Measuring subjective decision confidence

Arts, Sara and Ong, Qiyan and Qiu, Jianying

Radboud University Nijmegen, National University of Singapore, Radboud University Nijmegen

August 2020

Measuring decision confidence

Sara Arts* Qiyan Ong† Jianying Qiu‡

May 21, 2024

Abstract

We examine whether the way individuals randomize between options captures their decision confidence. In two experiments in which subjects faced pairs of options (a lottery and a varying sure payment), we allowed subjects to choose randomization probabilities according to which they would receive each option. Separately, we obtained two measures of self-reported confidence - confidence statements and probabilistic confidence - for choosing between the two options. Consistent with the predictions of two theoretical frameworks incorporating preference uncertainty, the randomization probabilities correlated strongly with both self-reported measures (median Spearman correlations between 0.86 to 0.89) and corresponded in absolute levels to probabilistic confidence. This relationship is robust to two exogenous manipulations of decision confidence, where we varied the complexity of the lottery and subjects' experience with the lottery.

Keywords: decision confidence, randomization, incentivized approach, preference uncertainty

JEL Classification B40, C91, D81

---

*Department of Economics, IMR, Radboud University, Heyendaalseweg 141, 6525 AJ Nijmegen, the Netherlands Email: sara.arts@ru.nl.

†Department of Economics, IMR, Radboud University, Heyendaalseweg 141, 6525 AJ Nijmegen, the Netherlands Email: qiyan.ong@ru.nl.

‡Corresponding author. Jianying Qiu, email: jianying.qiu@ru.nl, Department of Economics, IMR, Radboud University, EOS 02.577, Heyendaalseweg 141, 6525 AJ Nijmegen, the Netherlands.

1 Introduction

There are many decisions in life that people may not be able to make with full confidence. These decisions often involve difficult trade-offs among conflicting objectives, such as price vs. quality when buying goods, risk vs. return when investing, and efficiency vs. equality when making policy decisions. As more studies suggest that decision confidence has the potential to explain a wide range of behavioral anomalies, there is growing interest in eliciting and accounting for decision confidence when studying people's choices.[1]

Past studies so far have mostly relied on non-incentivized self-reports to elicit decision confidence. For example, some studies (Cohen et al., 1987; Dubourg et al., 1994, 1997; Cubitt et al., 2015) allowed subjects to indicate whether they were unsure of their choices. Butler and Loomes (2007, 2011) had subjects indicate their decision confidence using the ordinal terms "definitely" and "probably." More quantitatively, Enke and Graeber (2021, 2023) had subjects rate how certain (from 0% to 100% in increments of 5%) they were that their actual valuation of an option was within the interval obtained from the choices they made earlier in a price list.

While asking people explicitly about how confident they are about their decisions directly elicits decision confidence, incentivized behavioral measures that elicit decision confidence without referring to confidence may encourage people to contemplate their decisions carefully and thereby reduce behavioral biases such as priming effects and experimenter demand effects (Camerer and Hogarth, 1999). Yet, finding behavioral measures of decision confidence "in a form simple and transparent enough to work without creating additional uncertainty" is not trivial (Butler and Loomes, 2011, p. 516).

Building on earlier studies, we propose to use the randomization probability assigned to an option in a choice pair as a behavioral measure of decision confidence, and test its

[1] These anomalies include the willingness to accept (WTA) - willingness to pay (WTP) gap (Dubourg et al., 1994), preference reversals (Butler and Loomes, 2007), stochastic choices (Agranov and Ortoleva, 2017), insensitivity to variation in probabilities (Enke and Graeber, 2023), anomalies in intertemporal choices (Enke and Graeber, 2021), small-stakes risk aversion (Khaw et al., 2021), and many other violations of standard decision theory (Butler and Loomes, 2011).

validity in two experiments. For each pair of options (a lottery x and a sure payment y), we obtained subjects' binary choice, self-reported decision confidence, and randomization probabilities. Self-reported decision confidence was elicited by having subjects select a confidence statement from "Surely x," "Probably x," "Unsure," "Probably y," "Surely y," after they chose an option and by having them report how confident they were in choosing either options as a probabilistic confidence of p% x and 100-p% y, where p ranged from 0 to 40 and 60 to 100 in increments of 10. Separately, subjects chose a randomization probability $0 \leq \lambda \leq 1$ with which they would receive x (and with probability $1 - \lambda$ receive y) for each pair of options (Miao and Zhong, 2018; Agranov and Ortoleva, 2023; Feldman and Rehbeck, 2022; Ong and Qiu, 2023). Unlike the two self-reported confidence measures, the elicitation of randomization probabilities made no reference to decision confidence.

We exogenously manipulated decision confidence by a) having a simple lottery with two outcomes and a complex lottery with the same expected value but with more payoff outcomes over a wider range of possible values (Fudenberg and Puri, 2021), and by b) increasing subjects' experience with the lottery by allowing them to either observe the outcome draws of the lottery or to make hypothetical choices and observe the payoffs of their choices and the counterfactual (Myagkov and Plott, 1997; Plott and Zeiler, 2005; van de Kuilen and Wakker, 2006; van de Kuilen, 2009).

To ensure that experimenter demand effects and order effects were not driving our results, in one of our experiments, we elicited the three measures separately over three sessions (at least seven days apart). Subjects were randomly assigned to one of the three decision orders, which differed by whether the confidence statements, the probabilistic confidence, or the randomization probabilities were elicited in the first session.

We structure the analyses of our experiments through two theoretical frameworks, one based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015) and the other based on Fudenberg et al. (2015). We illustrate how randomization emerges from the optimization behavior of an individual who faces uncertainty regarding her preference between the two options. Based on our expectations that randomization probabilities serve as a good proxy for decision confidence, and that complexity increases preference uncertainty while experience decreases it, we have four hypotheses. First, randomization probabilities correlate positively with both self-reported measures; second, subjects choose randomization probabilities around 0.5 for choice pairs in which the sure payment has a similar decision utility as the lottery (and one does not dominate the other); third, subjects randomize over a wider range of sure payments and choose randomization probabilities closer to 0.5 for the complex lottery than for the simple lottery; finally, subjects randomize over a smaller range of sure payments and choose randomization probabilities further away from 0.5 in the experience treatments.

Our experimental results support all four hypotheses. Subjects' randomization probabilities were strongly and positively correlated with both confidence statements and probabilistic confidence (median Spearman correlation between 0.86 to 0.89). In line with our expectations, the two exogenous manipulations affected self-reported decision confidence. Increasing the complexity of the lottery led to a decrease in self-reported decision confidence, while increasing experience with lotteries led to an increase in self-reported decision confidence. These exogenous changes in self-reported decision confidence were met with corresponding changes in randomization probabilities: subjects randomized over a larger range of sure payments and the randomization probabilities were closer to 0.5 for the complex lottery than for the simple lottery, while the opposite occurred when subjects had more experience. As a result, the correlations between randomization probabilities and self-reported decision confidence measures were robust to the exogenous manipulations of decision confidence.

Our study builds on the growing literature on preferences for randomization, implying preference functionals that are convex with respect to probabilistic mixing, which is a violation of the betweenness axiom (Chew, 1983; Dekel, 1986; Chew, 1989). Preferences for randomization have been documented over wide ranges, across different domains, in experimental settings as well as in real life decisions. In a multiple-decision setting, Rubinstein (2002) suggested that randomization (diversification in his term) by choosing differently across five independent and identical decisions is "an expression of a more general phenomenon in which people tend to diversify their choices when they face a sequence of similar decision problems and are uncertain about the right action" (Rubinstein, 2002, p.1370). Dwenger et al. (2018) found that their experimental subjects preferred to randomize via an external randomization device rather than making choices themselves, and the authors reported similar behavior among German university applicants. Miao and Zhong (2018) showed that randomization could be used to balance ex-ante and ex-post social preferences. Feldman and Rehbeck (2022) elicited individuals' attitudes toward reduced mixtures over two lotteries in the space of three-outcome lotteries (the Marschak-Machina triangle) and found pervasive evidence of a preference for non-degenerate mixing over lotteries. The studies closest to ours are Agranov and Ortoleva (2023) and Ong and Qiu (2023), who also allowed subjects to choose randomization probabilities when deciding between two options. Both studies found that subjects often randomized and did so over large ranges. Ong and Qiu (2023) further found that subjects were willing to pay to randomize, suggesting that randomization was deliberate and not merely a result of indifference. Popular explanations for convex preferences include hedging in the face of preference uncertainty (Cerreia-Vioglio et al., 2015; Fudenberg et al., 2015; Cerreia-Vioglio et al., 2019), non-linear probability weighting (Kahneman and Tversky, 1979; Quiggin, 1982; Tversky and Kahneman, 1992), and responsibility aversion (Dwenger et al., 2014). Our study is the first to provide experimental evidence linking preference uncertainty and randomization behavior.[2]

We also contribute to the literature on stochastic choice, which examines why individuals change their decisions when they face the same decision situation repeatedly. The relationship we found between randomization probabilities and sure payments bears a remarkable resemblance to results reported in studies on stochastic choice, for example, Mosteller and Nogee (1951, Figure 2) and Loomes and Pogrebna (2017, Table 1).[3] The similarity between the choice proportion in repeated choices and the randomization probability in a one-shot decision suggests that decision confidence may have the potential to explain stochastic choices. Consistent with this interpretation, we find that, across subjects and

[2]Agranov and Ortoleva (2023) also explicitly discuss this link. Based on reports from the end-of-experiment questionnaire, they found that many of their subjects randomized because they were unsure of their preferences (Agranov and Ortoleva, 2023, Appendix A.8).

[3]Note that these results come from entirely different designs. In Mosteller and Nogee (1951) and Loomes and Pogrebna (2017) individuals repeatedly faced a lottery and a sure payment, with the sure payment varying from one question to another, and the results are about the proportion of accepting the lottery across decisions, whereas in our experiment subjects faced the lottery and a sure payment once and chose the randomization probability of receiving the lottery.

decisions, higher decision confidence in an option corresponded to choosing that option more frequently (but not always) in binary choices. Meanwhile, random (expected) utility models (see, e.g., Eliashberg and Hauser, 1985; Loomes and Sugden, 1995; Gul and Pesendorfer, 2006; Apesteguia and Ballester, 2018), which are the standard explanations for stochastic choices, do not predict randomization in a one-shot decision as observed in our experiment. This is because, while individuals may be considered to have a set of utility functions in this literature, at the moment of decision-making, they rely on one utility function randomly realized from the set.

Overall, our study provides direct evidence to the connections between some important concepts in the literature, such as decision confidence, cognitive uncertainty, preference uncertainty/imprecision, incomplete preference, preference for randomization, and stochastic choice. While there have been notable theoretical advancements and accumulating empirical evidence in this field, the precise interplay and relationships between these concepts as well as how they are connected to choices remain ambiguous. Our finding of a systematic relationship between the randomization, alternative measures of decision confidence, and stochastic choice suggests that there may exist a common psychological underpinning for these various concepts.

The rest of the paper proceeds as follows. Section 2 describes the experimental procedure. Section 3 provides the theoretical basis for how randomization probabilities may be linked to decision confidence. The results are reported in Section 4. Finally, Section 5 concludes the paper.

2 Experimental design

We had two experiments. We first describe the general structure of the experiments before detailing the differences.

2.1 General structure of the experiments

In each decision, subjects faced a pair of options: a lottery x and a sure payment y. The lottery was paired with 13 values of sure payments (0, 2, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, and 10 euros). For each type of decisions that we will describe below, subjects faced these pairs in a random sequence. Each decision was made on a separate screen, and subjects were not allowed to review or change their decisions once they were made. Each subject made three types of decisions: binary choices, self-reported decision confidence, and randomized choices.

Binary choices

The binary choices required subjects to choose either x or y. If x was chosen, the computer would draw a random number to determine x's outcome. For example, for x that has a 50% chance of paying 9 euro and a 50% chance of paying 1 euro, if the randomly drawn number falls between 1 and 50, the subject would receive 1 euro, and if the randomly drawn number falls between 51 and 100, the subject would receive 9 euros.

Two measures of self-reported decision confidence

After making the binary choices, we asked subjects how confident they felt about their choices. The confidence statements they could choose were "Surely x," "Probably x," "Unsure," "Probably y," or "Surely y." Similar statements were used in Dubourg et al. (1994), Butler and Loomes (2007), and Butler and Loomes (2011). Confidence statements were not incentivized and could not affect payoffs.

In addition to the confidence statements, subjects in Experiment 2 also had to report their probabilistic confidence in a separate experimental decision. Instead of making a direct binary choice, subjects had to choose how confident they felt about choosing x versus y. They had to choose between ten levels of probabilistic confidence: "100% x, 0% y," "90% x, 10% y," ... "60% x, 40% y," "40% x, 60% y," ..., "0% x, 100% y." Subjects were considered to have chosen the option for which they indicated more than 50% probabilistic confidence. For example, if a subject chose "60% x, 40% y," she was considered to have chosen x over y in that decision. To use the probabilistic confidence as a measure of decision confidence

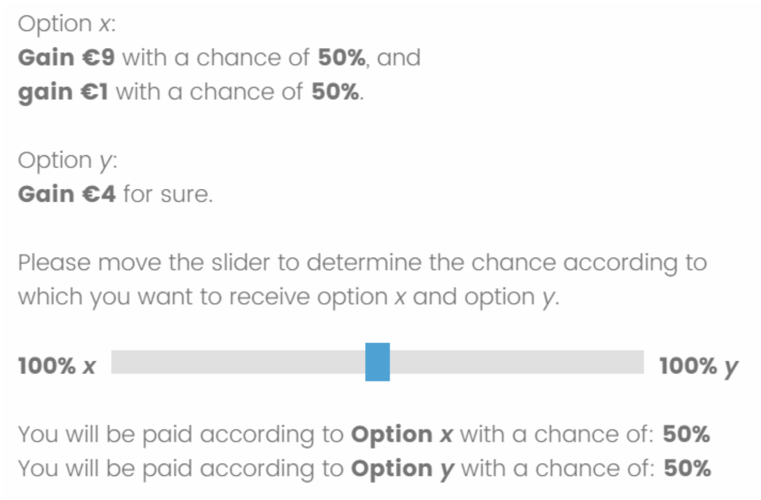


Figure 1: An example of the randomized choice decision screen, where option x is a lottery with a 50% chance of gaining 9 euro and a 50% chance of gaining 1 euro. Option y is a sure payment and varies across choices.

as well as an indicator of a subject's choice between x and y, we omitted the option "50% x, 50% y".

Randomized choices

The randomized choices required subjects to choose a randomization probability λ, based on which they would receive x (and hence with a probability $1 - \lambda$ of receiving y). For example, choosing $\lambda = 0.40$ means the subject would receive x with a chance of 40% and y with a chance of 60%. The subjects used a slider from 0% to 100% with increments of 1% to choose the randomization probability in each choice. In experiment 1, the slider was set in the middle at the start. To reduce anchoring bias, in experiment 2, the slider had no initial position, and subjects needed to click on the slider and move the bar to determine the randomization probability. If the randomized choice was chosen for payment, the computer would draw a random number between 1 and 100. If the drawn number was between 1 and 100λ, x would be chosen over y in that decision. If x was chosen for payment, a second random draw determined the outcome of the lottery. Figure 1 shows the decision screen for the randomized choice. To ensure that subjects understand the payoff structure of randomized choices, we provided two examples as well as reminders in the lower part of the decision screens on how randomized choices affect payment (see Figure C.4 and C.5 in Appendix C.3).

2.2 Manipulating decision confidence

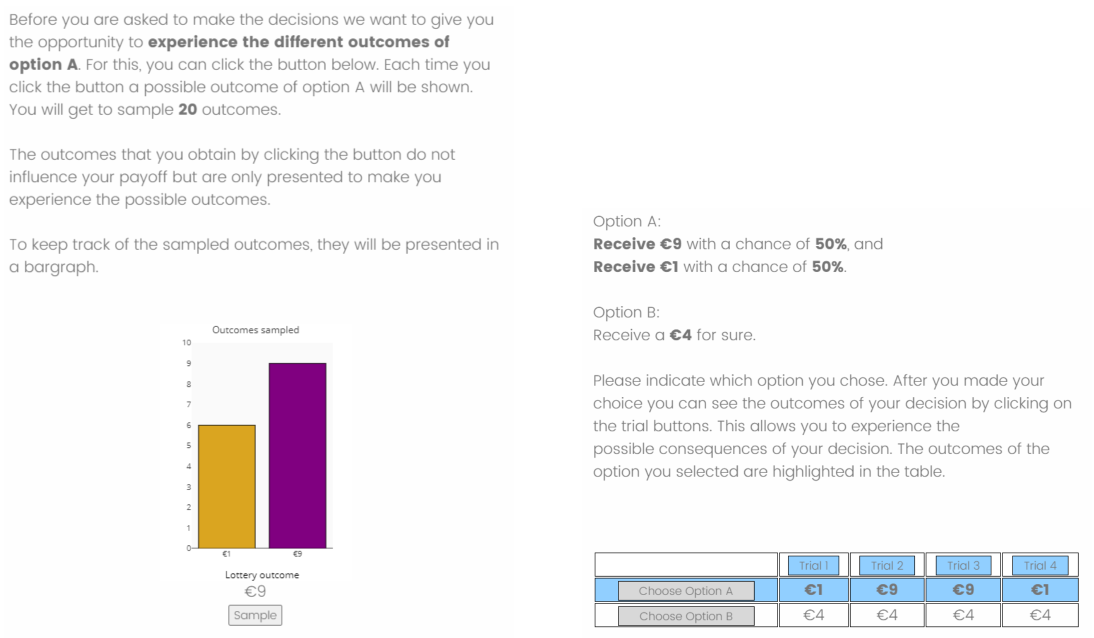
In the baseline treatment, the subjects faced a simple lottery with two outcomes (a 50% chance of 9 euro and a 50% chance of 1 euro). They received a complete description of the lottery before making their decisions.

Varying the complexity of the lottery: We manipulated, within subject, the complexity of the lottery by asking all subjects to make decisions involving a complex lottery with four outcomes in addition to decisions involving the simple lottery. The complex lottery has the same expected value as the simple lottery. It offers 9.75 euros with a chance of 20%, 7.50 euros with a chance of 30%, 2.50 euros with a chance of 30%, and 0.25 euros with a chance of 20%. The order of the lotteries was randomized at individual subject level: some subjects proceeded from the simple lottery to the complex lottery, while others completed the decisions in the reverse order.[4]

Varying subjects' experience with the lottery: For this manipulation, we adopted a between-subject design, where subjects were randomly assigned to either the baseline treatment or the experience treatment. In Experiment 1, after learning the probability distribution of the lottery and prior to making actual decisions, subjects in the partial-experience treatment had to click and view 20 draws of the lottery. As the subjects viewed each lottery draw, an accompanying bar chart which recorded each lottery outcome was updated. By the 20th draw, the bar chart reflected the probability distribution of the lottery. Figure 2(a) shows an example of the partial-experience treatment.

In Experiment 2, subjects in the full-experience treatment had to make five hypothetical decisions per lottery, with each decision involving a different sure payment (3, 4, 5, 6 or 7 euros in a random sequence). For each decision, they observed four potential realizations of the payoff of their choice as well as the counterfactual in a payoff table. Figure 2(b)

[4]The subjects in Experiment 1 also made decisions involving a loss lottery and a mixed lottery. We included these lotteries because they might lead to larger preference uncertainty due to the additional uncertainty in attitudes toward losses. We omitted these two lotteries in Experiment 2 because, as pointed out by one reviewer, the theoretical analysis of these two lotteries requires a more general approach than what we had relied on. The main results involving these lotteries are provided in online Appendix C.2. Further details can be found in the working paper version (Arts et al., 2020).



(a) The partial-experience treatment

(b) The full-experience treatment

Figure 2: Panel (a) illustrates what subjects in the partial-experience treatment saw when they generated the outcomes of the lottery. Panel (b) illustrates what subjects saw in the full-experience treatment. The numbers highlighted in blue in the table show a subject's hypothetical decision and her four potential payoffs, and the not highlighted numbers show the counterfactuals.

shows the decision screen and the payoff table viewed by a subject who chose the lottery over the sure payment of 4 euros in the hypothetical decision.

2.3 Design considerations

A few design features deserve some discussion. First, our within-subject design of eliciting self-reported confidence measures and randomization probabilities for each subject raises the concern that experimenter demand effects or order effects may unintentionally influence subjects to give similar responses, resulting in a systematic relationship between them. We took several measures in Experiment 2 to make it more obscure and costly for subjects to connect self-reported confidence measures and randomization probabilities (Zizzo, 2010), such as spreading the decisions over three sessions (at least seven days apart), including a cost for randomizing, and randomly assigning subjects to one of the three different orders.

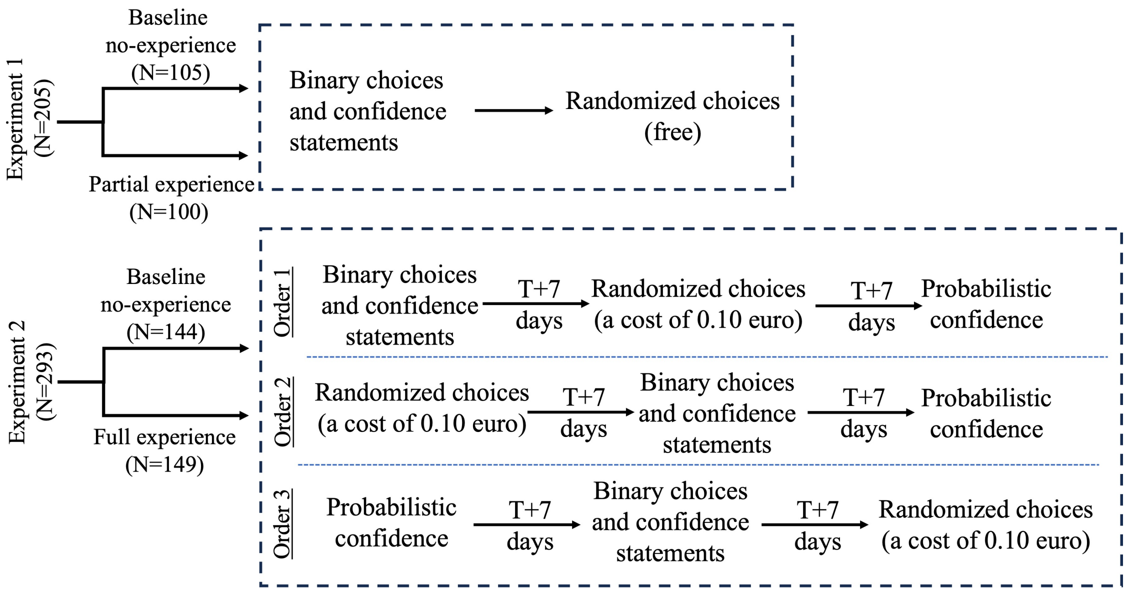


Figure 3: Summary of the treatments and experimental procedure in Experiment 1 and 2. The dotted rectangle highlights the types of decisions that subjects in each experiment made for both lotteries. The sequence of the simple lottery and the complex lottery in each type of decision was randomized at individual subject level.

