.46)
Adjusted R^2	0.20	0.19	0.23

Table F.8: Regression results for WTA and WTP correction for mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment) and the R-Range after mechanism complexity correction (continuous deviations from the optimal upper and lower bound of the sure payment) combined. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Explanatory variables	WTA L1	L2	Average
$(\text{WTA}_{sure} - 500)$	6.02***	4.95***	5.49***
	(0.44)	(0.46)	(0.41)
$\mathbf{D_{neg}}(\text{WTA}_{sure} - 500)$	−1.67*	−0.82	−1.24
	(0.85)	(0.90)	(0.79)
R-range after mechanism	0.04	0.03	0.06
complexity correction	(0.04)	(0.04)	(0.04)
Intercept	486.62***	488.30***	486.37***
	(8.74)	(9.33)	(8.29)
Adjusted R^2	0.21	0.15	0.21
		WTP	
Explanatory variables	L1	L2	Average
$(500-\text{WTP}_{sure})$	−5.26***	−4.60***	−4.91***
	(0.46)	(0.47)	(0.42)
$\mathbf{D_{neg}}(500-\text{WTP}_{sure})$	0.78	−0.72	0.02
	(0.83)	(0.84)	(0.76)
R-range after mechanism	−0.09**	−0.04	−0.09**
complexity correction	(0.04)	(0.04)	(0.04)
Intercept	443.26***	420.73***	433.52***
	(9.70)	(9.76)	(8.96)
Adjusted R^2	0.20	0.19	0.23

Table F.9: Regression results for WTA and WTP correction for mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment while allowing for different slopes for positive and negative deviations), and the R-range after mechanism complexity correction (continuous deviations from the optimal upper and lower bound of the sure payment) combined. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Explanatory variables	WTA L1	L2	Average
λ	−0.26	0.95	0.35
	(6.40)	(6.55)	(6.04)
Intercept	517.39***	512.21***	514.80***
	(17.71)	(18.12)	(16.74)
Adjusted R^2	0.000	0.000	0.000
		WTP	
Explanatory variables	L1	L2	Average
$1/\lambda$	50.24	49.90	50.07
	(38.70)	(39.47)	(36.67)
Intercept	444.56***	434.57***	439.56***
	(24.38)	(24.87)	(23.10)
Adjusted R^2	0.001	0.001	0.002

Table F.10: Regression results for WTA and WTP correction for loss aversion. WTA is corrected for λ and WTP for $1/\lambda$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Explanatory variables	WTA L1	L2	Average
λ	1.28 (5.71)	2.20 (6.06)	1.79 (5.42)
$\text{WTA}_{sure} - 500$	5.37*** (0.46)	4.63*** (0.49)	4.99*** (0.44)
R-range after mechanism complexity correction	0.03 (0.05)	0.01 (0.06)	0.06 (0.06)
Intercept	489.03*** (16.37)	488.99*** (17.38)	486.79*** (15.69)
Adjusted R^2	0.20	0.14	0.20
		WTP	
Explanatory variables	L1	L2	Average
$1/\lambda$	15.62 (34.70)	14.60 (35.76)	15.53 (32.51)
$(500 - \text{WTP}_{sure})$	−4.96*** (0.43)	−4.76*** (0.44)	−4.86*** (0.40)
R-range after mechanism complexity correction	−0.05 (0.05)	−0.11 (0.06)	−0.09 (0.06)
Intercept	439.10*** (22.16)	432.12*** (22.73)	436.39*** (20.83)
Adjusted R^2	0.20	0.19	0.22

Table F.11: Regression results for WTA and WTP correction for loss aversion, mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment), and the R-range after mechanism complexity correction (continuous deviations from the optimal upper and lower bound of the sure payment) combined. WTA is corrected for λ and WTP for $1/\lambda$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