The key features of the two experiments are summarized in Figure 3.

Second, varying the complexity of the lottery could induce changes in behavior through channels other than decision confidence. For example, subjects may valuate the two lotteries differently, as found in studies documenting complexity seeking and complexity averse behaviors (see e.g. Abdellaoui et al., 2020, and the references therein). This difference is less relevant for our purpose because our focus is on decision confidence (e.g., the range of sure payments that subjects do not have full confidence) rather than the average valuation of the lotteries. Another concern is that varying the complexity of the lottery induces different randomization behavior. While this may occur in some non-EUT models (Machina, 1985; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Quiggin, 1982), they do not predict a close relationship between randomization probability and self-reported decision confidence, or how this relationship changes with the complexity of the lottery. We will return to this point in subsection 4.4 to discuss other interpretations of randomization probabilities that could predict different randomization behavior between the two lotteries.

A final concern is the choice of the experience treatments. The description-experience gap literature suggests that description and experience may induce different risk preferences (see e.g., Hertwig et al., 2004; Wulff et al., 2018), which could reduce decision confidence. Our partial experience design is unlikely to have this issue because subjects' experience of the lottery realizations were equivalent to its description. Recent studies by Aydogan and Gao (2020) and Cubitt et al. (2022) show that the description-experience gap should be small in this case. Examining more closely, our experience treatments differ from the standard design in the Description-Experience gap literature as we provided subjects with the full description of the lottery's probability distribution in addition to the opportunity to experience the realizations of the lottery. In this sense, our experience treatments are closer to studies showing that experience in addition to a full description of the lottery could help subjects develop a better understanding about their preference (e.g., van de Kuilen and Wakker, 2006; van de Kuilen, 2009).

2.4 Sample and procedure

The data were collected from a sample of 498 subjects of the ID lab at Radboud University. A total of 205 subjects participated in Experiment 1 and 293 in Experiment 2. Invitations were sent in batches via ORSEE (Greiner, 2015). The experiment was conducted using Qualtrics and lasted approximately 20 minutes for Experiment 1 and about 30 minutes in total for Experiment 2. In Experiment 2, subjects were also asked to answer a post-experiment questionnaire at the end of each of the three sessions based on the type of decision confidence they reported in that session. Online Appendix C.3 contains the experimental instructions and decision screens. Each subject received a participation fee of 1 euro and an additional payment based on one of the decisions they made in the experiment. In Experiment 1, the additional monetary compensation was based on a decision randomly selected from their binary choices or randomized choices. In Experiment 2, it was based on a decision randomly selected from their binary choices, randomized choices, or probabilistic confidence decisions. The average additional payment was 6.28 euros. We made the payment via bank transfers.

3 Theoretical analysis

Under expected utility theory (EUT) which ignores decision confidence and assumes that a unique utility function (subject to positive affine transformation) captures an individual's preference, it is straightforward to show that the individual chooses $\lambda^* \in (0,1)$ for at most one value of the sure payment in the 13 choice pairs. Thus, under EUT, strict randomization ($\lambda^* \in (0,1)$) rarely occurs, and randomization probabilities do not contain additional information beyond indifference.

We present two theoretical analyses of our experiments that provide an explicit link between randomization probabilities and decision confidence. Both analyses assume that the individual is uncertain about her preference. Appendix A.1 presents a theoretical framework based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015). In this framework, the individual has multiple utility functions that we call multiple selves. She is not fully confident about her choice when some selves prefer one option while others prefer the other. In such instances, the individual prefers randomization over selecting a particular option because it offers a "fair" way to resolve internal conflicts among her different selves. This approach of capturing the lack of decision confidence from unsureness about preferences is closely related but is different from models of ambiguity (e.g., Gilboa and Schmeidler, 1989; Klibanoff et al., 2005), which focus on unsureness about beliefs (e.g., Halevy, 2007; Chew et al., 2017; Cubitt et al., 2020, and the references therein).

Appendix A.2 presents the extension of Fudenberg et al.'s (2015) model. Fudenberg et al. (2015) axiomatized a choice rule of deliberate randomization called additive perturbed utility (APU). Their representation corresponds to a form of ambiguity-averse preferences for an individual who is uncertain about her true utility function. The individual randomizes to balance the probability of errors due to preference uncertainty against the cost of avoiding them (Fudenberg et al., 2015, p. 2373).

Both analyses suggest that the preference for randomization is motivated by the hedging of preference uncertainty. In particular, randomization probabilities are affected by the perceived preference uncertainty of the options, attitudes towards preference uncertainty, as well as the utility difference between the options. The two theoretical analyses suggest that our proposed link between randomization probabilities and decision confidence could hold under a broad class of decision models that incorporate preference uncertainty.

We expect that subjects perceive more preference uncertainty with the complex lottery than with the simple lottery, and experience with the lottery reduces preference uncertainty regarding the lottery. With these expectations, our theoretical analyses show that randomization probabilities share three important properties of decision confidence: a) subjects choose randomization probabilities close to 0.5 when the sure payment has a similar decision utility to the lottery; b) they randomize over a wider range of sure payments, with randomization probabilities closer to 0.5, when they face the complex lottery compared to the simple lottery; and c) with experience and less preference uncertainty about the lottery, subjects' randomization probabilities may be stretched away from 0.5 as they randomize over a smaller range of sure payments. If randomization probabilities and the two self-report measures both capture decision confidence, we expect the following:

Hypotheses.

1. *Randomization probabilities are positively correlated with the self-reported confidence measures.*

2. *When two choice options are more similar, for example, around the switching choices where subjects switch between the lottery and the sure payment, the subjects have lower decision confidence, a higher likelihood of randomizing, and randomization probabilities closer to 0.5.*

3. *Compared to the decisions about the simple lottery, the decisions about the complex lottery exhibit lower decision confidence, as measured by the self-reported confidence measures, and randomization probabilities are affected in the same direction, maintaining a strong association between them.*

4. *Compared to the decisions in the no-experience treatment, the decisions in the experience treatments exhibit higher decision confidence, as measured by the self-reported confidence measures, and randomization probabilities are affected in the same direction, maintaining a strong association between them.*

4 Experimental results

We report the results in two steps. We begin by showing the systematic link between randomization probabilities and the two measures of self-reported confidence in the baseline no-experience treatment for decisions about the simple lottery (Hypotheses 1 and 2). We then show that decision confidence responded to our treatment manipulations in the expected direction by comparing the two measures of self-reported decision confidence across treatments. We demonstrate that exogenous shifts in self-reported decision confidence are paired with corresponding shifts in randomization probabilities, maintaining their systematic link (hypotheses 3 and 4). We pool subjects in different orders in our main analyses and discuss order effects in subsection 4.3.

4.1 Randomization probabilities and self-reported confidence

Below, we report two empirical observations that are consistent with Hypothesis 1 and 2.

Result 1. *In the baseline no-experience-simple-lottery treatment, randomization probabilities were significantly and positively correlated with confidence statements and probabilistic confidence among the large majority of subjects. Further, on average, randomization probabilities corresponded to probabilistic confidence in absolute levels.*

To obtain the correlation between randomization probabilities and confidence statements, we transformed the confidence statements to a scale of 1 to 5, with "Surely y" taking the value of 1 and "Surely x" taking the value of 5 to represent one's decision confidence in choosing x. We computed for each subject the nonparametric Spearman correlation between confidence statements and randomization probabilities in Experiment 1 and 2, and between probabilistic confidence and randomization probabilities in Experiment 2. The results for the baseline treatment are illustrated in Figure 4, Panel (a). Table B.1 in Appendix B summarizes the cross-measure correlations across treatments. Consistent with Hypothesis 1, confidence statements and randomization probabilities have a high and positive correlation. Moderate to strong correlations of 0.60 in Experiment 1 and

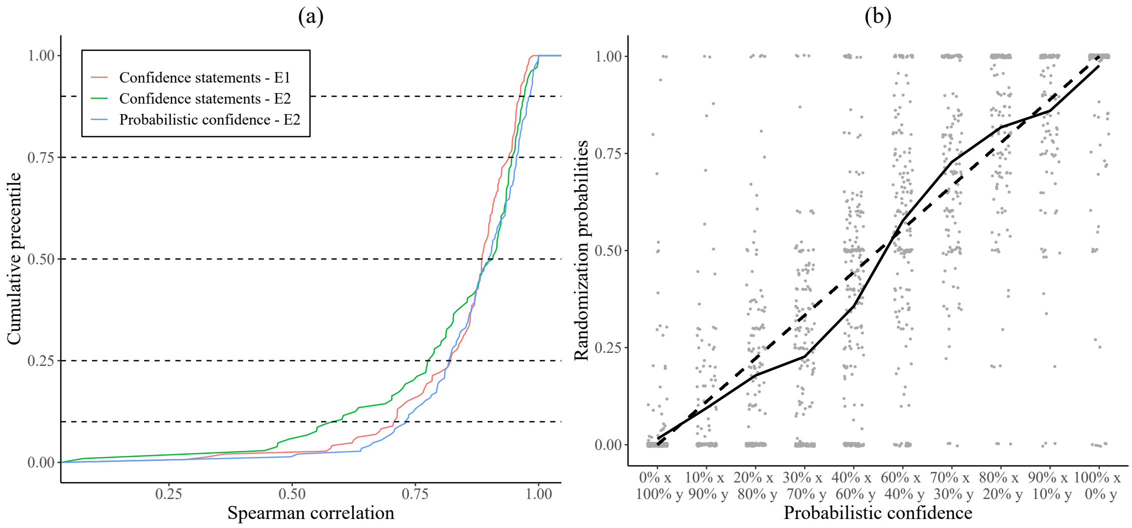


Figure 4: (a) Cumulative distributions of subjects' Spearman correlations between randomization probabilities and confidence statements or probabilistic confidence in the baseline no-experience treatment for decisions about the simple lottery. E1 and E2 refer to data from Experiment 1 and 2 respectively. (b) Scatter plot of randomization probabilities, with the mean randomization probability (in solid line) at each probabilistic confidence level in Experiment 2 in the baseline no-experience treatment for decisions about the simple lottery. The dashed line is a 45-degree line.

0.71 in Experiment 2 were found at the 10th percentile level, which increased to 0.91 in Experiment 1 and 0.89 in Experiment 2 at the median level. Since the subjects in Experiment 2 reported confidence statements and randomization probabilities in different sessions separated by at least seven days and in different orders, the similarities between the correlations found in Experiments 1 and 2 suggest that confidence statements and randomization probabilities are associated in ways beyond experimenter demand effects and order effects. In Experiment 2, we also found high correlation between self-reported probabilistic confidence and randomization probabilities: the correlation is 0.73 at the 10th percentile and 0.90 at the median.

As correlations do not describe the correspondence between randomization probabilities and self-reported decision confidence in absolute levels, we also computed the mean randomization probability at each level of probabilistic confidence for each subject and took the mean across subjects. This is shown in Panel (b) of Figure 4. Overall, the mean randomization probability for x is close to the probabilistic confidence of choosing x: subjects who chose a randomization probability of, for example, 0.7 would report probabilistic confidence of 70% on average. Examining the absolute correspondence between randomization probabilities and confidence statements gives a similar result, as summarized in Figure B.1 and Table B.2. These results suggest randomization probabilities can be used as a direct proxy for probabilistic confidence.

Next, we turn to Hypothesis 2 and examine the randomization probabilities around the switching choices. Intuitively, x and y are harder to compare around the switching choices. Reflecting this, subjects reported lower decision confidence and chose randomization probabilities close to 0.5 around the switching choices, as indicated in Result 2.

Result 2. *On average, the subjects reported lower decision confidence around the switching choices based on the self-reported confidence measures and were more likely to randomize and chose randomization probabilities close to 0.5 around the switching choices.*

We study the switching choice of each subject by considering two levels of sure payments: $\underline{y_b}$ and $\bar{y_b}$. We let $\underline{y_b}$ denote the highest sure payment at and below which the subject always preferred x over y, and $\bar{y_b}$ denote the lowest sure payment amount at and above which the subject always chose y over x in the binary choices. We henceforth refer to the values of y between $\underline{y_b}$ and $\bar{y_b}$ as the subject's switching range. This approach allows us to accommodate subjects who switched once as well as those who switched multiple times between lottery x and the sure payments (for the simple and complex lotteries, respectively, 19% and 25% in Experiment 1 and 14% and 23% in Experiment 2).[5]

As expected, decision confidence was lower within the switching range than outside it. In Experiment 2 (Experiment 1), 88% (85%) of the confidence statements within the switching range were "Probably x," "Unsure," or "Probably y," compared to 41% (40%) outside the switching range. In Experiment 2, "60% x, 40% y" and "40% x, 60% y" were selected for 53% of the values within the switching range, compared to 13% outside the switching range. Table 1 shows the median randomization probabilities, probabilistic confidence,

[5]It is important to include these subjects, because when subjects are not fully confident about their choices, they may switch between x and y multiple times. For the subjects who switched from the lottery to the sure payments once, the switching range simply includes the two sure payments around the switching choice (e.g., if a subject chooses the lottery at $y = 4$ and switches to the sure payment at $y = 4.5$ euros, this means that $\underline{y_b} = 4$, $\bar{y_b} = 4.5$, and the switching range is $[4, 4.5]$).

Experiment	Median behavior around the switching range					
	Confidence statements		Probabilistic Confidence		Randomization probabilities	
	$\underline{y_b}$	$\bar{y_b}$	$\underline{y_b}$	$\bar{y_b}$	$\underline{y_b}$	$\bar{y_b}$
Experiment 1 (N = 105)	Probably x	Probably y	-	-	0.67	0.46
Experiment 2 (N = 145)	Probably x	Probably y	60% x	40% x	0.65	0.45

Table 1: Median behavior around the switching choices in the baseline no-experience treatment for decisions about the simple lottery.

and confidence statements around the switching range. The median responses to the self-reported confidence measures indicate a lack of confidence around the switching range.

The randomization probabilities within the switching range resemble the two self-reported confidence measures. In Experiment 2 (Experiment 1), 67% (85%) of randomization probabilities reported for values of y within the switching range fell between 0.1 and 0.9, whereas this only holds for 33% (47%) outside the switching range. Further, in Experiment 2 (Experiment 1), the subjects assigned a median randomization probability of 0.65 (0.67) to x at $\underline{y_b}$, and a median randomization probability of 0.45 (0.46) to x at $\bar{y_b}$. The median randomization probability for all the choices that fell within the switching range is 0.5. These results are consistent with Hypothesis 2: subjects were more likely to choose randomization probabilities close to 0.5 for choices that they found difficult to compare.

4.2 Manipulating decision confidence

In this section, we examine whether our exogenous manipulations of the decision situation affected self-reported decision confidence in the expected direction and whether randomization probabilities were affected in similar ways to maintain a systematic relationship with the self-reported confidence measures.

4.2.1 The complex lottery versus the simple lottery

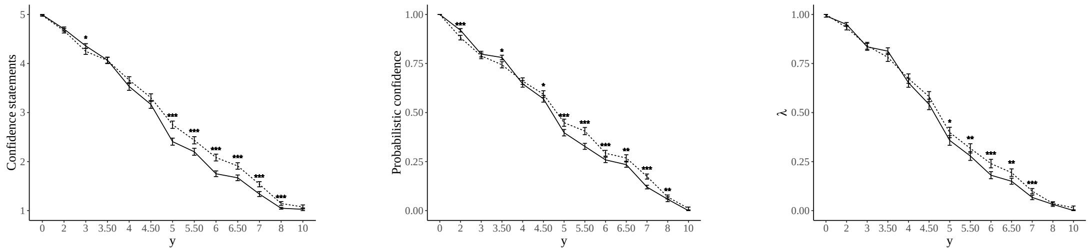
Consistent with Hypothesis 3, we find that subjects had lower decision confidence when making decisions about the complex lottery compared to the simple lottery. Result 3 summarizes our finding.

Result 3. *Compared to decisions about the simple lottery, the subjects revealed less than full decision confidence over a wider range of sure payments for decisions about the complex lottery, and their reported decision confidence were more compressed toward "Unsure" or (50% x, 50% y). Likewise, the subjects randomized over a wider range of sure payments and chose randomization probabilities closer to 0.5 for decisions about the complex lottery.*

We find that the range of sure payments over which the subjects chose confidence statements "Unsure" or "Probably" is larger for decisions about the complex lottery than for decisions about the simple lottery in both Experiment 1 and 2, and it is statistically significant in Experiment 2 (Experiment 1: 3.62 vs 3.36, Wilcoxon signed-rank test $p = 0.150$; Experiment 2: 3.58 vs 3.15, Wilcoxon signed-rank test $p < 0.01$). Comparing the range of sure payments for which subjects did not indicate probabilistic confidence of (100% x, 0% y) or (0% x, 100% y) gives similar results: the subjects were not fully confident over a wider range of sure payments for decisions about the complex lottery than decisions about the simple lottery (4.63 vs. 4.37, Wilcoxon signed-rank test p<0.01).

Panel (a) and Panel (b) of Figure 5 illustrate how confidence statements and probabilistic confidence varied with different sure payment amounts across the two lotteries in Experiment 2. Compared to the simple lottery, self-reported decision confidence measures were more compressed towards "Unsure" or (50% x, 50% y) when the subjects faced the complex lottery. The difference in decision confidence across the two lotteries is statistically significant between sure payments of 5 and 8 euros, and is less often statistically significant for lower sure payment amounts. The results for Experiment 1 are similar, albeit weaker, and can be found in Figure B.2 of Appendix B.

We proceed to examine the randomization probabilities chosen for each lottery. As we can see from Table B.3 in Appendix B, in both experiments, the range of sure payments over which subjects chose a randomization probability between 0.1 to 0.9 was significantly larger for decisions about the complex lottery than for decisions about the simple lottery (Experiment 1: 4.06 vs 3.63, Wilcoxon signed-rank test $p < 0.01$; Experiment 2: 3.19 vs 3.03, Wilcoxon signed-rank test $p < 0.10$). This is consistent with the findings from



(a) Confidence statements (b) Probabilistic confidence (c) Randomization probabilities

Figure 5: The mean self-reported decision confidence and randomization probabilities for each value of y for the simple lottery (solid line) and complex lottery (dashed line). Wilcoxon signed-rank tests were performed to test the treatment difference for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

self-reported decision confidence reported above.

Further, the randomization probabilities were also more compressed towards 0.5 when the subjects faced the complex lottery compared to the simple lottery, as shown in Panel (c) of Figure 5. The difference in randomization probabilities across the two lotteries is statistically significant between sure payments of 5 and 7 euros, coinciding with the range obtained from probabilistic confidence. Panel (c) of Figure 5 also shows asymmetric treatment effects on randomization probabilities for sure payments above 5 euros and sure payments below 5 euros. We show in Appendix A.3 that this asymmetric treatment effect can be consistent with the theoretical analysis when the treatment manipulation affects both preference uncertainty and the average valuation of the lotteries.

Despite our manipulation, the correlations between the two decision confidence measures and randomization probabilities remain similar. Comparing decisions about the simple lottery with those about the complex lottery, the median correlations between randomization probabilities and confidence statements are 0.86 vs 0.82 in Experiment 1 and 0.89 vs 0.88 in Experiment 2. The median correlations between randomization probabilities and probabilistic confidence are 0.90 vs 0.89 in Experiment 2. More results can be found in Table B.1, B.2, and B.4 in Appendix B.

4.2.2 Experience and no experience

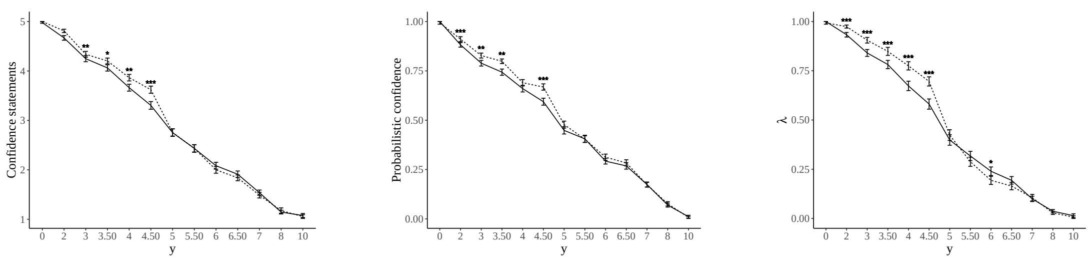
Hypothesis 4 states that, compared to the baseline no-experience treatment, gaining experience with the lotteries increases decision confidence. The results about Hypothesis 4 are summarized in Result 4.

Result 4. *Decisions in the partial-experience treatment and decisions about the simple lottery in the full-experience treatment did not exhibit significant treatment effects. Comparing decisions about the complex lottery in the full-experience treatment and the no-experience treatment, the subjects (1) revealed less than full decision confidence over a narrower range of sure payments, and their self-reported decision confidence were stretched further away from "Unsure" or (50% x, 50% y); and (2) they randomized over a narrower range of sure payments and chose randomization probabilities further away from 0.5.*

We report the results about the partial-experience treatment and the simple lottery in Figure B.3, Figure B.4 and Table B.5 in Appendix B, and report the comparison between the full-experience treatment and the no-experience treatment about the complex lottery here. We find that the range of sure payments over which subjects reported confidence statements of "Probably" or "Unsure" is significantly narrower (3.16 vs. 3.58, Wilcoxon rank-sum test $p < 0.05$). The range of sure payments over which subjects chose probabilistic confidence between 0.1 and 0.9 did not differ significantly between the full-experience treatment and the no-experience treatment (4.58 vs. 4.63, Wilcoxon rank-sum test $p = 0.590$).

Panels (a) and (b) of Figure 6 show how self-reported decision confidence differs between the full-experience and the no-experience treatment. Compared to the no-experience treatment, self-reported decision confidence was stretched further away from "Unsure" or (50% x, 50% y) for the subjects in the full-experience treatment and these differences in decision confidence were significantly different for sure payments between 2 and 4.5 euros. This implies that the subjects in the full-experience treatment were more confident about which option they preferred than subjects in the no-experience treatment.