		WTA	
Explanatory variables	L1	L2	Average
λ	1.18	2.12	1.70
	(5.71)	(6.07)	(5.42)
$\text{WTA}_{sure} - 500$	5.90^{***}	5.08^{***}	5.47^{***}
	(0.69)	(0.73)	(0.65)
$\mathbf{D_{neg}}(\text{WTA}_{sure} - 500)$	-1.32	-1.11	-1.22
	(1.28)	(1.35)	(1.21)
R-range after mechanism complexity correction	0.03	0.01	0.06
	(0.05)	(0.06)	(0.06)
Intercept	480.83^{***}	481.78^{***}	479.06^{***}
	(18.18)	(19.46)	(17.45)
Adjusted R^2	0.20	0.14	0.20
		WTP	
Explanatory variables	L1	L2	Average
$1/\lambda$	16.82	13.87	15.74
	(34.79)	(35.89)	(32.60)
$(500-\text{WTP}_{sure})$	-5.29^{***}	-4.55^{***}	-4.92^{***}
	(0.70)	(0.73)	(0.66)
$\mathbf{D_{neg}}(500-\text{WTP}_{sure})$	0.72	-0.44	0.13
	(1.22)	(1.26)	(1.15)
R-range after mechanism complexity correction	-0.05	-0.11^{*}	-0.09
	(0.05)	(0.06)	(0.06)
Intercept	445.41^{***}	428.17^{***}	437.50^{***}
	(24.65)	(25.40)	(23.22)
Adjusted R^2	0.20	0.19	0.22

Table F.12: Regression results for WTA and WTP correction for loss aversion, mechanism complexity (continuous deviations from the optimal WTA and WTP of the sure payment while allowing for different slopes for positive and negative deviations), and the R-range after mechanism complexity correction (continuous deviations from the optimal upper and lower bound of the sure payment) combined. WTA is corrected for λ and WTP for $1/\lambda$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Appendix: Experimental materials

U krijgt aan het begin van deze vraag een loterij. **Deze loterij is nu in uw bezit.** De loterij geeft u:

- 50% kans op 900 punten,
- en 50% kans op 100 punten.

U hebt steeds 2 mogelijkheden:

1. U houdt de loterij
2. U verkoopt de loterij voor een bepaald aantal punten.

Geef hieronder aan wat u voor ieder aantal punten kiest. Het aantal punten neemt in iedere keuze toe. De loterij blijft steeds hetzelfde. De computer vult automatisch de keuzes aan als u in een rij een keuze maakt. Controleert u daarna of u voor de overige rijen inderdaad die keuzes zou maken.

(om de loterij nogmaals te bekijken kunt u in iedere keuze met uw muis over '**loterij**' bewegen).

Optie 1		Optie 2	
◉	Houd de loterij	Verkoop de loterij voor **0 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **100 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **200 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **250 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **300 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **325 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **350 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **375 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **400 punten**	○
○	Houd de loterij	Verkoop de loterij voor **425 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **450 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **475 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **500 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **525 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **550 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **600 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **700 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **800 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **900 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **1000 punten**	◉

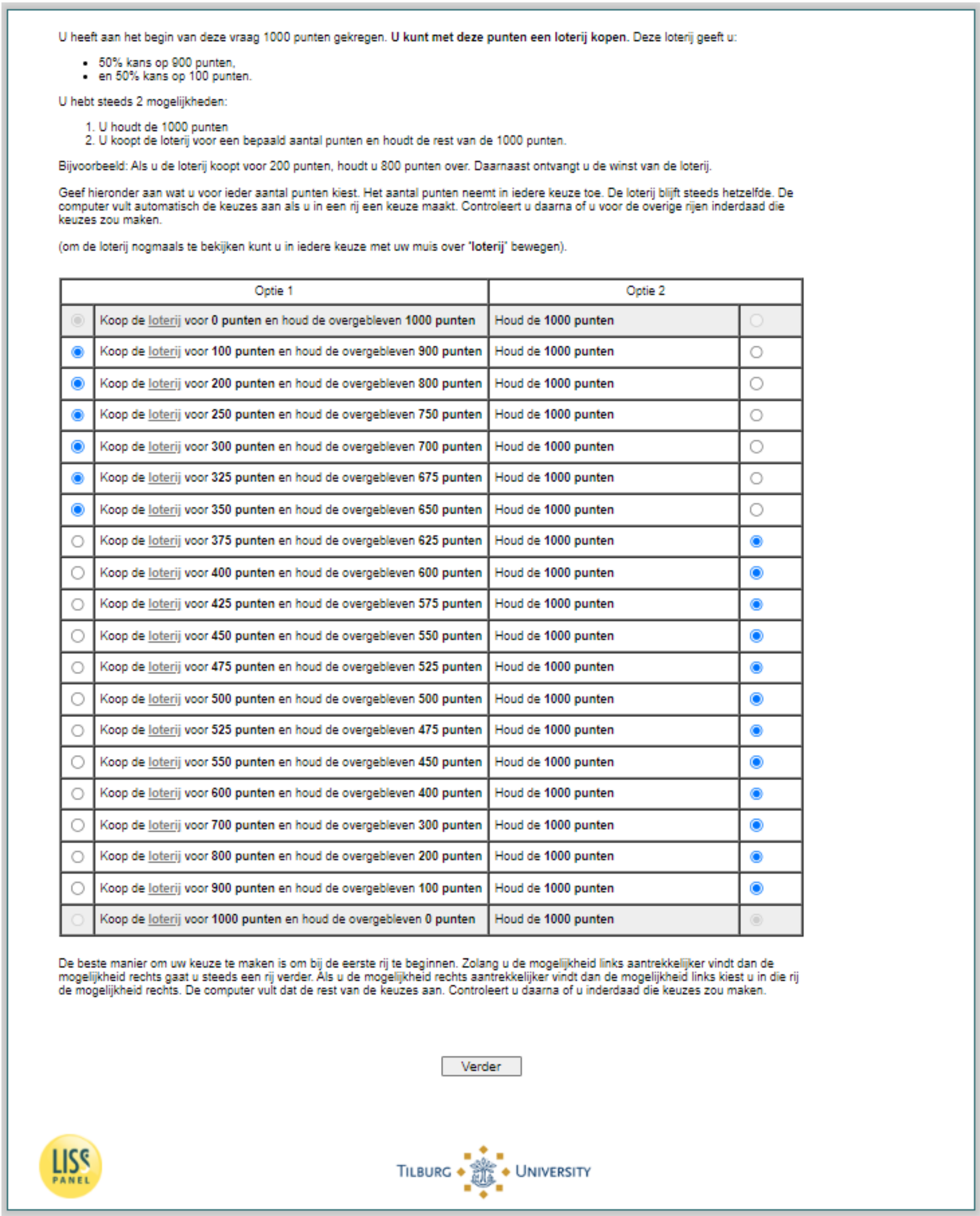
De beste manier om uw keuze te maken is om bij de eerste rij te beginnen. Zolang u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheid rechts gaat u steeds een rij verder. Als u de mogelijkheid rechts aantrekkelijker vindt dan de mogelijkheid links kiest u in die rij de mogelijkheid rechts. De computer vult dan de rest van de keuzes aan. Controleert u daarna of u inderdaad die keuzes zou maken.

Verder

LISS PANEL

Tilburg University

Figure G.1: Example of the WTA task for L1.



U heeft aan het begin van deze vraag 1000 punten gekregen. **U kunt met deze punten een loterij kopen.** Deze loterij geeft u:

- 50% kans op 900 punten,
- en 50% kans op 100 punten.

U hebt steeds 2 mogelijkheden:

1. U houdt de 1000 punten
2. U koopt de loterij voor een bepaald aantal punten en houdt de rest van de 1000 punten.

Bijvoorbeeld: Als u de loterij koopt voor 200 punten, houdt u 800 punten over. Daarnaast ontvangt u de winst van de loterij.

Geef hieronder aan wat u voor ieder aantal punten kiest. Het aantal punten neemt in iedere keuze toe. De loterij blijft steeds hetzelfde. De computer vult automatisch de keuzes aan als u in een rij een keuze maakt. Controleert u daarna of u voor de overige rijen inderdaad die keuzes zou maken.

(om de loterij nogmaals te bekijken kunt u in iedere keuze met uw muis over '**loterij**' bewegen).