Next, we examine these treatment effects on randomization probabilities. Like decision confidence, we find that the range of sure payments over which subjects chose random-



(a) Confidence statements (b) Probabilistic confidence (c) Randomization probabilities

Figure 6: The mean self-reported decision confidence and randomization probabilities for each value of y for the complex lottery in Experiment 2. The graphs show the baseline no-experience treatment (solid line) compared to the full-experience treatment (dashed line). Wilcoxon rank-sum tests were performed to test the treatment difference for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ization probabilities between 0.1 to 0.9 was significantly narrower in the full-experience treatment than in the no-experience treatment for decisions about the complex lottery (2.67 vs. 3.19, Wilcoxon rank-sum test $p < 0.05$). Panel (c) of Figure 6 shows that the difference in mean randomization probabilities across sure payments between the full-experience treatment and the no-experience treatment resembles that shown in Panel (a) and Panel (b). Compared to the no-experience treatment, randomization probabilities were also stretched further away from 0.5 among the subjects in the full-experience treatment. Significant differences in the randomization probabilities between subjects in the full-experience treatment and the no-experience treatment were also observed between 2 euros to 4.5 euros. Asymmetric treatment effects on randomization probabilities could also be observed here for sure payments above 5 compared to those below 5, which we discuss further in the theoretical models in Appendix A.3.

The increase in decision confidence from subjects' experience with the lotteries did not affect the high correlation between self-reported decision confidence and randomization probabilities. The median correlation between self-reported decision confidence and randomization probabilities was broadly similar in the two treatments. More details can be found in Table B.1, B.2, and B.4 in Appendix B.

4.3 Order effects

An important concern of the within-subject design is order effects: subjects' earlier decisions may affect their subsequent decisions. We are especially concerned with the order effects arising from priming: when randomization probabilities were elicited after self-reported confidence measures, subjects could be primed to link randomization probabilities to decision confidence and consequently reported randomization probabilities that cohered with self-reported decision confidence measures. Below we highlight the key findings about order effects and leave the details to Appendix B.1.

We find some order effects, suggesting that priming effects on randomized choices could be present. For example, subjects randomized strictly ($0 <$randomization probability< 1) in fewer choices when randomized choices were presented first (Order 2) compared to later (Order 1 and 3) in all treatments: averaging across treatments, 34% in Order 2 compared to 41% and 45% in Order 1 and 3 respectively (see Table B.6 and Figure B.5 in Appendix B.1 for details). Also, the cumulative distributions of the correlations in Order 2 tend to be lower (on the left of the other two orders), with a larger difference at the lower percentiles and in the complex lottery treatments.

Despite these differences, we find support for our hypotheses among the subjects in Order 2 where the aforementioned priming effects were absent, although the support is sometimes weaker than in the full sample, possibly due to the reduction of sample size. Figure B.6 in Appendix B.1 shows that across treatments, the median cross-measure correlations in Order 2 are high, consistent with H1. Subjects reported low decision around the switching choices and chose randomization probabilities around 0.5 in these choices, supporting H2 (see Table B.7). Table B.8 suggests that subjects in Order 2 had lower decision confidence in decisions involving the complex lottery compared to the simple lottery, consistent with H3. In Order 2, similar to the other two orders, decisions involving the complex lottery showed higher confidence in the full experience treatment than in the non-experience treatment, consistent with H4, although the treatment effect is not statistically significant in any individual order. This high level of consistency between randomization probabilities and self-reported confidence measures as well as their similar reactions to exogenous change of decision confidence suggest that these measures likely share common psychological foundations, even if not identical.

Finally, it is worth noting that the order effects discussed above do not necessarily suggest that randomization probabilities are a poorer proxy for decision confidence than the self-reported confidence measures. Confidence statements and probabilistic confidence are also noisy proxies of decision confidence, and it is not obvious what the "right" amount of strict randomization is. When we assess the value of decision confidence based on its correspondence with actual choices (when subjects report lower confidence in choosing an option, they should be less likely to choose that option), we find both randomization probabilities and probabilistic confidence corresponded to actual choices. Importantly, randomization probabilities exhibited a closer correspondence to binary choices than probabilistic confidence, especially in Order 2. See Appendix B.1 for more details.

4.4 Alternative interpretations of randomization probabilities

We have interpreted randomization behavior as a lack of decision confidence in the face of preference uncertainty. This is consistent with the findings from our post-experiment questionnaire as well as findings in Agranov and Ortoleva (2023) where many subjects explicitly mentioned unsureness, complexity, difficulty, and hedging as reasons for randomization (see Online Appendix C.1 for details). However, subjects may have other reasons for randomization. While it is not possible to eliminate all alternative interpretations, we show that indifference, random errors, or utility difference alone cannot be the driving force behind subjects' randomization behavior and the treatment effects.

First, indifference is not the driving reason for randomization because the majority of the subjects randomized at least twice (see Table B.9 in Appendix B.2), while randomization from indifference should occur for at most one value of sure payments. Second, randomization was unlikely a result of random errors because subjects' randomization probabilities of choosing x decreased monotonically with the value of y, even though they faced a random sequence of y (see e.g., Panel (c) in Figure 5). Further, the treatment effects on randomization probabilities in the expected directions suggest that subjects' randomization was likely a deliberate choice. Third, if randomization probabilities were due to utility differences alone, randomization probabilities should increase for each value of sure payments when the utility over the lottery increases (e.g., moving from the simple lottery to the complex lottery or the no-experience treatment to the full-experience treatment), which increases the lower bound and the upper bound of randomization as well. Our results clearly reject these predictions. Randomization probabilities were compressed toward 0.5 rather than increased monotonically, and the two bounds often moved in opposite directions. We elaborate on this in Appendix B.2.

5 Conclusion

We propose that letting individuals assign randomization probabilities according to which they receive options can be an incentivized way to elicit decision confidence. In two experiments, we elicited randomization probabilities as well as two self-reported confidence measures and further manipulated decision confidence exogenously.

We find that most subjects randomized frequently, and their randomization probabilities and self-reported confidence measures were linked in ways that are consistent with the hypotheses derived from two theoretical analyses. While there were some order effects depending on whether randomization probabilities were elicited before or after the self-reported confidence, cross-measure correlations were high, and randomization probabilities corresponded closely to probabilistic confidence in absolute levels, with a high randomization probability assigned to an option associated with high self-reported probabilistic confidence. Our further examination suggests that alternative interpretations of randomization such as indifference, random errors, or differences in utility alone are unlikely to be the driving factors. Overall, our results suggest that decision confidence can be meaningfully and accurately inferred from randomization probabilities.

References

Abdellaoui, M., Li, C., Wakker, P., and Wu, G. (2020). A defense of prospect theory in Bernheim and Sprenger's experiment. Working paper, Erasmus University Rotterdam.

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.

Apesteguia, J. and Ballester, M. A. (2018). Monotone stochastic choice models: The case of risk and time preferences. *Journal of Political Economy*, 126(1):74–106.

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

Aydogan, I. and Gao, Y. (2020). Experience and rationality under risk: Re-examining the impact of sampling experience. *Experimental Economics*, 23(4):1100–1128.

Bewley, T. F. (2002). Knightian Decision Theory: Part I. *Decisions in Economics and Finance*, 25:79–110.

Butler, D., Isoni, A., Loomes, G., and Tsutsui, K. (2014). Beyond choice: Investigating the sensitivity and validity of measures of strength of preference. *Experimental Economics*, 17:537–563.

Butler, D. and Loomes, G. (2011). Imprecision as an account of violations of independence and betweenness. *Journal of Economic Behavior & Organization*, 80(3):511–522.

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

Camerer, C. F. and Hogarth, R. (1999). The effects of financial incentives in experiments: A review and capital-labor-roduction framework. *Journal of Risk and Uncertainty*, 19:7–42.

Cerreia-Vioglio, S., Dillenberger, D., and Ortoleva, P. (2015). Cautious expected utility and the certainty effect. *Econometrica*, 83(2):693–728.

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., and Riella, G. (2019). Deliberately stochastic. *American Economic Review*, 109(7):2425–2445.

Chew, S. H. (1983). A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica*, 51(4):1065–1092.

Chew, S. H. (1989). Axiomatic utility theories with the betweenness property. *Annals of operations Research*, 19(1):273–298.

Chew, S. H., Miao, B., and Zhong, S. (2017). Partial ambiguity. *Econometrica*, 85(4):1239–1260.

Cohen, M., Jaffray, J.-Y., and Said, T. (1987). Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Decision Processes*, 39(1):1–22.

Cubitt, R., Kopsacheilis, O., and Starmer, C. (2022). An inquiry into the nature and causes of the Description - Experience gap. *Journal of Risk and Uncertainty*, 65(2):105–137.

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

Cubitt, R., van de Kuilen, G., and Mukerji, S. (2020). Discriminating between models of ambiguity attitude: A qualitative test. *Journal of the European Economic Association*, 18(2):708–749.

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40(2):304–318.

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

Dubourg, W. R., Jones-Lee, M. W., and Loomes, G. (1997). Imprecise preferences and survey design in contingent valuation. *Economica*, 64(256):681–702.

Dubra, J., Maccheroni, F., and Ok, E. A. (2004). Expected utility theory without the completeness axiom. *Journal of Economic Theory*, 115(1):118–133.

Dwenger, N., Kübler, D., and Weizsäcker, G. (2014). Flipping a coin: Theory and evidence. CESifo Working Paper 4740, Center for Economic Studies and ifo Institute (CESifo), Munich.

Dwenger, N., Kübler, D., and Weizsäcker, G. (2018). Flipping a coin: Evidence from university applications. *Journal of Public Economics*, 167:240–250.

Eliashberg, J. and Hauser, J. (1985). A measurement error approach for modeling consumer risk preference. *Management Science*, 31(1):1–25.

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.

Feldman, P. and Rehbeck, J. (2022). Revealing a preference for mixtures: An experimental study of risk. *Quantitative Economics*, 13(2):761–786.

Fudenberg, D., Iijima, R., and Strzalecki, T. (2014). Stochastic choice and revealed perturbed utility. Working paper.

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

Fudenberg, D. and Puri, I. (2021). Evaluating and extending theories of choice under risk. Working paper.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Gul, F. and Pesendorfer, W. (2006). Random expected utility. *Econometrica*, 74(1):121–146.

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

Hertwig, R., Barron, G., Weber, E. U., and Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15(8):534–539. PMID: 15270998.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

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.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Levitt, S. D. (2021). Heads or tails: The impact of a coin toss on major life decisions and subsequent happiness. *The Review of Economic Studies*, 88(1):378–405.

Loomes, G. and Pogrebna, G. (2017). Do preference reversals disappear when we allow for probabilistic choice? *Management Science*, 63(1):166–184.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, 92(368):805–824.

Loomes, G. and Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39(3-4):641–648.

Luce, R. and Suppes, P. (1965). *Preference, utility, and subjective probability*, chapter 3. in Handbook of Mathematical Psychology, Wiley, New York.

Machina, M. J. (1985). Stochastic choice functions generated from deterministic preferences over lotteries. *Economic Journal*, 95(379):575–94.

Miao, B. and Zhong, S. (2018). Probabilistic social preference: how Machina's mom randomizes her choice. *Economic Theory*, 65(1):1–24.

Mosteller, F. and Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, 59(5).

Myagkov, M. and Plott, C. R. (1997). Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments. *American Economic Review*, 87(5):801–828.

Ong, Q. and Qiu, J. (2023). Paying for randomization and indecisiveness. *Journal of Risk and Uncertainty*, 67:45–72.

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. *American Economic Review*, 95(3):530–545.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343.

Rubinstein, A. (2002). Irrational diversification in multiple decision problems. *European Economic Review*, 46(8):1369 – 1378.

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

van de Kuilen, G. (2009). Subjective probability weighting and the discovered preference hypothesis. *Theory and Decision*, 67:1–22.

van de Kuilen, G. and Wakker, P. (2006). Learning in the Allais paradox. *Journal of Risk and Uncertainty*, 33:155–164.

Vanberg, C. (2008). Why do people keep their promises? An experimental test of two explanations. *Econometrica*, 76(6):1467–1480.

Wulff, D. U., Mergenthaler-Canseco, M., and Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, 144(2):140–176.

Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

Appendices

A Appendix: Theoretical analysis

A.1 The theoretical analysis based on Cerreia-Vioglio et al. (2015) and Klibanoff et al. (2005)

To accommodate the potential that a decision-maker might not be fully confident about her choices, we assume an individual has multiple utility functions that we call multiple selves, with each self representing one particular way to trade off conflicting objectives in choices. Such a modelling technique has been used in models of incomplete preferences (see e.g., Bewley, 2002; Dubra et al., 2004; Cerreia-Vioglio et al., 2015).

Specifically, let u_τ denote the utility function of the self τ, and $\mathcal{T}$ denote the set of selves. Let π denote the subjective probability distribution over $\mathcal{T}$, which, similar to the modelling technique of Loomes and Sugden (1982), represents "the individual's degree of belief or confidence in the occurrence of the corresponding states" (Loomes and Sugden, 1982, p. 807). This belief could come from introspection or experiences with similar options. Given a utility function u_τ, we follow the standard assumption that the self behaves according to EUT. Let $U_\tau(l)$ denote the expected utility of an option $l \in L$.[6] We further assume that the individual dislikes disagreement among selves. This is because, to arrive at a choice when there are multiple selves with different preferences is, in essence, similar to situations where a group of people with different opinions tries to reach a consensus. The more strongly group members disagree with each other, the harder it is for the group to make compromises and agree on a single opinion. Hence, aversion to disagreement among selves can be interpreted as the cost of forcing different selves to reach a consensus. With

[6]The function $U(\cdot)$ could be made more general to allow for non-EUT preferences to incorporate unsureness about how strongly to weight the extra factor, such as probability weighting or loss aversion, in a non-EU model.

the above assumptions, we can write the individual's preference over an option l as:

$$V(l) = \int_{\mathcal{T}} \phi\left[U_{\tau}(l)\right] d\pi, \tag{1}$$

where concave $\phi(\cdot)$ implies an aversion to disagreement - deviations from the mean expected utility - among different selves. Similar to the connection between the concavity of the utility function and risk aversion, the concavity of $\phi(\cdot)$ implies that the individual places more weight on the selves who have lower value for l. Such a cautious attitude is consistent with Levitt (2021) who showed that subjects who have difficulties making a decision are often excessively cautious in the sense of preferring to maintain the status quo.

Equation 1 extends directly from Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015). It can be seen as a smooth version of the cautious expected utility model (Cerreia-Vioglio et al., 2015). It is also a parallel of the smooth ambiguity model of Klibanoff et al. (2005). Indeed, in the smooth ambiguity model, an individual is unsure about the probability distribution of the states of nature, and she has a subjective belief over these probability distributions. Likewise, in this model, an individual is unsure about her utility function, and she has a subjective belief over her multiple selves. Note that this does not mean this model only applies to decision-making under risk. If there is preference uncertainty under risk (or even under certainty, e.g., over options about experience goods) because individuals have difficulties evaluating options, this uncertainty is also likely to be present in more complex situations of decision-making under ambiguity. In this sense, this model complements the smooth model of ambiguity and general models about uncertainty in beliefs. Ultimately, the lack of decision confidence arises from the difficulties in evaluating options, which may be due to uncertainty in both beliefs and preferences. A general model accommodating both sources of uncertainty could be written as:

$$V(a) = \int_{\mathcal{M}} \int_{\mathcal{T}} \phi\left[U_{\tau,\mu}(a)\right] d\pi d\mu,$$

where a represents an act, and μ is a subjective probability distribution over M, the set of probability distributions of the states of nature.

We are now ready to establish the link between decision confidence and the randomization probability in the randomized choices. Specifically, recall that in our mechanism, the individual chooses a randomization probability $\lambda \in [0,1]$ and builds a lottery $(\lambda, x; (1-\lambda), y)$: She receives x with probability λ and y with probability $1-\lambda$. Since for any given self τ, the individual's preference over the lottery $(\lambda, x; (1-\lambda), y)$ satisfies EUT, we have $U_\tau [\lambda x + (1-\lambda)y] = \lambda U_\tau(x) + (1-\lambda)U_\tau(y)$. The individual's decision is then to maximize her utility by choosing the optimal randomization probability $0 \le \lambda \le 1$:

$$Max_\lambda \; V[\lambda x + (1-\lambda)y] = \int_{\mathcal{T}} \phi [\lambda U_\tau(x) + (1-\lambda)U_\tau(y)] \, d\pi.$$

In the experiment, y is a sure payment. Sure monetary payments are probably the easiest options to evaluate, hence we assume the individual is always confident about her evaluation of a sure payment: $U_\tau(y) = u(y), \; \forall \tau \in \mathcal{T}$. Applying the Taylor expansion to the above equation at y, we can derive the optimal λ as:[7]

$$\lambda^* \approx \frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}} \times \frac{E_\pi [U_\tau(x)] - u(y)}{\sigma_x^2} \tag{2}$$

where $\sigma_x^2 = E_\pi [U_\tau(x) - E_\pi(U_\tau(x))]^2$ is the standard deviation of the valuation of the lottery across multiple selves and approximates how strongly different selves disagree with each other. Similar to decision-making under risk, $-\frac{\phi''(u(y))}{\phi'(u(y))}$ can be interpreted as a metric of attitudes towards disagreement among selves. Thus, the randomization probability aggregates the three important determinants of decision confidence: preference uncertainty, the utility difference between the two options, and her attitude toward preference uncertainty. It is in this sense we argue that the randomization probability captures decision confidence.

Deriving the hypotheses

To see how the individual may randomize for sure payments that yield similar utility as

[7]More precisely, since $0 \le \lambda \le 1$, $\lambda^* \approx min\left\{max\left\{0, \frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}} \times \frac{\Delta_u}{\sigma_x^2}\right\}, 1\right\}$. The detailed derivation can be found below.

the lottery, notice that the certainty equivalent of the lottery is

$$\begin{aligned}
u(CE_x) &= \int \phi\left[U_\tau(l)\right] d\pi \\
&\approx E_\pi \left\{ E_\pi\left[U_\tau(x)\right] + \phi'\left(E_\pi\left[U_\tau(x)\right]\right)\left[U_\tau - E_\pi\left[U_\tau(x)\right]\right] + \frac{\phi''\left(E_\pi\left[U_\tau(x)\right]\right)}{2}\left[U_\tau - E_\pi\left[U_\tau(x)\right]\right]^2 \right\} \\
&= E_\pi\left[U_\tau(x)\right] + \frac{\phi''\left(E_\pi\left[U_\tau(x)\right]\right)}{2}\sigma_x^2.
\end{aligned}$$

The optimal randomization probability at the sure payment which is equal to the certainty equivalent of the lottery ($u(y) = u(CE_x) = E_\pi\left[U_\tau(x)\right] + \frac{\phi''(E_\pi[U_\tau(x)])}{2}\sigma_x^2$) is

$$\lambda^* \approx \frac{1}{-\frac{\phi''[u(CE_x)]}{\phi'[u(y)]}} \times \frac{E_\pi\left[U_\tau(x)\right] - u(CE_x)}{\sigma_x^2} = \frac{1}{2} \times \frac{\phi'\left[u(CE_x)\right]\phi''\left(E_\pi\left[U_\tau(x)\right]\right)}{\phi''\left[u(CE_x)\right]}.$$

When $\phi'\left[u(CE_x)\right]$ is close to one and the function $\phi(\cdot)$ is smoothly concave, which is likely to hold for options with moderate payoffs, the randomization probability is around 0.5. This implies that the individual would choose randomization probabilities close to 0.5 when two options yield similar utilities. Furthermore, the smallest sure payment that the individual chooses $\lambda^* < 1$ (the lower bound), and the largest sure payment that the individual chooses $\lambda^* > 0$ (the upper bound) are defined by $u(\underline{y_x}) = E_\pi\left[U_\tau(x)\right] - \frac{-\phi''[u(y)]}{\phi'[u(y)]}\sigma_x^2$, and $u(\bar{y_x}) = E_\pi\left[U_\tau(x)\right]$. The range of sure payments that the individual randomizes strictly is

$$u(\bar{y_x}) - u(\underline{y_x}) = \frac{-\phi''\left[u(y)\right]}{\phi'\left[u(y)\right]}\sigma_x^2,$$

which varies with preference uncertainty (σ_x^2).

Relating these results to our experiment, we expect subjects to have more preference uncertainty about a complex lottery than a simple lottery, as the individual may find it harder to evaluate a complex lottery. She considers relevant a larger set of utility functions and the subjective belief π becomes flatter. This translates into larger preference uncertainty (δ_x increases). Experience with a lottery, on the other hand, reduces preference uncertainty about the lottery because the individual attains clearer preferences about the lottery when she gains more experience (the set of utility functions becomes smaller and

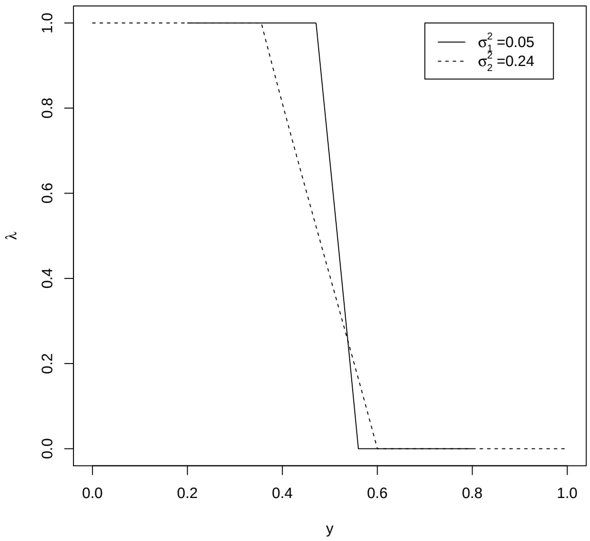


Figure A.1: The relationship between the randomization probability λ and the sure payment y. The figure is produced by assuming $\phi(U_\tau) = 1 - e^{-U_\tau}$, $\pi(u_1) = 0.6$, $\pi(u_2) = 0.4$, $U_1(x_1) = 0.8$ and $U_2(x_1) = 0.2$, $U_1(x_2) = 1.0$ and $U_2(x_2) = 0$, and $U_1(y) = U_2(y) = y$.

δ_x decreases). These lead to the hypotheses in the main text.

As a concrete illustration, consider the following numerical example: the individual has two selves $\tau = 1, 2$, and $\pi(u_1) = 0.6$, $\pi(u_2) = 0.4$. The individual's preference over the lottery x_1 is such that $U_1(x_1) = 0.8$ and $U_2(x_1) = 0.2$. Her preference over the lottery x_2 is such that $U_1(x_2) = 1.0$ and $U_2(x_2) = 0$. Thus, the individual perceives more preference uncertainty about the lottery x_2 than the lottery x_1 ($\sigma_{x_1} = 0.05 < \sigma_{x_2} = 0.24$). Option y is a sure payment, and $u_1(y) = u_2(y) = y$. The function $\phi(U_\tau) = 1 - e^{-U_\tau}$. Simple calculation shows that $\lambda_{x_1} = -\frac{1}{0.8-0.2} ln(\frac{0.4}{0.6} \times \frac{y}{1-y})$ and $\lambda_{x_2} = -\frac{1}{1-0} ln(\frac{0.4}{0.6} \times \frac{y}{1-y})$, subject to $0 \leq \lambda \leq 1$. Figure A.1 shows the relationship between the optimal λ and sure payment y. The Figure shows that the randomization probability decreases with y, and approaches to 0.5 for y that yields similar decision utility as the lottery ($y = 0.515$ for x_1 and $y = 0.476$ for x_2). Furthermore, the individual randomizes over a wider range of y for x_2 which she perceives higher preference uncertainty compared to x_1.

Derivation of the optimal λ^*

Taking the first order derivative of the optimization equation gives:[8]

$$\frac{dV\left[\lambda x+(1-\lambda)y\right]}{d\lambda}=\int_{\mathcal{T}}\phi'\left[\lambda U_{\tau}(x)+(1-\lambda)u(y)\right]\times\left[U_{\tau}(x)-u(y)\right]d\pi=0.$$

Note that $U_{\tau}(x)$ is a random variable governed by the subjective probability distribution π. Let $X=U_{\tau}(x)$, and $\Delta_{\tau}=X-u(y)$. With these notations, we have

$$\phi'\left[\lambda U_{\tau}(x)+(1-\lambda)u(y)\right]=\phi'\left[u(y)+\lambda\Delta_{\tau}\right].$$

We are most interested in scenarios where the individual finds it difficult to choose between x and y, i.e., when the two options are close and Δ_{τ} is small relative to X and $u(y)$. When this is the case, we can use the Taylor expansion at y and obtain

$$\phi'\left[u(y)+\lambda\Delta_{\tau}\right]=\phi'(u(y))+\phi''(u(y))\lambda\Delta_{\tau}+O\left(\lambda\Delta_{\tau}\right)\approx\phi'(u(y))+\phi''(u(y))\lambda\Delta_{\tau},$$

where $O\left(\lambda\Delta_{\tau}\right)$ is the sum of the terms that have $\lambda\Delta_{\tau}$ with a power of two or higher. The above first order condition can be written as

$$\begin{aligned}\tfrac{dV[\lambda x+(1-\lambda)y]}{d\lambda} &= \textstyle\int_{\mathcal{T}}\phi'\left[u(y)+\lambda\Delta_{\tau}\right]\Delta_{\tau}d\pi,\\ &\approx \textstyle\int_{\mathcal{T}}\left[\phi'(u(y))+\phi''(u(y))\lambda\Delta_{\tau}\right]\Delta_{\tau}d\pi\\ &= E_{\pi}\left[\phi'(u(y))\Delta_{\tau}\right]+\lambda E_{\pi}\left[\phi''(u(y))\Delta_{\tau}^{2}\right]=0,\end{aligned}$$

where $E_{\pi}(\cdot)$ is the expectation operator with respect to the distribution π. Solving for λ, we have:

$$\lambda^{*}\approx min\left\{max\left\{0,\frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}}\times\frac{\Delta_{u}}{\sigma_{x}^{2}-\Delta_{u}^{2}}\right\},1\right\}\approx min\left\{max\left\{0,\frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}}\times\frac{\Delta_{u}}{\sigma_{x}^{2}}\right\},1\right\},$$

where $\Delta_{u}=E_{\pi}\left[U_{\tau}(x)\right]-u(y)$ is the (expected) utility difference of x and y, $\sigma_{x}^{2}=E_{\pi}\left[U_{\tau}(x)-E_{\pi}(U_{\tau}(x))\right]^{2}$ is the standard deviation of $U_{\tau}(x)$.

---

[8]The second-order derivative is $\frac{d^{2}V[\lambda x+(1-\lambda y)]}{d\lambda^{2}}=\int_{\mathcal{T}}\phi''\left[\lambda U_{\tau}(x)+(1-\lambda)u(y)\right]\times\left[U_{\tau}(x)-u(y)\right]^{2}d\pi$. Since $\phi(\cdot)$ is concave, $\phi''(\cdot)$ is negative. We are interested in situations where options x and y are not the same, i.e., $U_{\tau}(x)\neq u(y)$ for some $\tau\in\mathcal{T}$. Together we have $\phi''\left[\lambda U_{\tau}(x)+(1-\lambda)u(y)\right]\times\left[U_{\tau}(x)-u(y)\right]^{2}\leq 0$, and the inequality is strict for some $\tau\in\mathcal{T}$. Consequently, $\frac{d^{2}V[\lambda x+(1-\lambda y)]}{d\lambda^{2}}=\int_{\mathcal{T}}\phi''\left[\lambda U_{\tau}(x)+(1-\lambda)u(y)\right]\times\left[U_{\tau}(x)-u(y)\right]^{2}d\pi<0$. This ensures we are indeed seeking for the maximum.

A.2 The theoretical analysis based on Fudenberg et al. (2015)

Below, we perform a theoretical analysis of our experiment based on Fudenberg et al. (2015) to demonstrate the links between randomization probabilities and decision confidence.[9] Fudenberg et al.'s (2015) original representation concerns final outcomes. To apply their model to our experiments with lotteries, we write the individual's preference over randomizing between lottery x and sure payment y as:[10]

$$V(\lambda, x; 1-\lambda, y) = \lambda U(x) - c(\lambda) + (1-\lambda)u(y) - c(1-\lambda),$$

where $U(x)$ is the expected utility of the lottery x and $c(\lambda)$ is a weak cost function with finite steepness (the first order derivative of the cost function at the limit of 0 is not infinite). Using the weak cost function allows the model to accommodate zero choice probability that is present in our experiment. The cost function captures the implementation costs of making the desired choice, such as time and cognitive resources. In the Fudenberg et al.'s (2015) main representation, the cost function is independent of the option and the choice set. In an earlier version of their paper (Fudenberg et al., 2014), they proposed two extensions (item-invariant and menu-invariant APU) in which the cost function may depend on the preference uncertainty over options or the choice problem. We consider these two extensions to examine the effects of our treatments (increasing the complexity of the lottery or increasing subjects' experience with the lottery) on the cost function.

When $c(\lambda)$ is strictly convex, there exists an optimal randomization probability λ^* which maximizes the individual's utility, as defined by the equation $c'(\lambda^*) - c'(1-\lambda^*) = U(x) - u(y)$, where $c'(\lambda^*) - c'(1-\lambda^*)$ measures the convexity of the cost function $c''(\cdot)$. While

[9] Cerreia-Vioglio et al. (2019) predict preference for randomization when the individual faces non-degenerated lotteries. However, when one of the two options is a sure payment, as in our experiment, the individual has no preference for randomization. This follows directly from the axiom of Weak Stochastic Certainty Effect.

[10] Cerreia-Vioglio et al. (2019, Footnote 22, p.2437) proposed an alternative approach in which the individual integrates the lottery and the sure payment into a compound lottery, applies the reduction of the compound lottery, and implements the cost function to each outcome. We illustrate their approach and point out the differences between the two below. In particular, that approach predicts that the optimal randomization probability for the pair of the lottery and the sure payment that the individual is indifferent with depends on the number of outcomes in the lottery.

the exact value of the optimal randomization probability depends on the cost function, some observations are in order. First, the optimal randomization probability approaches 0.5 when $U(x)$ is close to $u(y)$. Second, for the same utility difference between the two options, the individual chooses a randomization closer to 0.5 when the cost function is more convex. More generally, as Proposition 3 in Fudenberg et al. (2015) demonstrates, the individual becomes less selective and randomizes more when $c''(\cdot)$ increases. Third, simple calculations show that the largest sure payment that the individual chooses $\lambda^* = 0.9$ (the lower bound) is $u(\underline{y}) = U(x) - \Delta$, and the smallest sure payment she chooses $\lambda^* = 0.1$ (the upper bound) is $u(\bar{y}) = U(x) + \Delta$, where $\Delta = c'(0.9) - c'(0.1) > 0$.[11] Thus, the individual randomizes over a larger range of sure payments when the cost function is more convex $(u(\bar{y}) - u(\underline{y})) = 2\Delta)$. According to Fudenberg et al. (2015), the cost function may depend, among other things, on the individual's perceived preference uncertainty over the options and her attitude towards uncertainty. Using this interpretation of the cost function, the three properties of randomization probabilities correspond to the three properties of decision confidence we outlined in the main body of the paper. It is in this sense that we say randomization probabilities measure decision confidence.

If we are willing to make more specific assumptions about the cost function, we can obtain a direct solution of the optimal randomization probability. For example, when the cost function takes the form of $c(\lambda) = \eta\lambda log(\lambda)$, we can derive the familiar logit/logistic choice rule:

$$\lambda^* = \frac{e^{U(x)/\eta}}{e^{U(x)/\eta} + e^{u(y)/\eta}}. \tag{3}$$

As shown by Holman and Marley, the parameter η 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). In the context of our study, η can be interpreted as the individual's preference uncertainty about lottery x. Figure A.2 depicts the relationship between the optimal randomization probability λ^* and the sure payments y. As we can see, randomization probabilities decrease with the value of y and approach 0.5 when the two options have

[11]The values of 0.1 and 0.9 were chosen to accommodate experimental data.

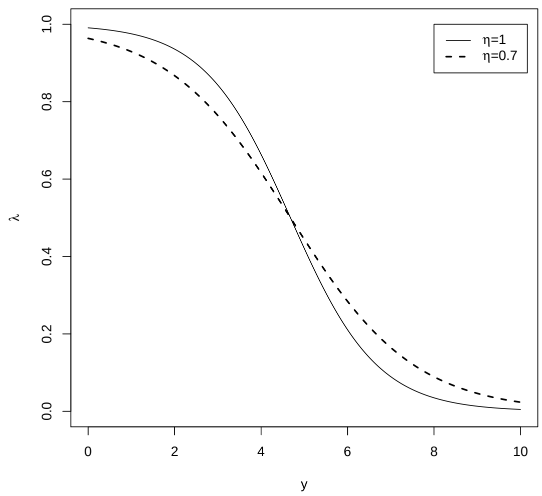
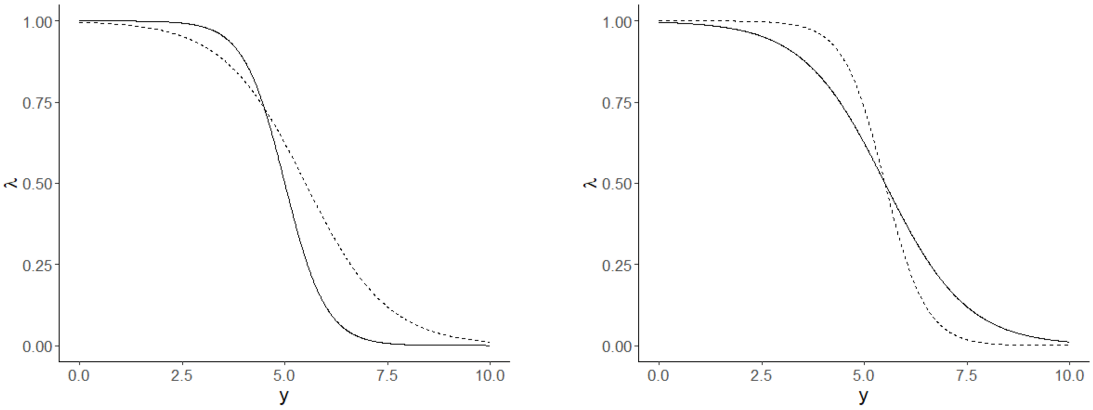


Figure A.2: The relationship between the optimal randomization probability λ^* and the sure payments y. The figure is produced according to the logit/logistic choice rule $\lambda^* = \frac{e^{U(x)/\eta}}{e^{U(x)/\eta}+e^{u(y)/\eta}}$. The parameter η captures the preference uncertainty over lottery x, with a larger η implying more convexity in the cost function and thus more preference uncertainty.

similar utilities. Furthermore, when η increases, the cost function becomes more convex and the individual's randomization probabilities become more compressed (the dashed line) and closer to 0.5.

Individuals may perceive more preference uncertainty over the complex lottery than over the simple lottery ($\Delta_c > \Delta_s$, where c denotes the complex lottery and s denotes the simple lottery), and experience with the lottery may reduce preference uncertainty about the lottery ($\Delta_e < \Delta_n$, where e denotes experience and n denotes no experience). In light of our analysis above, we expect that subjects' randomization probabilities are closer to 0.5 and that they randomize strictly over a wider range of sure payments when they make decisions about the complex lottery than when they make decisions about the simple lottery. In addition, compared to the no-experience treatment, randomization probabilities of subjects in the experience treatments are stretched away from 0.5, and subjects randomize strictly over a smaller range of sure payments. Figure A.3 demonstrates the effects.

Cerreia-Vioglio et al. (2019)'s approach



(a) Simple versus complex

(b) No experience versus full experience

Figure A.3: The effects of complexity and experience on the lower bound, the upper bound, and the size of randomization range.

Cerreia-Vioglio et al. (2019, Footnote 22, p.2437) proposed an alternative approach to apply Fudenberg et al.'s (2015) model to lotteries. We illustrate their approach with the following example. Consider an individual who faces a choice between a sure payment y and a lottery $x = 9_{0.5}1$ which pays 9 or 1 with equal likelihood. Cerreia-Vioglio et al. (2019) treat the randomized choice as a compound lottery. With the reduction of the compound lottery, the randomization of $(\lambda, x; 1-\lambda, y)$ becomes $9_{0.5\lambda}1_{0.5\lambda}y$, and the individual' preference over $9_{0.5\lambda}1_{0.5\lambda}y$ is

$$\begin{aligned} V(\lambda, x; 1-\lambda, y) &= 0.5\lambda u(9) - c(0.5\lambda) + 0.5\lambda u(1) - c(0.5\lambda) + (1-\lambda)u(y) - c(1-\lambda) \\ &= \lambda U(x) - 2c(0.5\lambda) + (1-\lambda)u(y) - c(1-\lambda) \end{aligned}$$

This formulation predicts an optimal randomization probability of 2/3 when the expected utility of the lottery is close to the utility of the sure payment ($c'(0.5\lambda) - c'(1-\lambda) = U(x) - u(y) = 0 \Rightarrow \lambda = 2/3$). The intuition is that the above formulation rewards the individual for randomizing over *more outcomes*, and thus the individual assigns a higher randomization probability to lotteries with more outcomes. It can be shown that, when the lottery x has four outcomes which are equally likely, the optimal randomization probability is $\lambda = 4/5$ when $U(x) = u(y)$. These predictions are different from those obtained based on Fudenberg et al. (2015)'s approach.

A.3 The asymmetric treatment effects on the lower and upper bound of randomization range

We illustrate the asymmetric treatment effects on the lower bound and the upper bound of randomization range in this section. Recall that $\underline{y}$ denotes the largest sure payment that the individual chooses $\lambda^* < 1$ (the lower bound) and $\bar{y}$ denotes the smallest sure payment she chooses $\lambda^* > 0$ (the upper bound).

In the model extended from Cerreia-Vioglio et al. (2015) and Klibanoff et al. (2005),

$$\begin{aligned} u(\bar{y_x}) &= E_\pi \left[U_\tau(x)\right], \\ \underline{y} &= E_\pi \left[U_\tau(x)\right] - \frac{-\phi'' \left[u(y)\right]}{\phi' \left[u(y)\right]}\sigma_x^2. \end{aligned}$$

The changes in the upper and lower bounds depend on both $E_\pi \left[U_\tau(x)\right]$ and σ_x^2. We observe that subjects on average valued the complex lottery higher than the simple lottery (mean CE of 4.68 for the simple lottery versus 4.98 for the complex lottery in Experiment 2, p<0.01). Since the complex lottery has a larger σ_x^2 and the average valuation of the lottery is $E_\pi \left[U_\tau(x)\right] - \frac{-\phi''(E_\pi[U_\tau(x)])}{2}\sigma_x^2$, this implies an increase in $E_\pi \left[U_\tau(x)\right]$ for the complex lottery. The increase in $E_\pi \left[U_\tau(x)\right]$ increases both the upper bound and the lower bound, while the increase in σ_x^2 decreases only the lower bound. Together, they imply that the treatment effect on the upper bound could be larger than on the lower bound. Similarly, we observe an increase, albeit small, in the valuation of the complex lottery in the full-experience treatment (mean CE of 4.98 in the no-experience treatment versus 5.07 in the full-experience treatment in Experiment 2, p>0.10). The increase in $E_\pi \left[U_\tau(x)\right]$ increases both the upper and lower bounds, and the decrease in σ_x^2 increases the lower bound further. Consequently, the treatment effect could be stronger on the lower bound than on the upper bound.

The analysis based on Fudenberg et al. (2015) follows similarly. In Fudenberg et al. (2015), $\bar{y} = EU(x) + \Delta$, $\underline{y} = EU(x) - \Delta$. The changes in the upper and lower bounds depend on both $EU(x)$ and Δ. Since the average valuation of the lottery is $EU(x)$, the higher average valuation of the complex lottery implies higher $EU(x)$ of the complex lottery compared to the simple lottery. Higher $EU(x)$ and Δ imply a stronger treatment effect on the upper bound than on the lower bound. Likewise, an increase in experience level is associated with an increase in $EU(x)$ and a decrease in Δ, which jointly imply a stronger treatment effect on the lower bound than the upper bound.

B Additional figures and tables

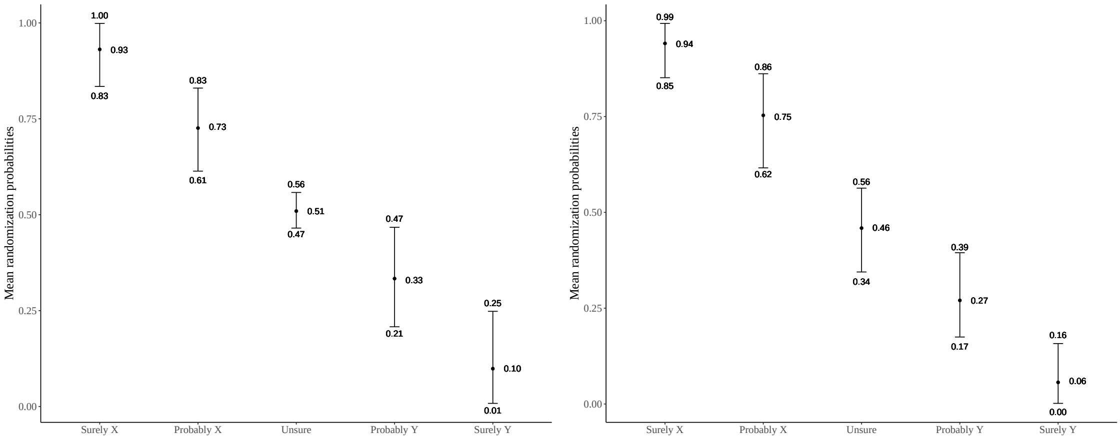
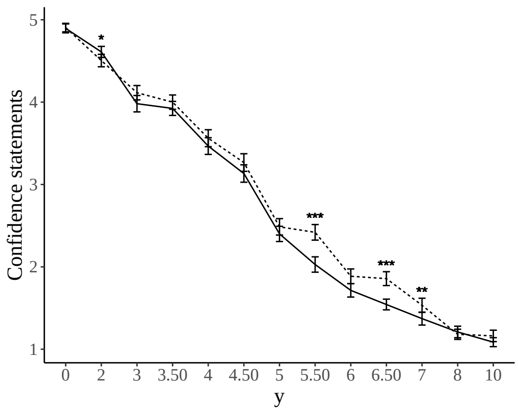
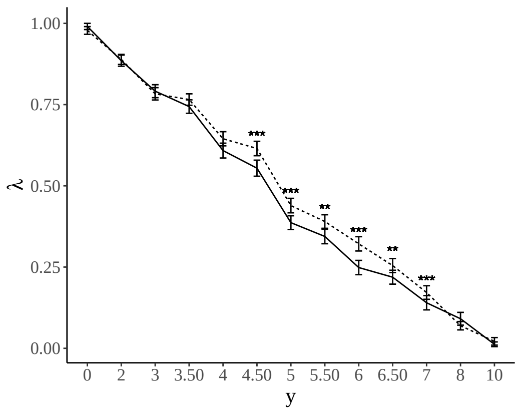


Figure B.1: The mean randomization probabilities at each confidence statement. The bars show the average minimum and maximum values. The values show the aggregate values for the baseline treatment – simple lottery, no-experience – in Experiment 1 (left) and Experiment 2 (right). The mean, minimum, and maximum values for the separate treatments in each of the experiments can be found in Table B.2. These values are broadly consistent with the cutoff probabilistic confidence levels of each confidence statement reported in Vanberg (2008, Footnote 10, p.1472: the probabilistic confidence level of 0.85 as the cutoff between surely and probably, 0.68 as the cutoff between probably and unsure, and 0.50 as the mean value for unsure). The minimum randomization probabilities were 0.83 and 0.85 for "Surely x" and 0.61 and 0.62 for "Probably x," and the mean randomization probabilities were 0.51 and 0.46 for "Unsure" in Experiments 1 and 2 respectively.