Optie 1		Optie 2	
○	Koop de loterij voor **0 punten** en houd de overgebleven **1000 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **100 punten** en houd de overgebleven **900 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **200 punten** en houd de overgebleven **800 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **250 punten** en houd de overgebleven **750 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **300 punten** en houd de overgebleven **700 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **325 punten** en houd de overgebleven **675 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor **350 punten** en houd de overgebleven **650 punten**	Houd de **1000 punten**	○
○	Koop de loterij voor **375 punten** en houd de overgebleven **625 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **400 punten** en houd de overgebleven **600 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **425 punten** en houd de overgebleven **575 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **450 punten** en houd de overgebleven **550 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **475 punten** en houd de overgebleven **525 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **500 punten** en houd de overgebleven **500 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **525 punten** en houd de overgebleven **475 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **550 punten** en houd de overgebleven **450 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **600 punten** en houd de overgebleven **400 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **700 punten** en houd de overgebleven **300 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **800 punten** en houd de overgebleven **200 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **900 punten** en houd de overgebleven **100 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor **1000 punten** en houd de overgebleven **0 punten**	Houd de **1000 punten**	◉

De beste manier om uw keuze te maken is om bij de eerste rij te beginnen. Zolang u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheid rechts gaat u steeds een rij verder. Als u de mogelijkheid rechts aantrekkelijker vindt dan de mogelijkheid links kiest u in die rij de mogelijkheid rechts. De computer vult dat de rest van de keuzes aan. Controleert u daarna of u inderdaad die keuzes zou maken.

Verder

LISS PANEL

Tilburg University

Figure G.2: Example of the WTP task for L1.

U krijgt aan het begin van deze vraag een loterij. **Deze loterij is nu in uw bezit.** De loterij geeft u:

- 100% kans op 500 punten.

U hebt steeds 2 mogelijkheden:

1. U houdt de loterij
2. U verkoopt de loterij voor een bepaald aantal punten.

Geef hieronder aan wat u voor ieder aantal punten kiest. Het aantal punten neemt in iedere keuze toe. De loterij blijft steeds hetzelfde. De computer vult automatisch de keuzes aan als u in een rij een keuze maakt. Controleert u daarna of u voor de overige rijen inderdaad die keuzes zou maken.

(om de loterij nogmaals te bekijken kunt u in iedere keuze met uw muis over **'loterij'** bewegen).

Optie 1		Optie 2	
○	Houd de loterij	Verkoop de loterij voor **450 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **475 punten**	○
◉	Houd de loterij	Verkoop de loterij voor **499 punten**	○
○	Houd de loterij	Verkoop de loterij voor **501 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **525 punten**	◉
○	Houd de loterij	Verkoop de loterij voor **550 punten**	○

De beste manier om uw keuze te maken is om bij de eerste rij te beginnen. Zolang u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheid rechts gaat u steeds een rij verder. Als u de mogelijkheid rechts aantrekkelijker vindt dan de mogelijkheid links kiest u in die rij de mogelijkheid rechts. De computer vult dan de rest van de keuzes aan. Controleert u daarna of u inderdaad die keuzes zou maken.

Verder

LISS PANEL

TILBURG UNIVERSITY

U heeft aan het begin van deze vraag 1000 punten gekregen. **U kunt met deze punten een loterij kopen.** Deze loterij geeft u:

- 100% kans op 500 punten.

U hebt steeds 2 mogelijkheden:

1. U houdt de 1000 punten
2. U koopt de loterij voor een bepaald aantal punten en houdt de rest van de 1000 punten.

Bijvoorbeeld: Als u de loterij koopt voor 200 punten, houdt u 800 punten over. Daarnaast ontvangt u de winst van de loterij.

Geef hieronder aan wat u voor ieder aantal punten kiest. Het aantal punten neemt in iedere keuze toe. De loterij blijft steeds hetzelfde. De computer vult automatisch de keuzes aan als u in een rij een keuze maakt. Controleert u daarna of u voor de overige rijen inderdaad die keuzes zou maken.

(om de loterij nogmaals te bekijken kunt u in iedere keuze met uw muis over **'loterij'** bewegen).

Optie 1		Optie 2	
○	Koop de loterij voor 450 punten en houd de overgebleven **550 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor 475 punten en houd de overgebleven **525 punten**	Houd de **1000 punten**	○
◉	Koop de loterij voor 499 punten en houd de overgebleven **501 punten**	Houd de **1000 punten**	○
○	Koop de loterij voor 501 punten en houd de overgebleven **499 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor 525 punten en houd de overgebleven **475 punten**	Houd de **1000 punten**	◉
○	Koop de loterij voor 550 punten en houd de overgebleven **450 punten**	Houd de **1000 punten**	○

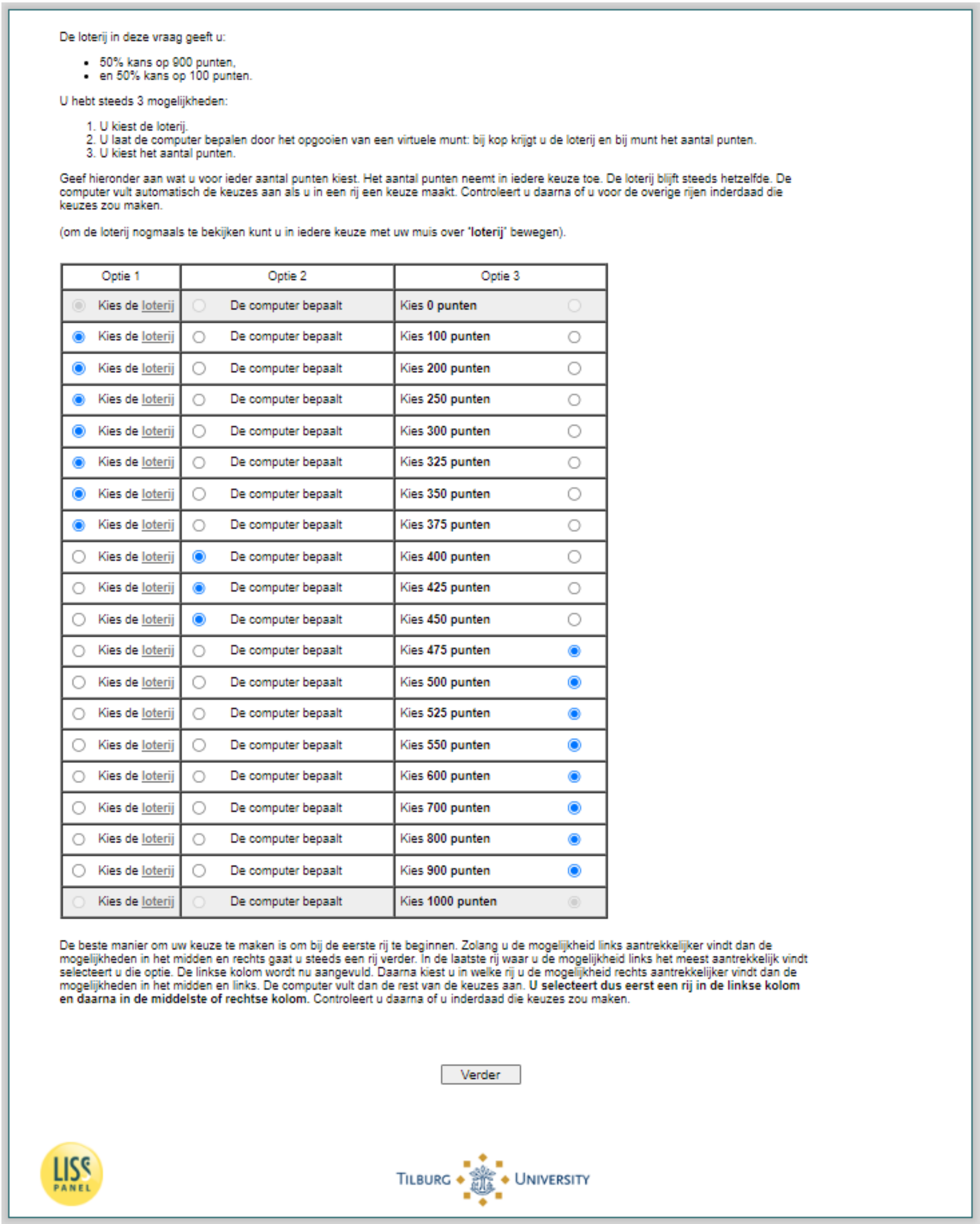
De beste manier om uw keuze te maken is om bij de eerste rij te beginnen. Zolang u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheid rechts gaat u steeds een rij verder. Als u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheid links kiest u in die rij de mogelijkheid rechts. De computer vult dat de rest van de keuzes aan. Controleert u daarna of u inderdaad die keuzes zou maken.

Verder

LISS PANEL

TILBURG UNIVERSITY

Figure G.3: Example of the WTA and WTP task for the sure payment.


De loterij in deze vraag geeft u:

- 50% kans op 900 punten,
- en 50% kans op 100 punten.