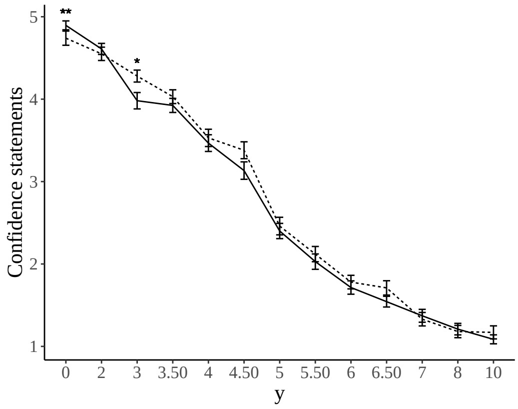


(a) Confidence statements

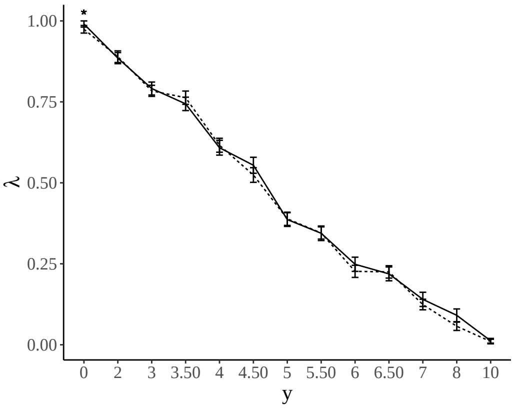


(b) Randomization probabilities

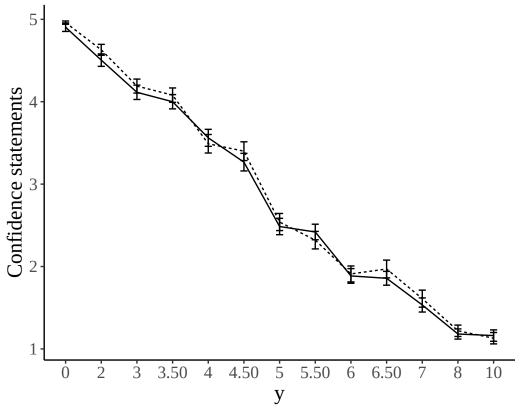
Figure B.2: The mean self-reported decision confidence and randomization probabilities for each value of y obtained from decisions about the simple lottery (solid line) and decisions about the complex lottery (dashed line) in Experiment 1. Wilcoxon signed-rank tests were performed to test the difference between the simple lottery and the complex lottery for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



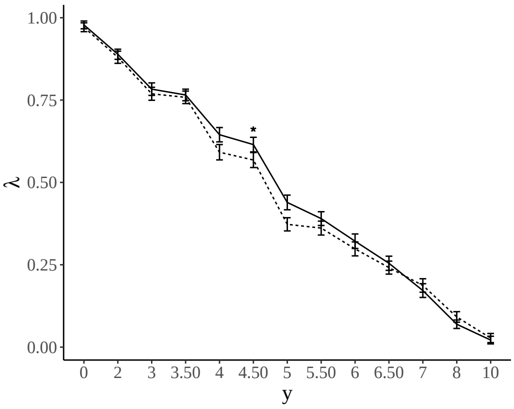
(a) Confidence statements, simple lottery



(b) Randomization probabilities, simple lottery

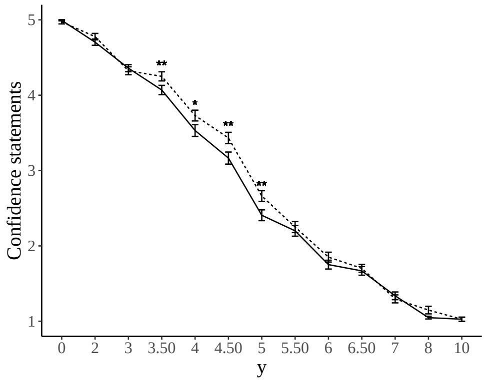


(c) Confidence statements, complex lottery

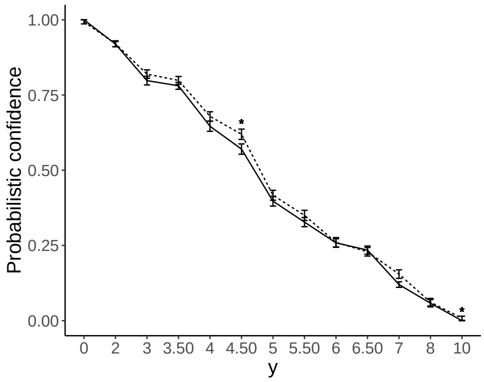


(d) Randomization probabilities, complex lottery

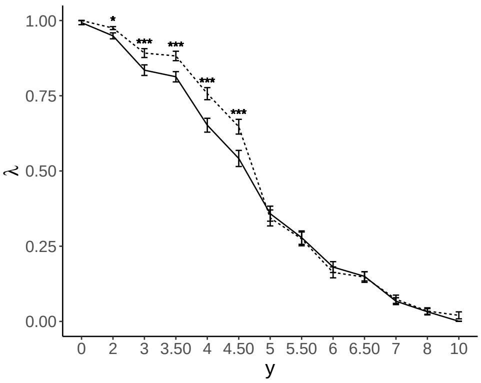
Figure B.3: The mean self-reported decision confidence and randomization probabilities for each value of y in the no-experience treatment (solid line) and the partial-experience treatment (dashed line) in Experiment 1. Wilcoxon rank-sum tests were performed to test the difference between the partial-experience treatment and no-experience treatment for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



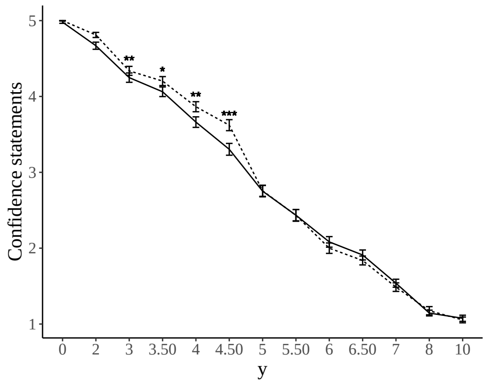
(a) Confidence statements, simple lottery



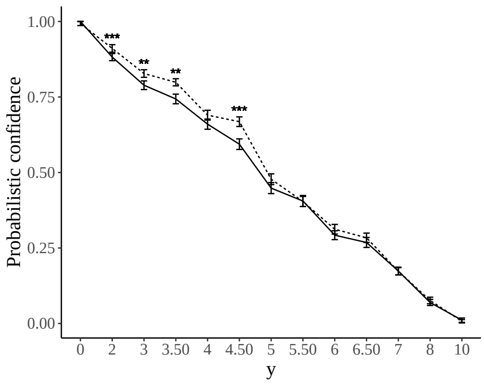
(b) Probabilistic confidence, simple lottery



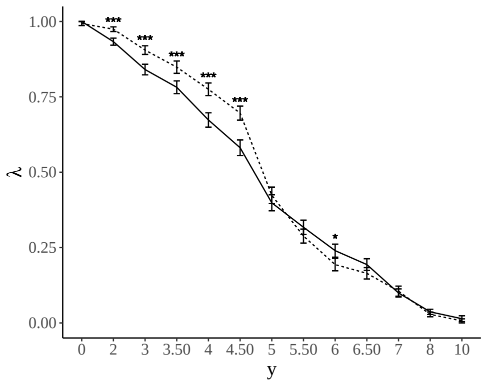
(c) Randomization probabilities, simple lottery



(d) Confidence statements, complex lottery



(e) Probabilistic confidence, complex lottery



(f) Randomization probabilities, complex lottery

Figure B.4: The mean self-reported decision confidence and randomization probabilities for each value of y in the no-experience treatment (solid line) and the full-experience treatment (dashed line) in Experiment 2. Wilcoxon rank-sum tests were performed to test the difference between full-experience treatment and no-experience treatment for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Lottery	Treatment		Correlation between randomization probabilities and confidence statements		Correlation between randomization probabilities and prob. confidence
			Experiment 1	Experiment 2	Experiment 2
Simple	No experience	10th percentile	0.60	0.71	0.73
		25th percentile	0.78	0.82	0.82
		median	0.91	0.89	0.90
		75th percentile	0.95	0.94	0.96
		90th percentile	0.97	0.96	0.98
	Experience	10th percentile	0.60	0.78	0.77
		25th percentile	0.85	0.85	0.85
		median	0.93	0.90	0.91
		75th percentile	0.96	0.95	0.96
		90th percentile	0.97	0.97	0.99
Complex	No experience	10th percentile	0.69	0.67	0.64
		25th percentile	0.83	0.77	0.83
		median	0.90	0.88	0.89
		75th percentile	0.94	0.93	0.95
		90th percentile	0.97	0.96	0.97
	Experience	10th percentile	0.62	0.69	0.77
		25th percentile	0.81	0.80	0.84
		median	0.88	0.90	0.90
		75th percentile	0.94	0.94	0.95
		90th percentile	0.96	0.97	0.97

Table B.1: Nonparametric Spearman correlation between randomization probabilities and the two self-reported confidence measures at the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile in the two experiments for each lottery and experience treatment group.

Treatment	Lottery		Surely x	Probably x	Unsure	Probably y	Surely y
Experiment 1							
No-experience	Simple	Mean	0.93 (0.011)	0.73 (0.016)	0.51 (0.023)	0.33 (0.018)	0.10 (0.010)
		Min	0.83 (0.027)	0.61 (0.022)	0.47 (0.027)	0.21 (0.019)	0.01 (0.005)
		Max	1 (0.001)	0.83 (0.015)	0.56 (0.025)	0.47 (0.026)	0.25 (0.025)
	Complex	Mean	0.92 (0.012)	0.72 (0.015)	0.56 (0.015)	0.35 (0.017)	0.09 (0.011)
		Min	0.82 (0.024)	0.62 (0.019)	0.49 (0.023)	0.24 (0.020)	0.01 (0.005)
		Max	0.99 (0.006)	0.82 (0.016)	0.63 (0.020)	0.46 (0.018)	0.22 (0.026)
Partial-experience	Simple	Mean	0.90 (0.018)	0.68 (0.017)	0.52 (0.021)	0.33 (0.018)	0.10 (0.012)
		Min	0.79 (0.0129)	0.57 (0.021)	0.45 (0.022)	0.24 (0.021)	0.01 (0.008)
		Max	0.98 (0.013)	0.79 (0.019)	0.59 (0.026)	0.42 (0.021)	0.24 (0.027)
	Complex	Mean	0.89 (0.015)	0.69 (0.016)	0.51 (0.023)	0.31 (0.016)	0.11 (0.013)
		Min	0.77 (0.029)	0.55 (0.023)	0.44 (0.027)	0.21 (0.018)	0.01 (0.008)
		Max	0.98 (0.009)	0.82 (0.016)	0.58 (0.028)	0.42 (0.022)	0.22 (0.024)
Experiment 2							
No-experience	Simple	Mean	0.94 (0.010)	0.75 (0.018)	0.46 (0.023)	0.27 (0.020)	0.06 (0.008)
		Min	0.85 (0.022)	0.62 (0.027)	0.34 (0.026)	0.17 (0.020)	0 (0.001)
		Max	0.99 (0.007)	0.86 (0.017)	0.56 (0.029)	0.39 (0.026)	0.16 (0.022)
	Complex	Mean	0.95 (0.008)	0.73 (0.019)	0.46 (0.025)	0.22 (0.017)	0.05 (0.009)
		Min	0.88 (0.018)	0.59 (0.028)	0.34 (0.027)	0.13 (0.015)	0 (0)
		Max	1 (0)	0.86 (0.017)	0.58 (0.029)	0.34 (0.026)	0.13 (0.021)
Full-experience	Simple	Mean	0.95 (0.008)	0.79 (0.017)	0.51 (0.026)	0.22 (0.019)	0.06 (0.010)
		Min	0.87 (0.022)	0.68 (0.026)	0.39 (0.031)	0.12 (0.018)	0.01 (0.007)
		Max	1 (0.002)	0.89 (0.015)	0.62 (0.031)	0.34 (0.026)	0.16 (0.023)
	Complex	Mean	0.95 (0.009)	0.78 (0.018)	0.50 (0.025)	0.24 (0.019)	0.51 (0.010)
		Min	0.87 (0.022)	0.65 (0.027)	0.39 (0.030)	0.15 (0.019)	0.01 (0.007)
		Max	1 (0.001)	0.89 (0.015)	0.62 (0.028)	0.36 (0.026)	0.13 (0.021)

Table B.2: The mean, minimum, and maximum randomization probabilities that correspond to each confidence statement for all treatments in the two experiments. The values in parentheses are the standard errors of the mean.

	Lottery	Randomization probabilities	Confidence statements	Probabilistic confidence
		Experiment 1		
Lower bound	Simple	2.99	2.95	
	Complex	2.84*	2.94	
Upper bound	Simple	6.61	6.30	
	Complex	6.90***	6.56**	
Range size	Simple	3.63	3.36	
	Complex	4.06***	3.62	
		Experiment 2		
Lower bound	Simple	3.16	3.03	2.63
	Complex	3.19	2.99	2.59
Upper bound	Simple	6.18	6.19	7.00
	Complex	6.38***	6.5***	7.21***
Range size	Simple	3.03	3.15	4.37
	Complex	3.19*	3.58***	4.63***

Table B.3: Comparisons of the lower bound, the upper bound, and the range size between the simple lottery and complex lottery in the no-experience treatment in the two experiments. The lower bound, the upper bound, and the range sizes are defined by randomization probabilities ($0.10 \leq \lambda \leq 0.90$), confidence statements ("Probably x", "Unsure", "Probably y") and probabilistic confidence (between "90% x, 10% y" and "10% x, 90% y"). Wilcoxon signed-rank tests were performed to test the difference between the simple lottery and the complex lottery for each measure: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	Self-reported probabilistic confidence									
	100% x 0% y	90% x 10% y	80% x 20% y	70% x 30% y	60% x 40% y	40% x 60% y	30% x 70% y	20% x 80% y	10% x 90% y	0% x 100% y
	Simple lottery, no-experience treatment									
Rand. prob.	0.98 (0.008)	0.86 (0.021)	0.82 (0.022)	0.73 (0.022)	0.58 (0.027)	0.36 (0.024)	0.23 (0.023)	0.18 (0.022)	0.09 (0.018)	0.02 (0.004)
	Complex lottery, no-experience treatment									
Rand. prob.	0.97 (0.007)	0.85 (0.024)	0.76 (0.026)	0.73 (0.024)	0.59 (0.025)	0.33 (0.024)	0.21 (0.021)	0.18 (0.024)	0.08 (0.016)	0.04 (0.012)
	Simple lottery, full-experience treatment									
Rand. prob.	0.97 (0.007)	0.92 (0.016)	0.85 (0.024)	0.76 (0.025)	0.63 (0.026)	0.37 (0.025)	0.23 (0.024)	0.16 (0.020)	0.07 (0.015)	0.03 (0.009)
	Complex lottery, full-experience treatment									
Rand. prob.	0.98 (0.007)	0.92 (0.017)	0.81 (0.025)	0.71 (0.028)	0.62 (0.027)	0.35 (0.023)	0.20 (0.024)	0.16 (0.027)	0.06 (0.015)	0.02 (0.008)

Table B.4: The mean randomization probabilities at each self-reported probabilistic confidence level in Experiment 2 for each lottery and experience treatment group. The standard errors of the mean are reported in the parentheses. We compute the mean randomization probability at each level of probabilistic confidence for each subject before taking its mean across subjects.

Lottery		Experience	Randomization probabilities	Confidence statements	Probabilistic confidence
		Experiment 1			
Simple	Lower bound	No	2.99	2.95	
		Partial	2.75	2.80	
	Upper bound	No	6.61	6.30	
		Partial	6.60	6.25	
	Range size	No	3.63	3.36	
		Partial	3.85	3.45	
Complex	Lower bound	No	2.84	2.94	
		Partial	2.74	3.03	
	Upper bound	No	6.90	6.56	
		Partial	6.90	6.52	
	Range size	No	4.06	3.62	
		Partial	4.16	3.49	
		Experiment 2			
Simple	Lower bound	No	3.16	3.03	2.63
		Full	3.44**	3.26*	2.71
	Upper bound	No	6.18	6.19	7.00
		Full	6.18	6.30	7.10
	Range size	No	3.03	3.15	4.37
		Full	2.74	3.04	4.38
Complex	Lower bound	No	3.19	2.99	2.59
		Full	3.60***	3.29**	2.73*
	Upper bound	No	6.38	6.57	7.21
		Full	6.27	6.45	7.31
	Range size	No	3.19	3.58	4.63
		Full	2.67**	3.16**	4.58

Table B.5: Comparisons of the lower bound, the upper bound, and the range size between the no-experience treatment and experience treatments by lottery type in the two experiments. The lower bound, the upper bound, and the range sizes are defined by randomization probabilities ($0.10 \leq \lambda \leq 0.90$), confidence statements ("Probably x", "Unsure", "Probably y") and probabilistic confidence (between "90% x, 10% y" and "10% x, 90% y"). Wilcoxon rank-sum tests were performed to test the difference between experience treatment and no-experience treatment for each measure: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.1 Order effects in experiment 2

We selected three different orders and randomly assigned subjects to each order: Order 1) binary choices and confidence statements → randomized choices → probabilistic confidence, Order 2) randomized choices → binary choices and confidence statements → probabilistic confidence, Order 3) probabilistic confidence → binary choices and confidence statements → randomized choices. Order 1 is similar to the task order in Experiment 1, allowing us to assess the robustness of the findings in Experiment 1. Order 2 removes the potential priming effects of the self-reported confidence measures on randomized choices. Order 3 preserves the potential of the priming effects, but allows us to look at the probabilistic confidence measure when it is elicited first.

Given our proposal to capture decision confidence with randomization probabilities, the question that is most relevant to us is whether the subjects randomized differently when they made randomization choices prior to and after they completed the self-reported confidence measures.

We find that subjects were less likely to randomize strictly ($0 < \lambda < 1$) when randomization probabilities were elicited before self-reported decision measures (Order 2). Table B.6(a) compares the proportion of decisions in which subjects randomized strictly when the randomized choices were made first versus when they were made later across orders and treatments. On average, subjects in Order 2 randomized strictly in fewer choices compared to subjects in the other two orders. This difference is statistically significant in three out of eight comparisons, but it is not significant when we aggregate across all treatments. Table B.6(b) shows that the above order effects were stronger in choices with high values ($y \geq 4$) than with low values ($y \leq 3.5$) of sure payments. Consistent with these findinds, Figure B.5 shows that the proportion of subjects with strict randomization probabilities tends to be lower at each value of y in Order 2 (the red lines) than in the other two orders, and more so for high values of sure payments.

Apart from the lower tendency to randomize, randomization probabilities corresponded to the probabilistic confidence less well at some probabilistic confidence values (e.g., 30% x

	Simple		Complex	
	No experience	Experience	No experience	Experience
Order 1	0.456	0.368	0.452	0.373
Order 2	0.345	0.331	0.369	0.328
Order 3	0.471	0.412	0.498	0.400
Wilcoxon rank-sum tests:				
Order 1 vs Order 2	$p < 0.10$	$p = 0.625$	$p = 0.217$	$p = 0.541$
Order 2 vs Order 3	$p < 0.05$	$p = 0.192$	$p < 0.05$	$p = 0.194$

Table (a)

	$y = 0$ to 3.5	$y = 4$ to 6	$y = 6.5$ to 10
Order 1	0.303	0.622	0.267
Order 2	0.272	0.514	0.197
Order 3	0.339	0.650	0.296
Wilcoxon rank-sum tests:			
Orders 2 vs 1	$p = 0.258$	$p < 0.05$	$p < 0.05$
Orders 2 vs 3	$p < 0.05$	$p < 0.01$	$p < 0.01$

Table (b)

Table B.6: Panel (a) reports the proportion of strict randomization choices ($0 < \lambda < 1$) across treatments in each order. Panel (b) the average proportions of strict randomiztion at different ranges of sure payments aggregated across treatments and lotteries in each order.

and 40% x) in Order 2 compared to the other two orders. Figure B.7 reports this result. Finally, we find that the cumulative distributions of the correlation between randomization probabilities and a self-reported decision confidence measure in Order 2 tend to be on the left of the other two orders (see Figure B.6). This implies that there were more subjects with lower correlation between randomization probabilities and self-reported decision confidence in Order 2 than in the other two orders. As a whole, the results suggest that randomized choices were affected by priming.

Despite the presence of the priming effects on randomized choices, we find support for our hypotheses when we restrict our analyses to subjects in Order 2. First, we find high correlations between randomization probabilities and self-reported confidence in Order 2, consistent with H1: the median correlation between randomization probabilities and confidence statement as well as the median correlation between randomization probabilities and probabilistic confidence in Order 2 ranges from 0.87 to 0.89 across treatments respectively. Second, subjects in Order 2 reported low decision confidence for choices around the switch-

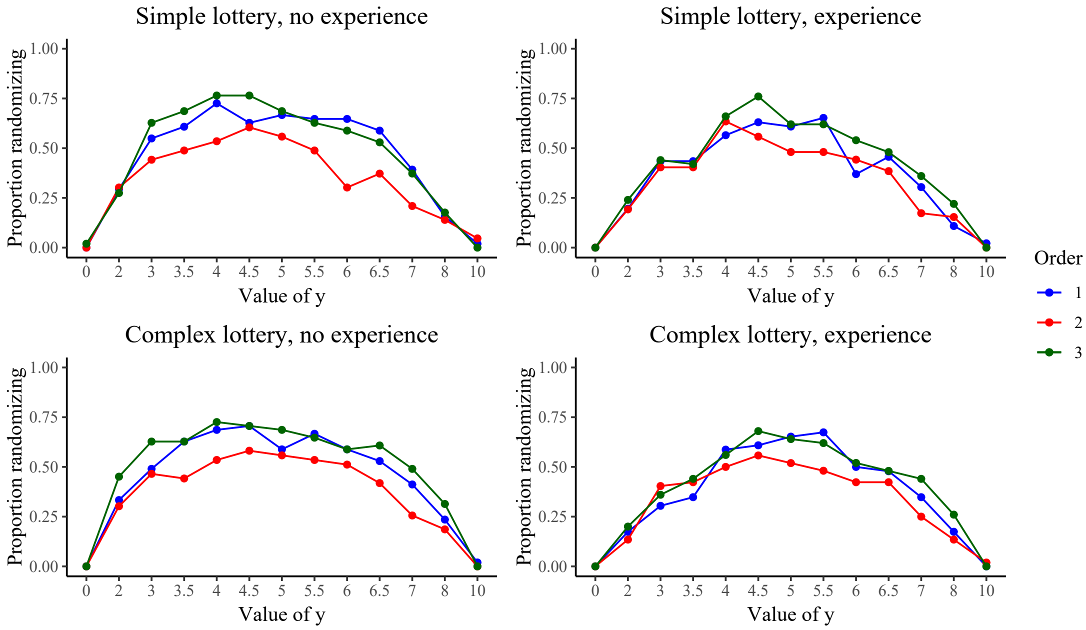


Figure B.5: The proportion of strict randomization choices ($0 < \lambda < 1$) at each value of y in the three orders.

ing range and chose randomization probabilities close to 0.5 in these decisions, supporting H2. Table B.7 shows that subjects' median randomization probability for all choices that falls within the switching range is between 0.48 and 0.50 across treatments, consistent with their median confidence statement of "Unsure," and their median probabilistic confidence which ranges from 40% x to 60% x across treatments.[12] Third, like in the full sample, we find significant treatment effects in Order 2. Table B.8 reports the range of sure payments over which subjects expressed less than full decision confidence in their decisions. The ranges based on self-reported decision measures and randomization choices suggest that subjects in Order 2 had lower decision confidence for decisions involving the complex lottery than for decisions involving the simple lottery, consistent with H3. We find similar support for H4 in Order 2 as in the full sample. Among subjects in Order 2, those in the full-experience treatment reported less than full decision confidence in smaller ranges of sure payments for decisions involving the complex lottery, but not for decisions involving

[12]The median randomization probability at the upper bound ($\bar{y}_b$) of the switching range is lower in Order 2 compared to the other two orders in most treatments. This is consistent with our earlier finding of the order effects that there were fewer strict randomization choices in Order 2, and more so for choices involving larger sure payments.

the simple lottery. In terms of randomization probabilities, these ranges are 2.30 vs 2.89 in Order 2. In comparison, they are 2.81 vs 3.09 in Order 1 and 2.90 vs 3.51 in Order 3. The difference is statistically significant in aggregate (2.67 vs 3.19, $p < 0.05$), but not when considering any order separately ($p > 0.10$).

Taken together, priming effects could have strengthened some of our aggregate findings. However, the findings that subjects in Order 2 randomized in ways that were broadly consistent with our hypotheses suggest that randomization probabilities and self-reported decision confidence measures are likely to share common psychological foundations.

It is worth noting that the aforementioned order effects do not automatically imply that randomization probabilities are a poorer proxy for decision confidence than self-reported decision confidence. Confidence statements and probabilistic confidence are also noisy proxies of decision confidence, and there is no objective criterion for the "right" amount of strict randomization. For example, it is possible that subjects may have not randomized too little in Order 2, but that they have randomized excessively in the other two orders due to the priming effects. Since there is no obvious benchmark to compare decision confidence measures, and decision confidence is not directly observable, we consider the value of decision confidence on the basis of its correspondence with actual choices. When decision confidence corresponds perfectly with choices, having decision confidence of $p\%$ for x would imply that x is chosen $p\%$ of time. Figure B.8 shows, across subjects and choices, the proportion of choices in which x was chosen based on subjects' binary choices by randomization probabilities as well as the probabilistic confidence in the three orders.[13] On average, both measures of decision confidence closely trace the proportions of x chosen in binary choices. Importantly, randomization probabilities exhibit a closer correspondence to binary choices than probabilistic confidence in all three orders. This alignment is particularly evident in Order 2, with randomization probabilities exhibiting a significantly closer match to these proportions than probabilistic confidence at nine levels of randomization probabilities/probabilistic confidence versus five levels in Order 1 and six levels in Order 3.

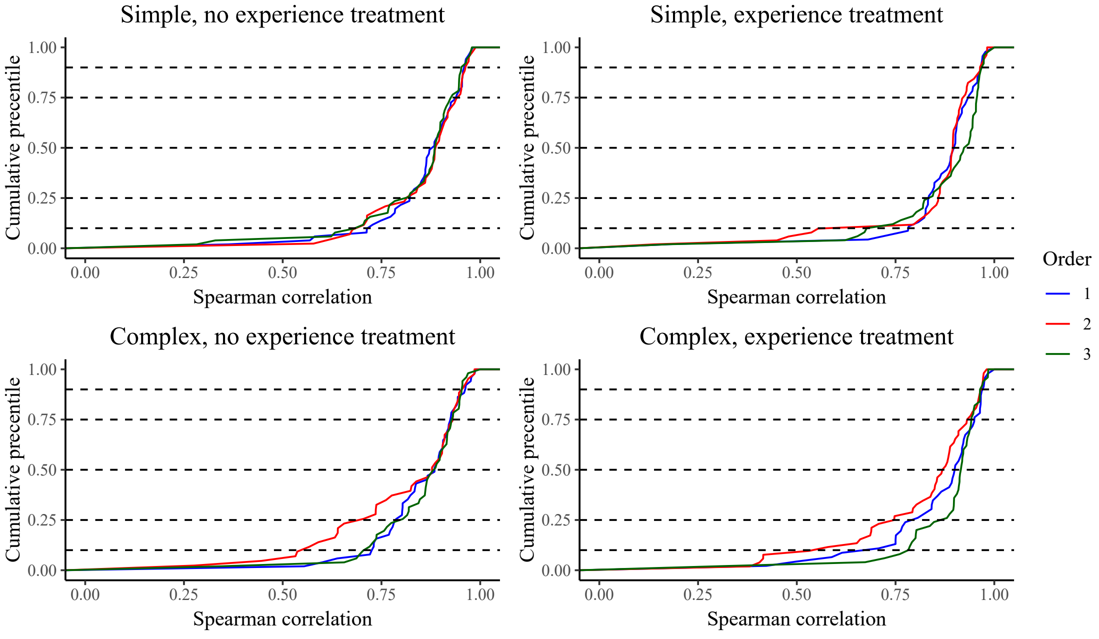
[13]We did not consider confidence statements here, because they were elicited on the same decision screen as binary choices.

Order	Simple			Complex		
			Confidence statements			
			No experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
Order 2	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
Order 3	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
			Experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
Order 2	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
Order 3	Probably x	Probably y	Unsure	Probably x	Probably y	Unsure
			Probabilistic confidence			
			No experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	60% x	40% x	40% x	60% x	40% x	60% x
Order 2	70% x	40% x	60% x	60% x	40% x	40% x
Order 3	60% x	40% x	60% x	60% x	30% x	40% x
			Experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	60% x	40% x	50% x	60% x	40% x	40% x
Order 2	70% x	40% x	60% x	60% x	40% x	40% x
Order 3	60% x	40% x	60% x	60% x	40% x	60% x
			Randomization probabilities			
			No experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	0.65	0.46	0.50	0.68	0.40	0.50
Order 2	0.68	0.27	0.50	0.63	0.30	0.48
Order 3	0.60	0.46	0.50	0.58	0.21	0.40
			Experience			
Order	Simple			Complex		
	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$	$\underline{y_b}$	$\bar{y_b}$	$[\underline{y_b}, \bar{y_b}]$
Order 1	0.68	0.30	0.55	0.60	0.33	0.50
Order 2	0.63	0.23	0.50	0.66	0.25	0.49
Order 3	0.70	0.43	0.51	0.61	0.40	0.50

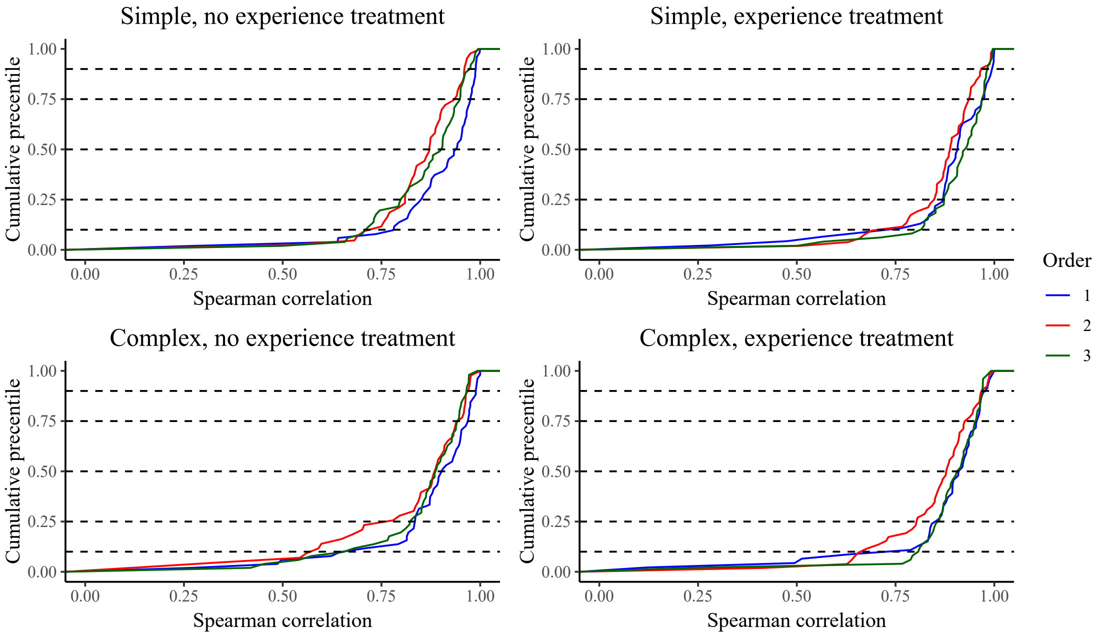
Table B.7: Median behavior around the switching choices ($\underline{y_b}$ and $\bar{y_b}$) and within the switching range ($[\underline{y_b}, \bar{y_b}]$) aggregated across lotteries and treatments for each order.

	Lottery	Treatment	Combined	Order 1	Order 2	Order 3
Confidence Statements	Simple	No experience	3.15	2.96	2.77	3.68
		Experience	3.04	2.84	2.92	3.35
	Complex	No experience	3.58***	3.25*	3.56***	3.93
		Experience	3.16_{**}	3.23	3.03_{*}	3.23_{*}
Probabilistic Confidence	Simple	No experience	4.37	4.28	4.06	4.72
		Experience	4.38	4.24	4.40	4.50
	Complex	No experience	4.63***	4.34	4.61***	4.93*
		Experience	4.58	4.53	4.50	4.70
Randomization Probabilities	Simple	No experience	3.03	3.17	2.67	3.14
		Experience	2.74	3.00	2.34	2.93
	Complex	No experience	3.19*	3.09	2.89*	3.51**
		Experience	2.67_{**}	2.81	2.30	2.90

Table B.8: The mean size of the range of sure payments over which subjects express that they are not fully confident about their decision based on each of the confidence measures, by the lottery and experience treatments in aggregate and in each order separately (randomization probabilities ($0.10 \leq \lambda \leq 0.90$), confidence statements ("Probably x", "Unsure", "Probably y") and probabilistic confidence (between "90% x, 10% y" and "10% x, 90% y")). Stars in the upper right corners of a cell denote statistical significance of Wilcoxon signed-rank tests between the simple lottery and the complex lottery, while stars in the lower right corner denote statistical significance of Wilcoxon rank-sum tests between the no experience and experience treatment: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Confidence statements



(b) Probabilistic confidence

Figure B.6: ECDF for the correlation with randomisation probabilities in each order across treatments.

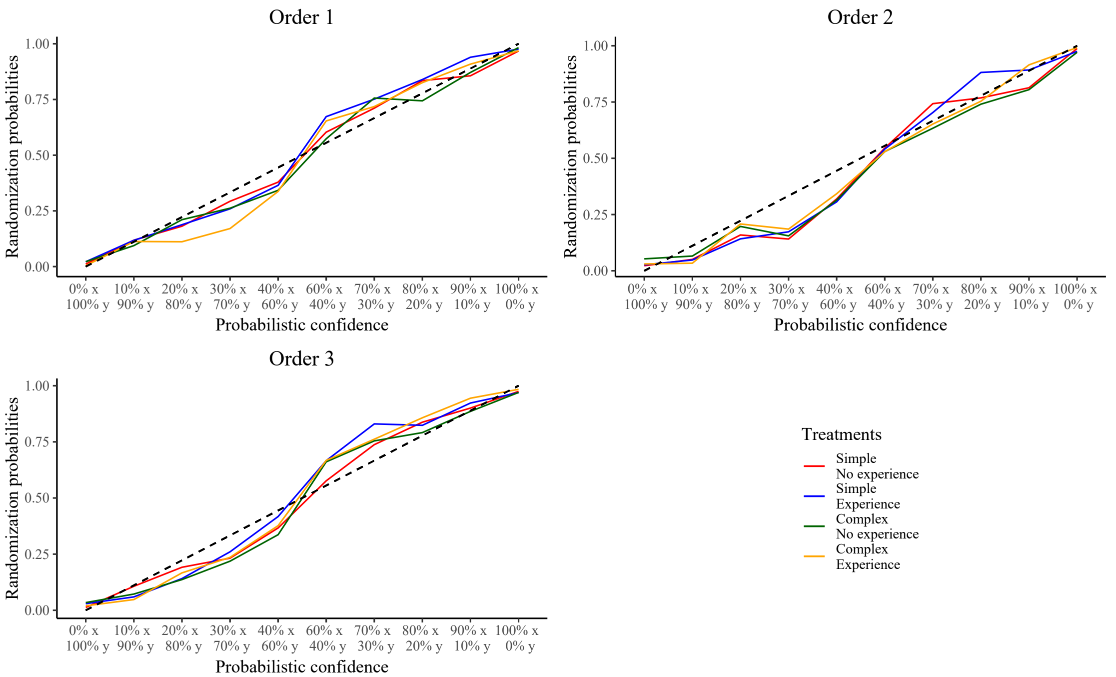


Figure B.7: The mean randomization probabilities at each self-reported probabilistic confidence level in Experiment 2 for each lottery and experience treatment in each order separately.

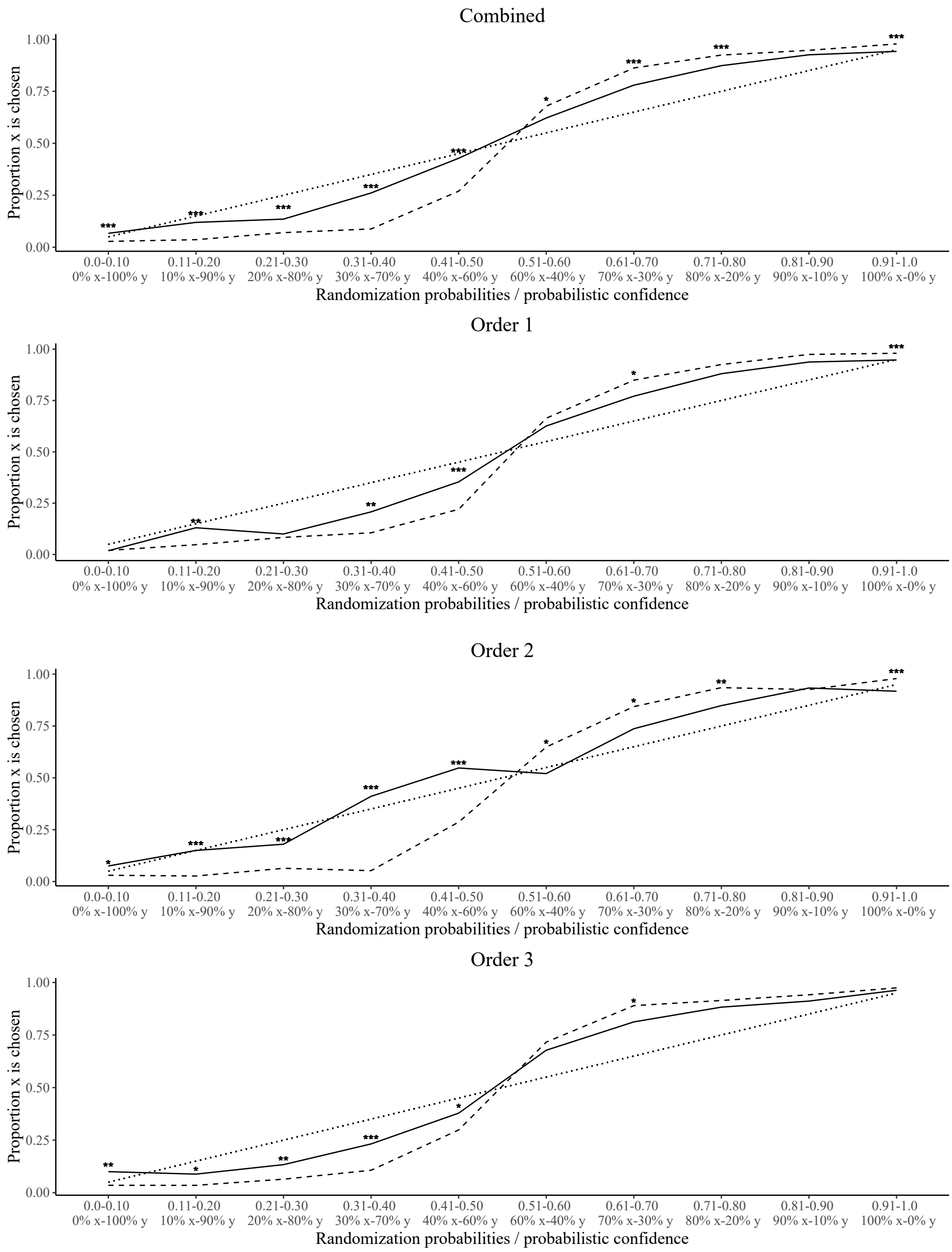


Figure B.8: Correspondence of randomization probabilities (solid line) and probabilistic confidence (dashed line) with the proportion of binary choices in which x is chosen (y-axis) across lotteries and treatments for each decision order. The dotted line represents a 45-degree line. Fisher's exact tests were performed to test the difference in choice proportions between the randomization probabilities and probabilistic confidence: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.2 Alternative interpretations of randomization probabilities

In the main text, we have shown that indifference and random errors cannot be the driving force behind subjects' randomization behavior. We elaborate why utility difference alone cannot explain randomization here.

Butler et al. (2014) call utility difference the strength of preferences: "the relative degree of difference between the two options as perceived by the decision maker" (Butler et al., 2014, p.538). For example, we can write this explicitly as a Fechnerian utility model $p = \phi\left[U(L) - u(y)\right]$, where $\phi : R \Rightarrow [0, 1]$ is a cumulative distribution function with $\phi(0) = 0.5$ (Luce and Suppes, 1965, p.334). The lower bound and the upper bound of randomization are then $u(\underline{y}) = U(L) - \phi^{-1}(0.90)$ and $u(\bar{y}) = U(L) - \phi^{-1}(0.10)$. If randomization probabilities depend only on utility differences, when $U(L)$ increases: (1) the randomization probability should increase for each value of sure payment; and (2) the lower bound and the upper bound of randomization should increase equally. Our results clearly reject these two predictions. Subjects' randomization probabilities did not shift horizontally but were instead compressed towards 0.5 when they faced the complex lottery compared to the simple lottery, and were stretched away from 0.5 in the full-experience treatment compared to the no-experience treatment. In addition, Table B.3 shows that while $\bar{y}$ for the complex lottery was significantly higher than $\bar{y}$ for the simple lottery in both experiments (Experiment 1: 6.90 vs 6.61, Wilcoxon signed-rank test, $p < 0.01$; Experiment 2: 6.38 vs 6.18, Wilcoxon signed-rank test $p < 0.01$), $\underline{y}$ for the complex lottery was significantly lower than $\underline{y}$ for the simple lottery in Experiment 1 at 10% significance level (2.84 vs 2.99, Wilcoxon signed-rank test $p < 0.10$), but was not significantly different in Experiment 2 (3.19 vs 3.16, Wilcoxon signed-rank test $p = 0.646$).

Importantly, although the difference in the mean valuation of the lottery with or without experience was similar to that of the complex lottery versus the simple lottery, Table B.5 shows that $\underline{y}$ were significantly lower for subjects in the full-experience treatment than in the no experience treatment (Simple lottery: 3.16 vs 3.44, Wilcoxon rank-sum test $p < 0.05$; Complex lottery: 3.19 vs 3.60, Wilcoxon rank-sum test $p < 0.01$) but not $\bar{y}$ (Simple lottery: 6.18 vs 6.18, Wilcoxon rank-sum test $p = 0.789$; Complex lottery: 6.27

Randomization	The number of subjects who chose randomization			
Interval	0 times	1 time	2 times or more	3 times or more
Experiment 1: Simple lottery, no-experience				
$0 < \lambda < 1$	2	6	97	95
$0.10 \leq \lambda \leq 0.90$	3	6	96	93
$0.40 \leq \lambda \leq 0.60$	22	26	57	35
Experiment 1: Complex lottery, no-experience				
$0 < \lambda < 1$	4	1	100	98
$0.10 \leq \lambda \leq 0.90$	4	2	99	98
$0.40 \leq \lambda \leq 0.60$	21	15	69	46
Experiment 1: Simple lottery, partial-experience				
$0 < \lambda < 1$	6	1	93	89
$0.10 \leq \lambda \leq 0.90$	6	1	93	88
$0.40 \leq \lambda \leq 0.60$	13	24	63	38
Experiment 1: Complex lottery, partial-experience				
$0 < \lambda < 1$	3	5	92	89
$0.10 \leq \lambda \leq 0.90$	3	7	90	89
$0.40 \leq \lambda \leq 0.60$	14	19	67	46
Experiment 2: Simple lottery, no-experience				
$0 < \lambda < 1$	25	8	112	106
$0.10 \leq \lambda \leq 0.90$	26	7	112	105
$0.40 \leq \lambda \leq 0.60$	42	24	79	44
Experiment 2: Complex lottery, no-experience				
$0 < \lambda < 1$	26	6	113	100
$0.10 \leq \lambda \leq 0.90$	26	7	112	98
$0.40 \leq \lambda \leq 0.60$	37	38	70	42
Experiment 2: Simple lottery, full-experience				
$0 < \lambda < 1$	32	11	105	98
$0.10 \leq \lambda \leq 0.90$	34	11	103	96
$0.40 \leq \lambda \leq 0.60$	55	36	57	32
Experiment 2: Complex lottery, full-experience				
$0 < \lambda < 1$	35	11	102	91
$0.10 \leq \lambda \leq 0.90$	35	12	101	90
$0.40 \leq \lambda \leq 0.60$	56	25	67	38

Table B.9: The distribution of subjects who chose $0 < \lambda < 1$, $0.10 \leq \lambda \leq 0.90$, and $0.40 \leq \lambda < 0.60$ zero times, one time, two times or more, and three times or more across treatments in the two experiments.

Probabilistic confidence associated with each confidence statement						
Levels	Median	10th	30th	70th	90th	SD
Surely (min)	85%	70%	80%	90%	100%	16.31%
Probably (max)	80%	70%	80%	90%	99%	11.10%
Probably (min)	55%	25%	50%	60%	65%	16.59%
Unsure (max)	54%	25%	50%	60%	64%	17.58%
Unsure (min)	35%	0%	0%	40%	50%	21.12%
Confidence statements associated with each probabilistic confidence level						
Levels	Median	10th	30th	70th	90th	SD
100%	Surely x	Surely x	Surely x	Surely x	Surely x	0.17
90%	Surely x	Probably x	Surely x	Surely x	Surely x	0.54
80%	Probably x	Probably x	Probably x	Surely x	Surely x	0.54
70%	Probably x	Probably x	Probably x	Probably x	Probably x	0.34
60%	Unsure	Unsure	Unsure	Surely x	Surely x	0.54

Table B.10: The median, 10th, 30th, 70th, 90th percentile, and standard deviation of probabilistic confidence associated with each confidence statement and the median, 10th, 30th, 70th, 90th percentile, and standard deviation of confidence statements associated with each probabilistic confidence level. Consistent with Result 1, we code confidence statements of surely x, probably x, unsure, probably y, and surely y as 5, 4, 3, 2, and 1 respectively. Standard deviations are calculated accordingly.

vs 6.38, Wilcoxon rank-sum test $p = 0.369$). Further, randomization probabilities were larger at low sure payments but smaller at high sure payments when we compare the full-experience treatment with the no-experience treatment. These results highlight the central role of preference uncertainty beyond utility difference in affecting decision confidence.