U hebt steeds 3 mogelijkheden:

1. U kiest de loterij.
2. U laat de computer bepalen door het opgooien van een virtuele munt: bij kop krijgt u de loterij en bij munt het aantal punten.
3. U kiest het aantal punten.

Geef hieronder aan wat u voor ieder aantal punten kiest. Het aantal punten neemt in iedere keuze toe. De loterij blijft steeds hetzelfde. De computer vult automatisch de keuzes aan als u in een rij een keuze maakt. Controleert u daarna of u voor de overige rijen inderdaad die keuzes zou maken.

(om de loterij nogmaals te bekijken kunt u in iedere keuze met uw muis over '**loterij**' bewegen).

Optie 1	Optie 2	Optie 3
Kies de loterij	De computer bepaalt	Kies **0 punten**
Kies de loterij	De computer bepaalt	Kies **100 punten**
Kies de loterij	De computer bepaalt	Kies **200 punten**
Kies de loterij	De computer bepaalt	Kies **250 punten**
Kies de loterij	De computer bepaalt	Kies **300 punten**
Kies de loterij	De computer bepaalt	Kies **325 punten**
Kies de loterij	De computer bepaalt	Kies **350 punten**
Kies de loterij	De computer bepaalt	Kies **375 punten**
Kies de loterij	De computer bepaalt	Kies **400 punten**
Kies de loterij	De computer bepaalt	Kies **425 punten**
Kies de loterij	De computer bepaalt	Kies **450 punten**
Kies de loterij	De computer bepaalt	Kies **475 punten**
Kies de loterij	De computer bepaalt	Kies **500 punten**
Kies de loterij	De computer bepaalt	Kies **525 punten**
Kies de loterij	De computer bepaalt	Kies **550 punten**
Kies de loterij	De computer bepaalt	Kies **600 punten**
Kies de loterij	De computer bepaalt	Kies **700 punten**
Kies de loterij	De computer bepaalt	Kies **800 punten**
Kies de loterij	De computer bepaalt	Kies **900 punten**
Kies de loterij	De computer bepaalt	Kies **1000 punten**

De beste manier om uw keuze te maken is om bij de eerste rij te beginnen. Zolang u de mogelijkheid links aantrekkelijker vindt dan de mogelijkheden in het midden en rechts gaat u steeds een rij verder. In de laatste rij waar u de mogelijkheid links het meest aantrekkelijk vindt selecteert u die optie. De linkse kolom wordt nu aangevuld. Daarna kiest u in welke rij u de mogelijkheid rechts aantrekkelijker vindt dan de mogelijkheden in het midden en links. De computer vult dan de rest van de keuzes aan. **U selecteert dus eerst een rij in de linkse kolom en daarna in de middelste of rechtse kolom.** Controleert u daarna of u inderdaad die keuzes zou maken.

Verder

Figure G.4: Example of the R-range task for L1.