One may perceive that self-reported decision confidence measures are easier to interpret than randomization probabilities because they ask about decision confidence explicitly. We show that self-reported decision confidence measures can be just as difficult to interpret by analyzing how subjects associate the two self-reported decision measures in the post-experiment questionnaire of Experiment 2. We asked subjects which confidence statement best described their probabilistic confidence $p\%$ in choosing x and 100-$p\%$ in choosing y for values $p = 60, 70, 80, 90, 100$. In a separate session, we asked subjects to state the minimum level of probabilistic confidence for "Surely", and the minimum and maximum levels of probabilistic confidence for "Probably" and "Unsure" on a scale from 0% to 100%.

Table B.10 summarizes the subjects' responses to the two questions. The top panel shows the range of probabilistic confidence levels associated with each confidence statement. Although the first column shows that the median probabilistic confidence thresholds are

Probabilistic confidence associated with each confidence statement						
For subjects: Unsure (max) ≥ 50% and Unsure(min) > 0%						
Levels	Median	10th	30th	70th	90th	S.D
Surely (min)	85%	75%	80%	90%	100%	14.80%
Probably (max)	85%	75%	80%	90%	99%	9.13%
Probably (min)	60%	41%	55%	60%	70%	13.26%
Unsure (max)	60%	50%	55%	60%	65%	9.62%
Unsure (min)	40%	30%	40%	45%	50%	10.46%
For subjects: Unsure (max) < 50%						
Levels	Median	10th	30th	70th	90th	S.D
Surely (min)	80%	50.3%	75%	85.5%	99%	20.22%
Probably (max)	80%	60%	75%	84%	95%	14.01%
Probably (min)	40%	20%	30%	50%	60%	17.57%
Unsure (max)	30%	10%	20%	35.4%	40%	12.94%
Unsure (min)	0%	0%	0%	0%	10%	8.46%

Table B.11: The median, 10th, 30th, 70th, 90th percentile, and standard deviation of probabilistic confidence associated with each confidence statement for subjects who fit the criteria specified in the table.

well-ordered (the median maximum probabilistic confidence of a lower ordered statement was always smaller than the median minimum probabilistic confidence of a higher ordered statement), the standard deviations reported in the last column as well as minimum and maximum probabilistic confidence assigned to each confidence statement at different percentile levels show the presence of substantial heterogeneity in the probabilistic confidence associated with each confidence statement.

Further, we find two different interpretations of the confidence statement “Unsure”. A large group of subjects (n=172) reported probabilistic confidence higher than 50% as the maximum of “Unsure” and higher than 0% as the minimum of “Unsure,” while another group of subjects (n=84) reported a probabilistic confidence level lower than 50% as the maximum of “Unsure” and close 0% as the minimum of “Unsure.” Table B.11 shows how different these two groups were in their associations of probabilistic confidence and other confidence statements. For example, the maximum level of probabilistic confidence for the statement “Probably” ranges from 75% to 99% among subjects who reported a probabilistic confidence level higher than 50% as the maximum of “Unsure,” and it ranges from 60% to 95% among subjects who reported a probabilistic confidence level lower than 50% as the maximum of “Unsure.”

C Online appendix

C.1 Reasons to randomize

At the end of the session on randomized choices in Experiment 2, we asked the subjects who had chosen to randomize at least once in the post-experiment questionnaire, what their reasons for randomizing were. Of the 120 subjects who provided an answer to this question, 22% stated that they randomized because they were unsure about their choice or found it difficult to compare the two options. Here are a few examples:

- "Because I was not completely sure whether I wanted to choose A or B.'
- "I was not sure exactly what the consequences of my decision was going to be and I was not 100% confident in choosing either A or B."
- "Its difficult to make a decision for sure, so a combination feels more safe."

Another group of subjects (22.5%) randomized for reasons related to hedging. Here are a few examples:

- "Even though the certain option was less valued, certainty is nice and preferred over risky options. Therefore, I chose to combine them some of the time."
- "To hedge my bets when the expected gains of A and B were similar, gaining a small chance for big gains or loses in option A, adding some suspense."
- "For example when I preferred A but B felt a little safer so I thought it wouldn't hurt adding a bit more security since a B amount for sure isn't bad."

Around 18% stated that they chose to randomize when the sure payment amount was close to the expected value of the lottery but did not explain why randomizing is better. In contrast, most of the subjects who did not to randomize at all stated that they did not randomize because they did not want to pay the cost of 0.10 euro for randomizing and/or that they made their choices solely based on the computation of the expected value of the lottery.

C.2 Results for the loss lottery and the mixed lottery in Experiment 1

Treatment	Lottery		Surely x	Probably x	Unsure	Probably y	Surely y
No-experience	Loss	Mean	0.94 (0.008)	0.73 (0.015)	0.50 (0.021)	0.29 (0.019)	0.07 (0.011)
		Min	0.84 (0.022)	0.61 (0.023)	0.40 (0.025)	0.19 (0.019)	0 (0.002)
		Max	1 (0.003)	0.85 (0.016)	0.60 (0.025)	0.41 (0.026)	0.15 (0.024)
	Mixed	Mean	0.90 (0.015)	0.71 (0.015)	0.53 (0.025)	0.32 (0.020)	0.10 (0.016)
		Min	0.75 (0.035)	0.56 (0.025)	0.44 (0.029)	0.17 (0.018)	0.02 (0.011)
		Max	0.99 (0.003)	0.85 (0.016)	0.61 (0.030)	0.46 (0.031)	0.23 (0.031)
Partial-experience	Loss	Mean	0.91 (0.014)	0.70 (0.018)	0.55 (0.019)	0.34 (0.019)	0.07 (0.011)
		Min	0.81 (0.026)	0.55 (0.026)	0.42 (0.027)	0.23 (0.021)	0.01 (0.006)
		Max	0.98 (0.012)	0.82 (0.016)	0.66 (0.022)	0.46 (0.025)	0.14 (0.022)
	Mixed	Mean	0.91 (0.014)	0.71 (0.019)	0.57 (0.025)	0.33 (0.023)	0.11 (0.014)
		Min	0.78 (0.031)	0.57 (0.027)	0.47 (0.032)	0.23 (0.026)	0.01 (0.006)
		Max	0.99 (0.006)	0.83 (0.018)	0.67 (0.026)	0.46 (0.030)	0.28 (0.035)

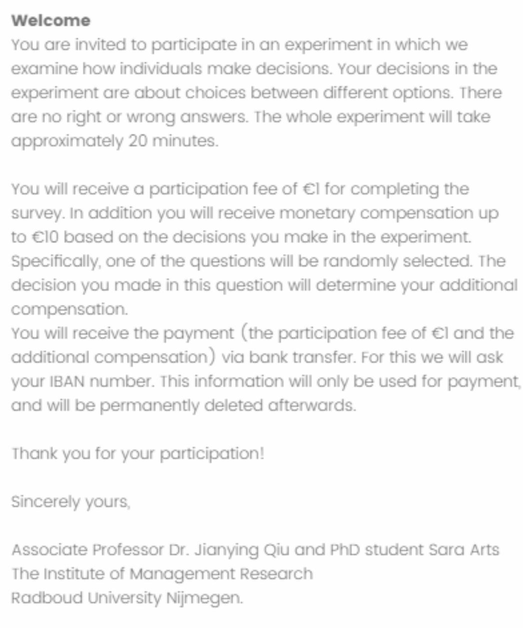
Table C.1: The mean, minimum, and maximum randomization probabilities that correspond to each confidence statement for the loss and mixed lottery in both treatments in Experiment 1. The values in parentheses are the standard errors of the mean.

Treatment		Correlation between randomization probabilities and confidence statements	
		The loss lottery	The mixed lottery
No experience	10th percentile	0.67	0.35
	25th percentile	0.81	0.64
	median	0.90	0.85
	75th percentile	0.95	0.91
	90th percentile	0.97	0.96
Experience	10th percentile	0.65	0.47
	25th percentile	0.80	0.66
	median	0.87	0.81
	75th percentile	0.94	0.92
	90th percentile	0.97	0.95

Table C.2: Nonparametric Spearman correlation at the 10th percentile, 25th percentile, median, 75th percentile, and 90th percentile for the loss lottery and mixed lottery in both treatments in Experiment 1.

C.3 Experimental materials

Experiment 1



Welcome
You are invited to participate in an experiment in which we examine how individuals make decisions. Your decisions in the experiment are about choices between different options. There are no right or wrong answers. The whole experiment will take approximately 20 minutes.

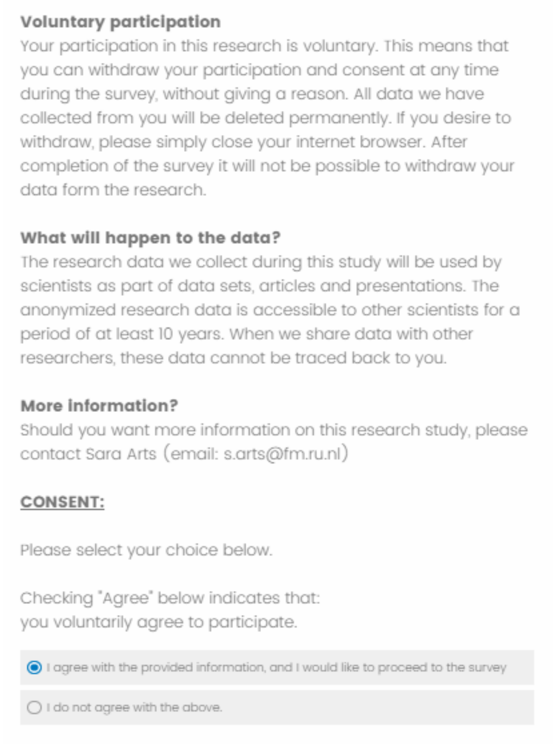
You will receive a participation fee of €1 for completing the survey. In addition you will receive monetary compensation up to €10 based on the decisions you make in the experiment. Specifically, one of the questions will be randomly selected. The decision you made in this question will determine your additional compensation.
You will receive the payment (the participation fee of €1 and the additional compensation) via bank transfer. For this we will ask your IBAN number. This information will only be used for payment, and will be permanently deleted afterwards.

Thank you for your participation!

Sincerely yours,

Associate Professor Dr. Jianying Qiu and PhD student Sara Arts
The Institute of Management Research
Radboud University Nijmegen.

(a)



Voluntary participation
Your participation in this research is voluntary. This means that you can withdraw your participation and consent at any time during the survey, without giving a reason. All data we have collected from you will be deleted permanently. If you desire to withdraw, please simply close your internet browser. After completion of the survey it will not be possible to withdraw your data form the research.

What will happen to the data?
The research data we collect during this study will be used by scientists as part of data sets, articles and presentations. The anonymized research data is accessible to other scientists for a period of at least 10 years. When we share data with other researchers, these data cannot be traced back to you.

More information?
Should you want more information on this research study, please contact Sara Arts (email: s.arts@fm.ru.nl)

CONSENT:

Please select your choice below.

Checking "Agree" below indicates that:
you voluntarily agree to participate.

I agree with the provided information, and I would like to proceed to the survey

I do not agree with the above.

(b)

Figure C.1: Welcome screen (a) and informed consent (b) of the experiment.

The following questions are about the options below:

Option A:
Gain €9,75 with a chance of **20%**,
gain €7,50 with a chance of **30%**,
gain €2,50 with a chance of **30%**, and
gain €0,25 with a chance of **20%.**

Option B:
Gain a **sure** amount (which varies across questions).

You will be asked to choose between the two options, and to describe how confident you feel about your choice.

(a)

The following questions are about the options below:

Option A:
Gain €9 with a chance of **50%,** and
gain €1 with a chance of **50%**.

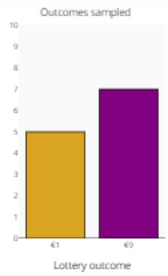
Option B:
Gain a **sure** amount (which varies across questions).

You will be asked to choose between the two options, and to describe how confident you feel about your choice.

Before you are asked to make the decisions we want to give you the opportunity to **experience the different outcomes of option A**. For this, you can click the button below. Each time you click the button a possible outcome of option A will be shown. You will get to sample **20** outcomes.

The outcomes that you obtain by clicking the button do not influence your payoff but are only presented to make you experience the possible outcomes.

To keep track of the sampled outcomes, they will be presented in a bargraph.

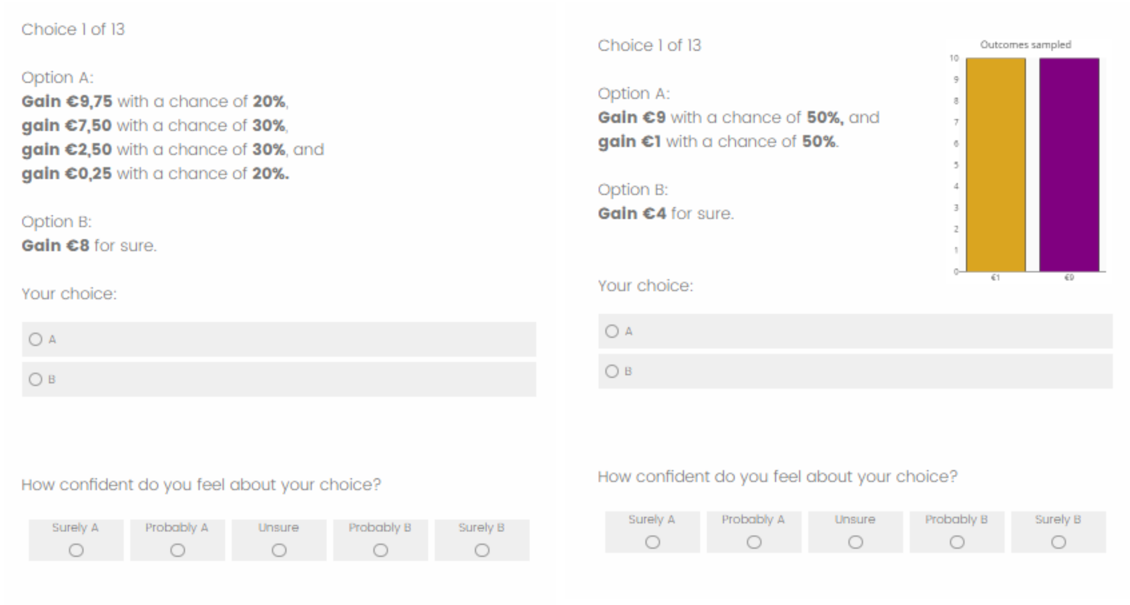


€1

Sample

(b)

Figure C.2: The introduction of the binary choices and confidence statements for the complex lottery in the no-experience treatment (a) and the simple lottery in the partial-experience treatment (b).



(a) (b)

Figure C.3: Examples of the decision screens for the binary choices and confidence statements for the complex lottery in the no-experience treatment (a) and the simple lottery in the partial-experience treatment (b).

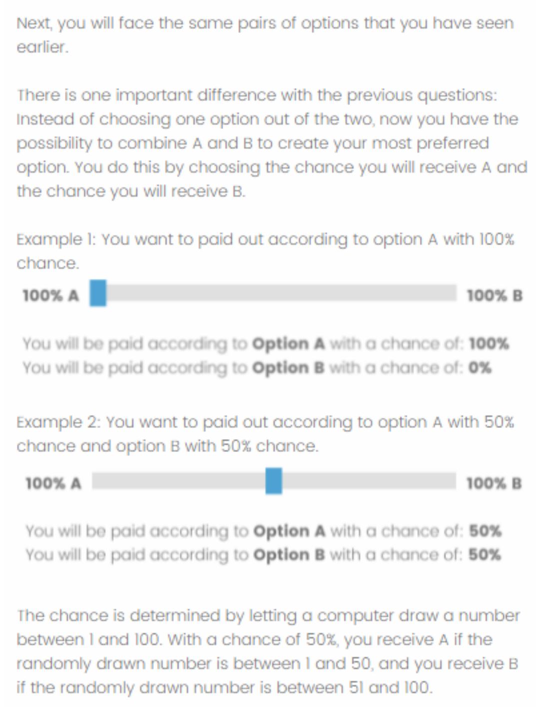
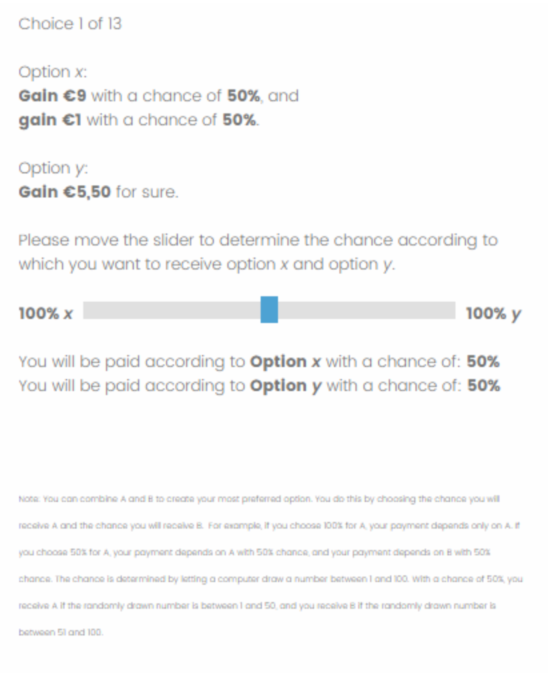
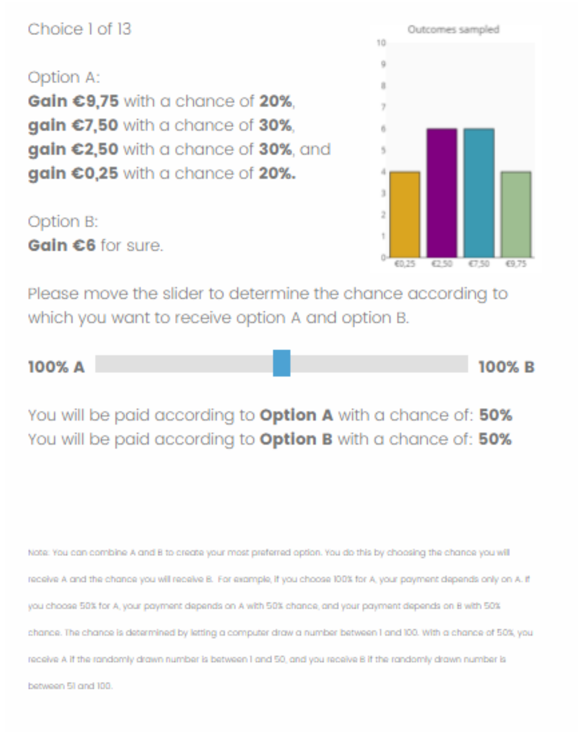


Figure C.4: Explanation of the randomized choices.

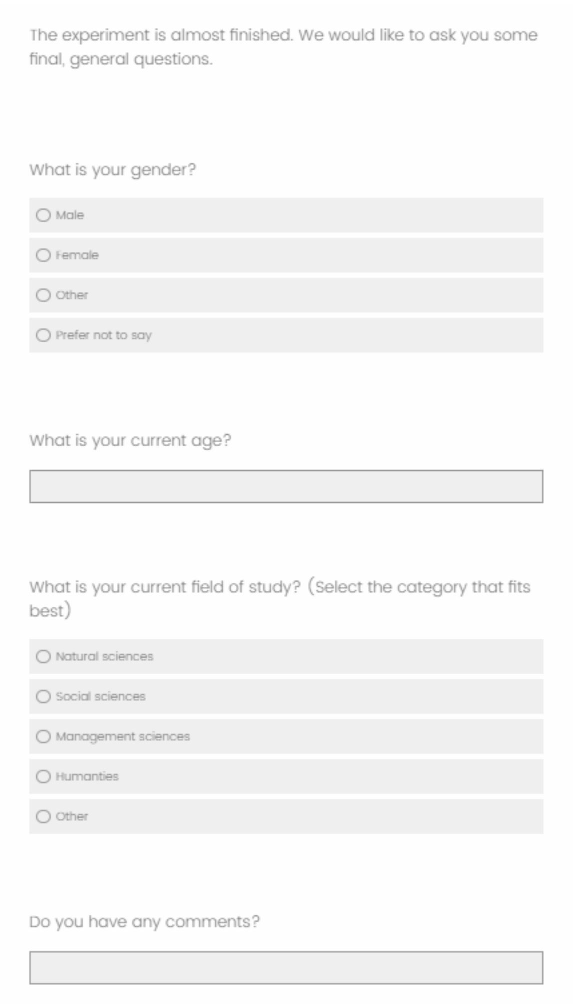


(a)



(b)

Figure C.5: Examples of the decision screens for the randomized choices for the simple lottery in the no-experience treatment (a) and the complex lottery in the partial-experience treatment (b).



The experiment is almost finished. We would like to ask you some final, general questions.

What is your gender?

- Male
- Female
- Other
- Prefer not to say

What is your current age?

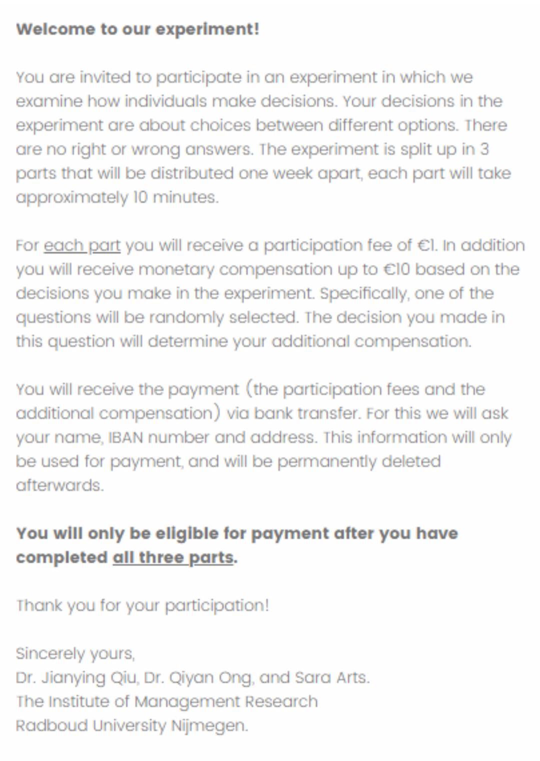
What is your current field of study? (Select the category that fits best)

- Natural sciences
- Social sciences
- Management sciences
- Humanties
- Other

Do you have any comments?

Figure C.6: Demographic questions asked at the end of the experiment.

Experiment 2



Welcome to our experiment!

You are invited to participate in an experiment in which we examine how individuals make decisions. Your decisions in the experiment are about choices between different options. There are no right or wrong answers. The experiment is split up in 3 parts that will be distributed one week apart, each part will take approximately 10 minutes.