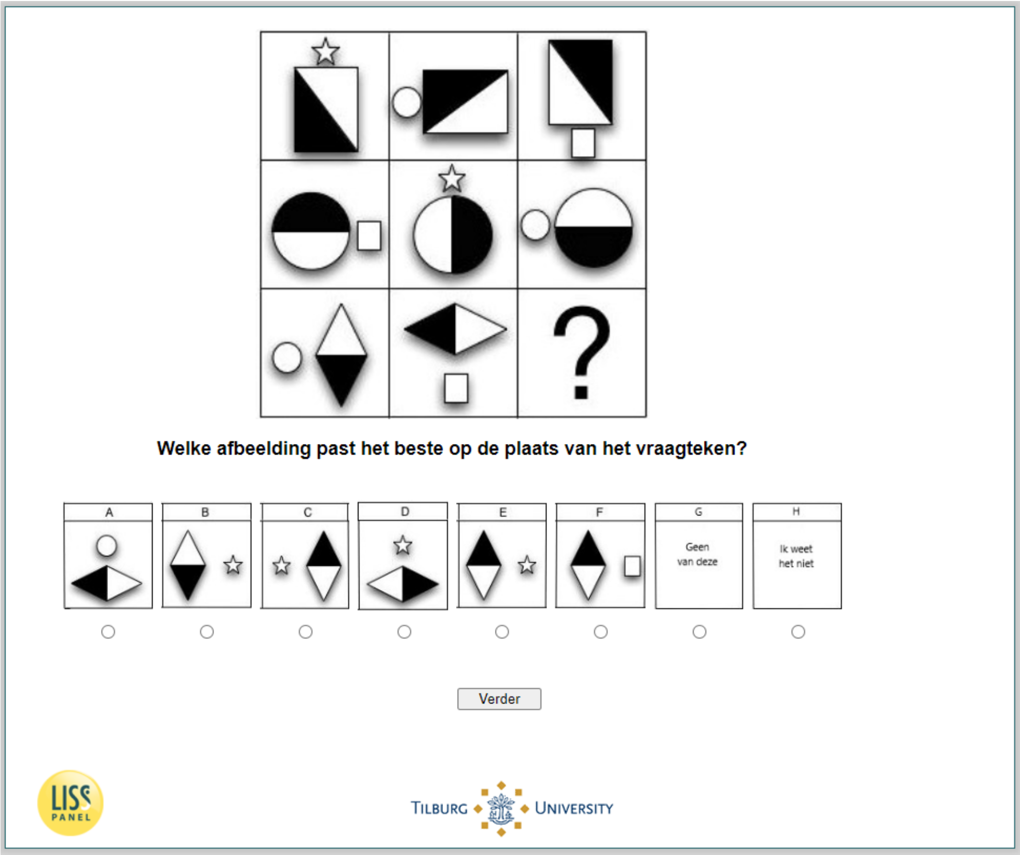
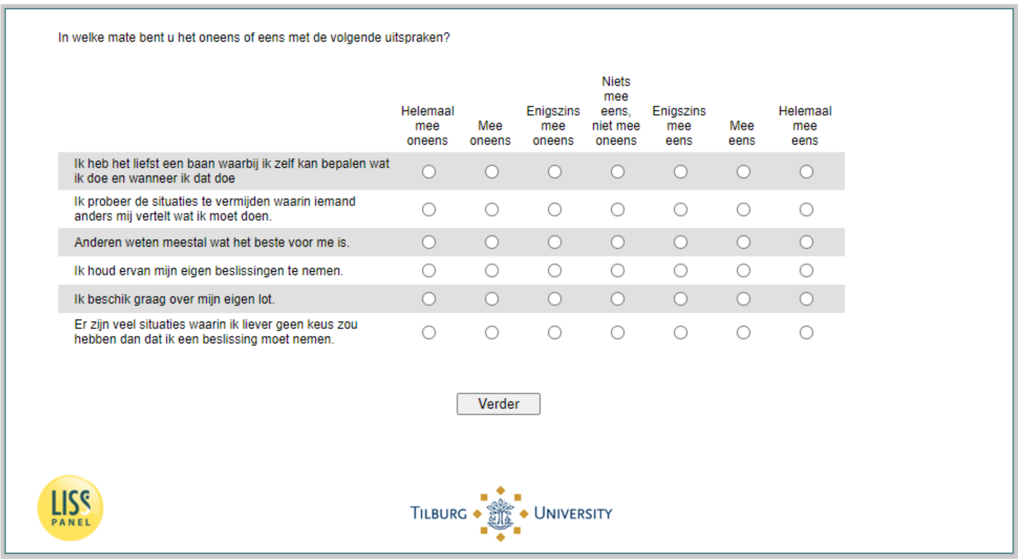


Figure G.5: Example of the matrix reasoning questions.



In welke mate bent u het oneens of eens met de volgende uitspraken?

	Helemaal mee oneens	Mee oneens	Enigszins mee oneens	Niets mee eens, niet mee oneens	Enigszins mee eens	Mee eens	Helemaal mee eens
Ik heb het liefst een baan waarbij ik zelf kan bepalen wat ik doe en wanneer ik dat doe							
Ik probeer de situaties te vermijden waarin iemand anders mij vertelt wat ik moet doen.							
Anderen weten meestal wat het beste voor me is.							
Ik houd ervan mijn eigen beslissingen te nemen.							
Ik beschik graag over mijn eigen lot.							
Er zijn veel situaties waarin ik liever geen keus zou hebben dan dat ik een beslissing moet nemen.							

Verder

LISS PANEL

TILBURG UNIVERSITY

Figure G.6: Desirability of Control statements.