For each part you will receive a participation fee of €1. In addition you will receive monetary compensation up to €10 based on the decisions you make in the experiment. Specifically, one of the questions will be randomly selected. The decision you made in this question will determine your additional compensation.

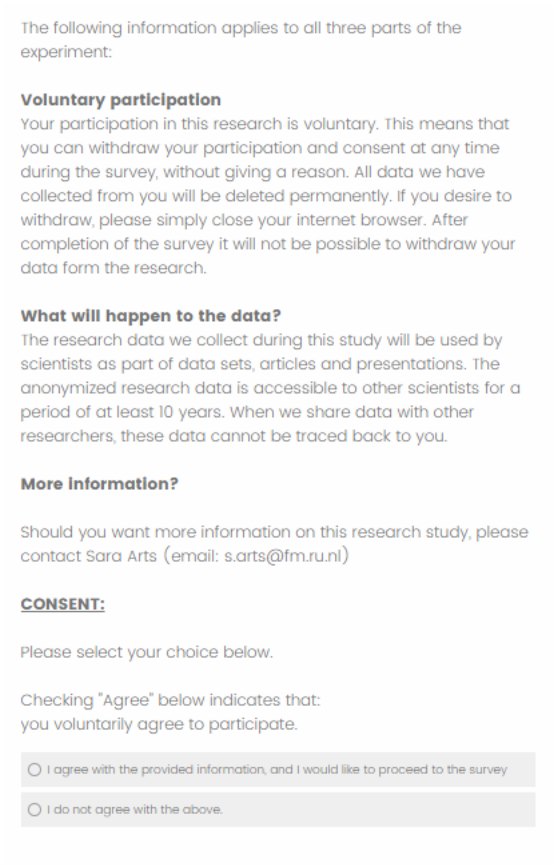
You will receive the payment (the participation fees and the additional compensation) via bank transfer. For this we will ask your name, IBAN number and address. This information will only be used for payment, and will be permanently deleted afterwards.

You will only be eligible for payment after you have completed all three parts.

Thank you for your participation!

Sincerely yours,
Dr. Jianying Qiu, Dr. Qiyan Ong, and Sara Arts.
The Institute of Management Research
Radboud University Nijmegen.

(a)



The following information applies to all three parts of the experiment:

Voluntary participation
Your participation in this research is voluntary. This means that you can withdraw your participation and consent at any time during the survey, without giving a reason. All data we have collected from you will be deleted permanently. If you desire to withdraw, please simply close your internet browser. After completion of the survey it will not be possible to withdraw your data form the research.

What will happen to the data?
The research data we collect during this study will be used by scientists as part of data sets, articles and presentations. The anonymized research data is accessible to other scientists for a period of at least 10 years. When we share data with other researchers, these data cannot be traced back to you.

More information?

Should you want more information on this research study, please contact Sara Arts (email: s.arts@fm.ru.nl)

CONSENT:

Please select your choice below.

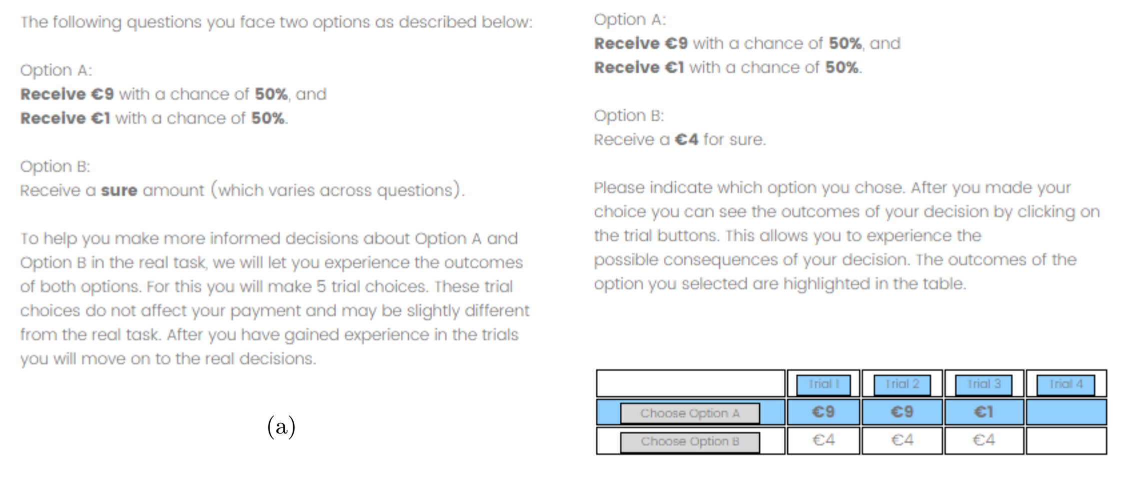
Checking "Agree" below indicates that:
you voluntarily agree to participate.

○ I agree with the provided information, and I would like to proceed to the survey

○ I do not agree with the above.

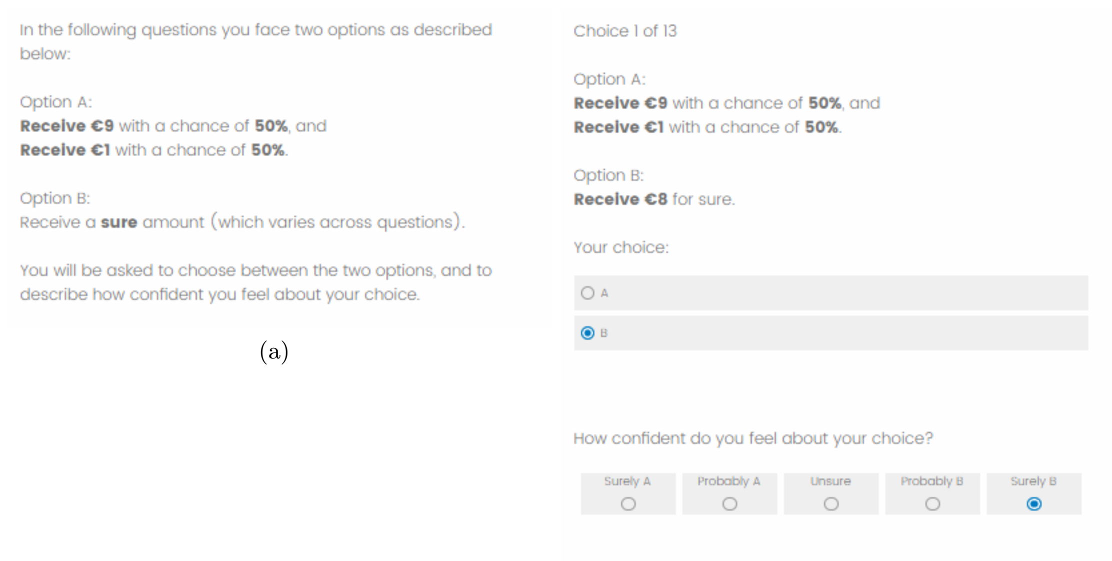
(b)

Figure C.7: Welcome screen (a) and informed consent (b) of the experiment.



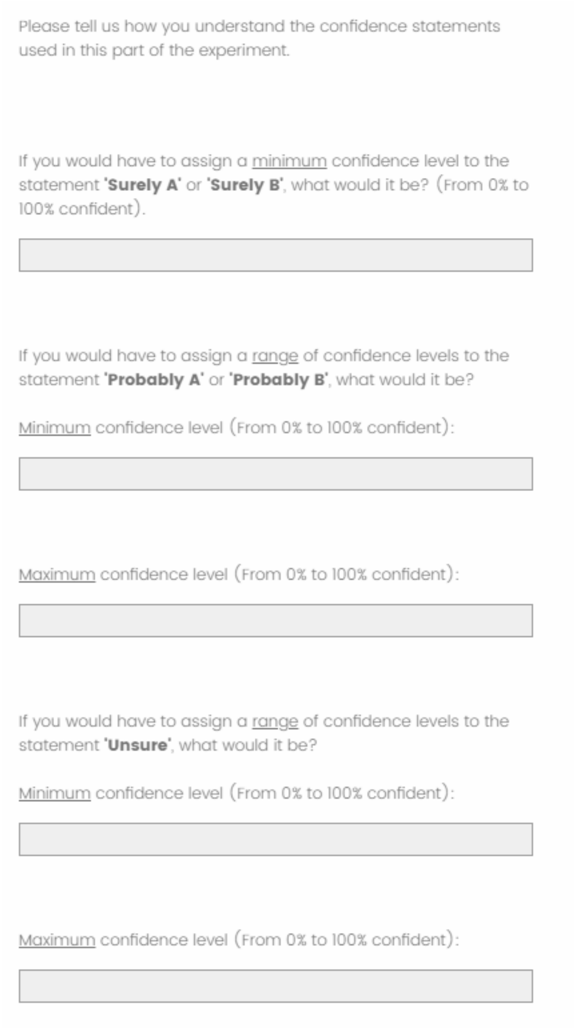
(a) (b)

Figure C.8: The introduction (a) and an example of the hypothetical decision screens (b) of the full-experience treatment.



(a) (b)

Figure C.9: The introduction (a) and an example of the decision screens (b) of binary choices and confidence statements for the simple lottery.



Please tell us how you understand the confidence statements used in this part of the experiment.

If you would have to assign a minimum confidence level to the statement **'Surely A'** or **'Surely B'**, what would it be? (From 0% to 100% confident).

If you would have to assign a range of confidence levels to the statement **'Probably A'** or **'Probably B'**, what would it be?

Minimum confidence level (From 0% to 100% confident):

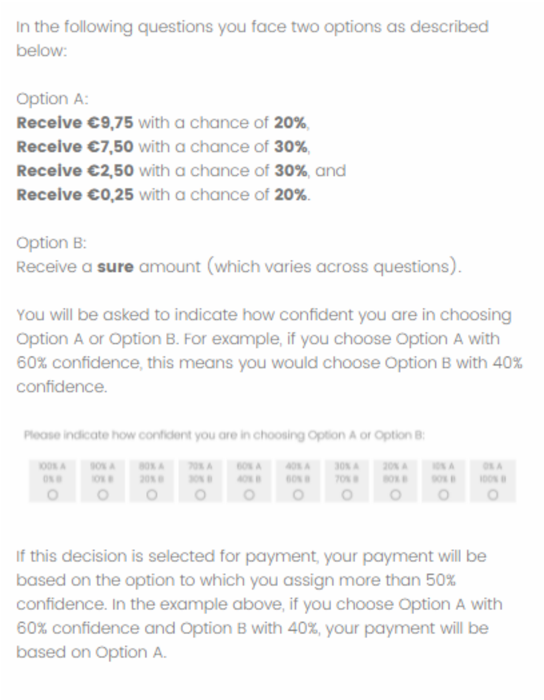
Maximum confidence level (From 0% to 100% confident):

If you would have to assign a range of confidence levels to the statement **'Unsure'**, what would it be?

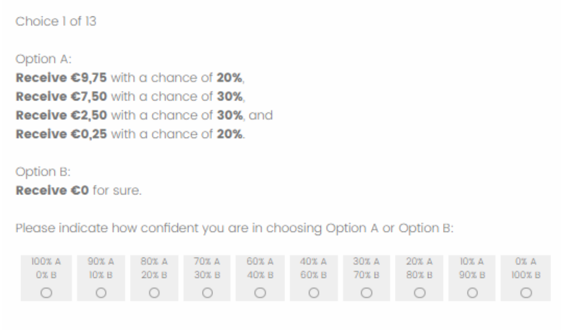
Minimum confidence level (From 0% to 100% confident):

Maximum confidence level (From 0% to 100% confident):

Figure C.10: post-experiment questionnaire after the binary choices and confidence statements.

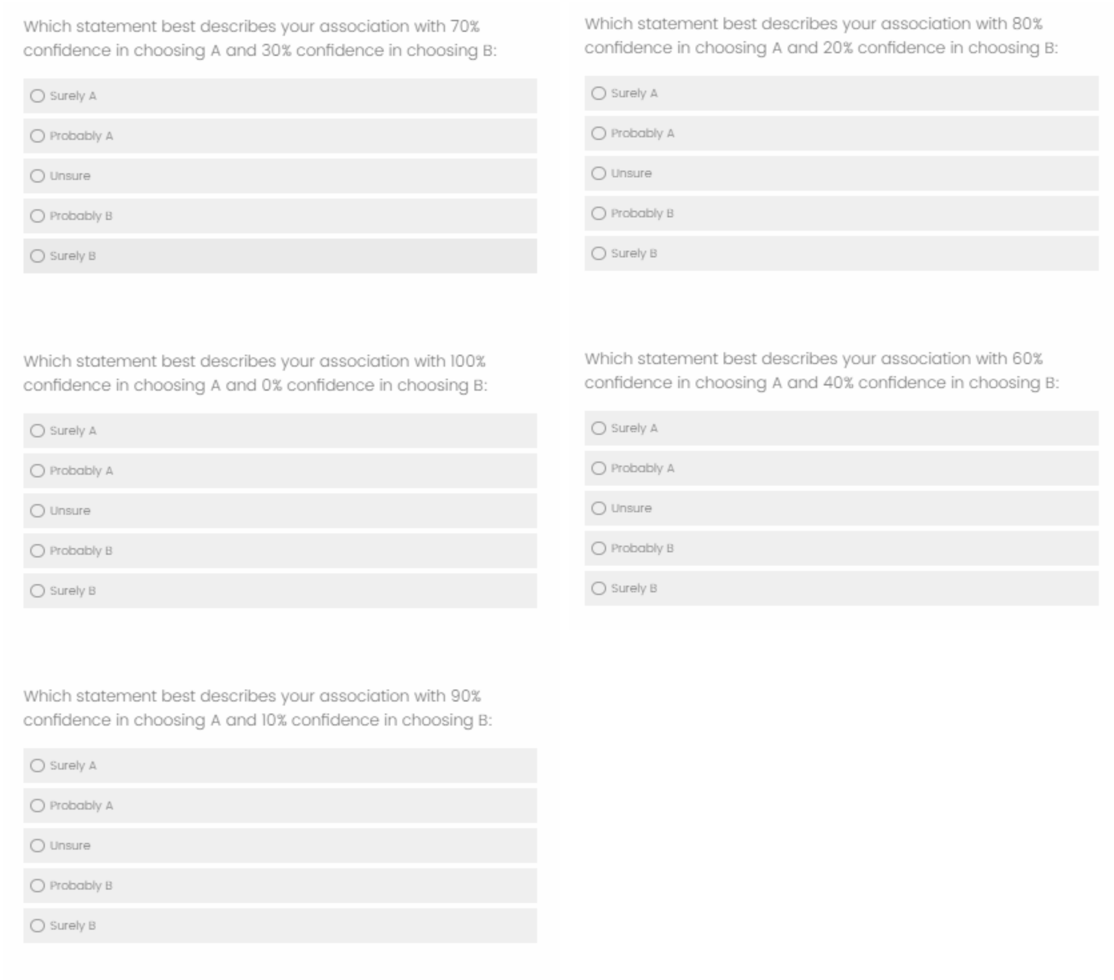


(a)



(b)

Figure C.11: The introduction (a) and an example of the decision screens (b) of probabilistic confidence choices for the complex lottery.


Which statement best describes your association with 70% confidence in choosing A and 30% confidence in choosing B:

- Surely A
- Probably A
- Unsure
- Probably B
- Surely B

Which statement best describes your association with 80% confidence in choosing A and 20% confidence in choosing B:

- Surely A
- Probably A
- Unsure
- Probably B
- Surely B

Which statement best describes your association with 100% confidence in choosing A and 0% confidence in choosing B:

- Surely A
- Probably A
- Unsure
- Probably B
- Surely B

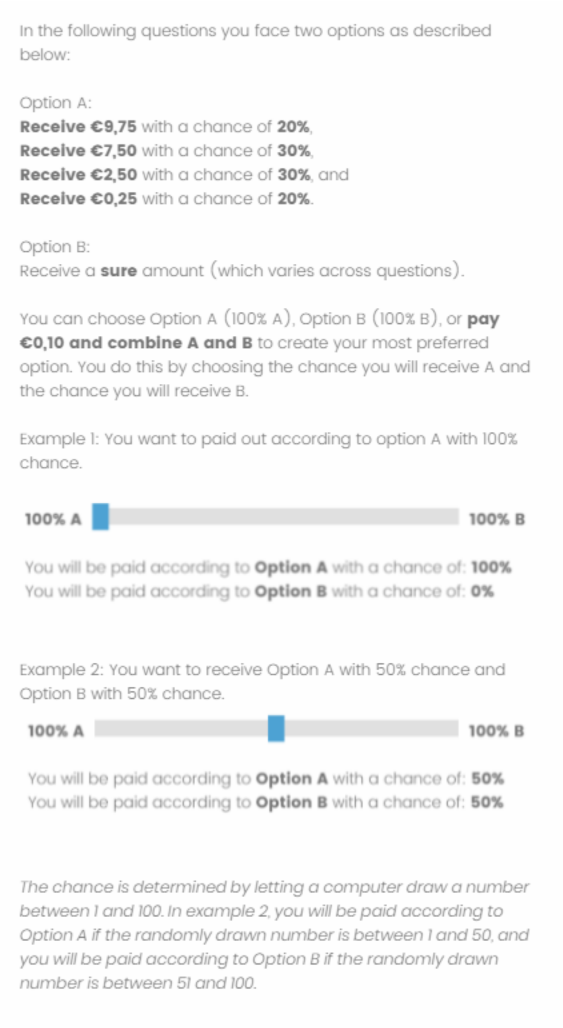
Which statement best describes your association with 60% confidence in choosing A and 40% confidence in choosing B:

- Surely A
- Probably A
- Unsure
- Probably B
- Surely B

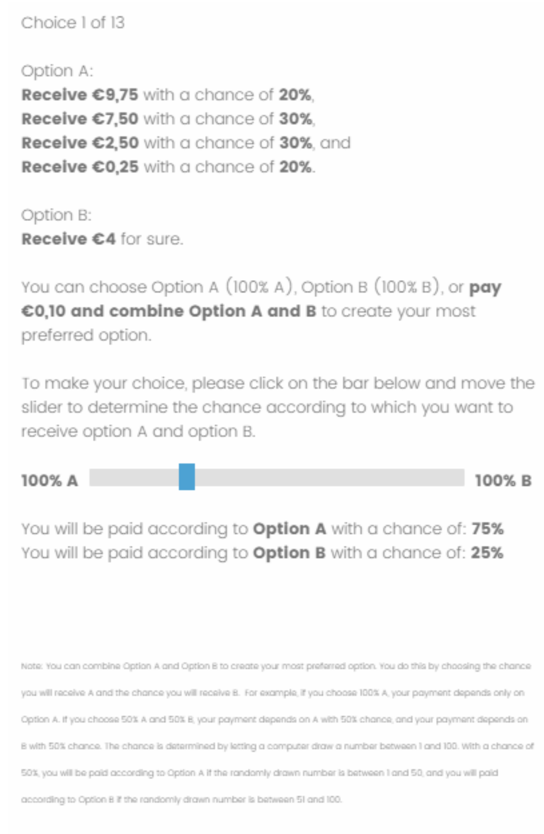
Which statement best describes your association with 90% confidence in choosing A and 10% confidence in choosing B:

- Surely A
- Probably A
- Unsure
- Probably B
- Surely B

Figure C.12: post-experiment questionnaire after the probabilistic confidence choices.

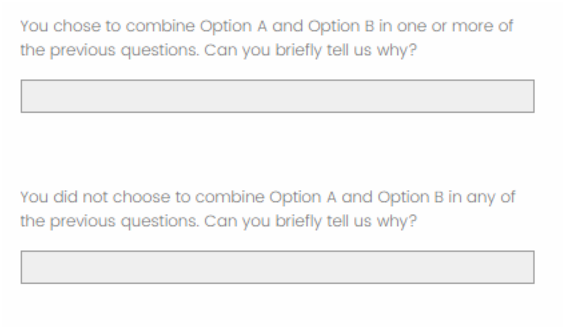


(a)



(b)

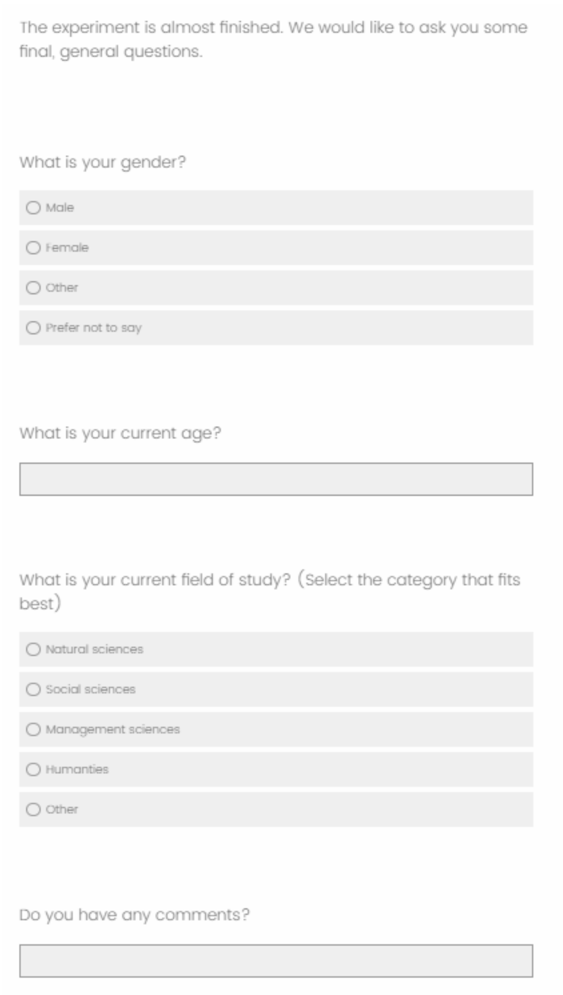
Figure C.13: The introduction (a) and an example of the decision screens (b) of randomized choices for the complex lottery.



You chose to combine Option A and Option B in one or more of the previous questions. Can you briefly tell us why?

You did not choose to combine Option A and Option B in any of the previous questions. Can you briefly tell us why?

Figure C.14: post-experiment questionnaire after the randomized choices. The first question was asked if a subject chose randomization probabilities other than 0 or 1 in at least 1 choice. The second question was asked if a subject only chose randomization probabilities of 0 or 1.



The experiment is almost finished. We would like to ask you some final, general questions.

What is your gender?

- Male
- Female
- Other
- Prefer not to say

What is your current age?

What is your current field of study? (Select the category that fits best)

- Natural sciences
- Social sciences
- Management sciences
- Humanties
- Other

Do you have any comments?

Figure C.15: Demographic questions asked at the end of the experiment.