MPRA
Munich Personal RePEc Archive

Measuring subjective decision confidence

Arts, Sara and Ong, Qiyan and Qiu, Jianying

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

August 2020

Online at https://mpra.ub.uni-muenchen.de/117907/
MPRA Paper No. 117907, posted 18 Jul 2023 06:36 UTC

Measuring subjective decision confidence

Sara Arts* Qiyan Ong† Jianying Qiu‡

November 27, 2022

Abstract

We examine whether the way individuals randomize between options captures their decision confidence. In two experiments in which subjects face pairs of options (a lottery and a varying sure payment), we allow subjects to choose randomization probabilities according to which they would receive each option. Separately, we obtain two measures of self-reported decision confidence for choosing between the two options. The randomization probabilities are well correlated and correspond with absolute levels of self-reported confidence measures. This relationship is robust to two exogenous manipulations of decision confidence, where we vary the complexity of the lottery and subjects' experience with the lottery, consistent with the predictions of two different theoretical frameworks incorporating preference uncertainty. Our findings suggest that randomization probabilities could serve as an incentivized quantitative measure of decision confidence.

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, EOS 02.577, 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 self-interest vs. social welfare when making policy decisions. As more studies are suggesting that decision confidence has the potential to explain a wide range of anomalies, there is growing interest in eliciting and accounting for decision confidence when studying people's choices.[1]

Most studies on decision confidence have relied on non-incentivized self-reported approaches. For example, Dubourg et al. (1994, 1997) allowed subjects to indicate whether they were unsure of their choices. Butler and Loomes (2007, 2011) asked subjects to indicate their decision confidence using the ordinal terms "definitely" and "probably." After subjects made choices from a multiple price list, Enke and Graeber (2021b) asked them how certain (from 0% to 100% in increments of 5%) they were that their actual valuation of an intertemporal payment/voucher was within their switching interval obtained from the multiple price list. Other methods examine decision confidence by having subjects state their confidence intervals. For example, Cohen et al. (1987) and Cubitt et al. (2015) had subjects report the range of choices over which they were unsure of their preferences, and Enke and Graeber (2021a) had subjects report the range of values over which they were certain of their preferences (75%, 90%, 95%, 99% and 100%).

We build on these studies and show that eliciting randomization probabilities of competing options could be a viable and incentivized approach to capture decision confidence in two experiments. Specifically, subjects faced pairs of options: a fixed lottery and varying amounts of a sure payment. Instead of requiring them to choose either the sure payment or the lottery, this approach allowed them to choose the randomization probabilities according to which they would receive each option (Qiu and Ong, 2017; Feldman and Rehbeck, 2022;

[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, 2021a), anomalies in intertemporal choices (Enke and Graeber, 2021b), small-stakes risk aversion (Khaw et al., 2021), and many other violations of standard decision theory (Butler and Loomes, 2011).

Miao and Zhong, 2018; Agranov and Ortoleva, 2020). By eliciting self-reported confidence measures separately, we are able to test whether the way subjects randomize between two options coheres with their self-reported decision confidence and compare how randomization probabilities and self-reported decision confidence relate to their choices between the options. We structure our analyses 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), to illustrate how an individual who is uncertain about her preferences would randomize. We show that randomization probabilities capture three important properties of decision confidence: first, subjects choose randomization probabilities around 0.5 for options that yield similar utility; second, when subjects face a lottery that is harder to evaluate, they randomize over a wider range of sure payments and choose randomization probabilities closer to 0.5; finally, when subjects gain more experience with the lottery and form clearer preferences, they randomize over a smaller range of sure payments and choose randomization probabilities further away from 0.5.

To test for the presence of these properties and show that randomization probabilities can measure decision confidence, we obtained subjects' binary choice, their self-reported decision confidence, and their randomization probabilities for each pair of options (a lottery x and a sure payment y) in the experiments. Self-reported decision confidence was elicited by having subjects select a confidence statement from "Surely x," "Probably x," "Unsure," "Probably y," "Surely y," in line with Dubourg et al. (1994) and Butler and Loomes (2007), and by having them select probabilistic confidences of p% x and 100-p% y, where p ranged from 0 to 40 and 60 to 100 in increments of 10. Separately, subjects had to choose 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. We exogenously manipulated decision confidence by a) having two lotteries: a simple lottery and a complex lottery with more payoff outcomes over a wider range of possible values, and b) increasing subjects' experience with the lottery by allowing them to observe the outcome draws of the lottery or to make hypothetical choices and observe the payoffs of their choices and the counterfactual. To ensure that experimenter demand effects and order effects were not the main drivers of our results, in one of our experiments, we elicited the measures separately over three sessions (at least seven days apart) in a random sequence.

Our experimental results suggest that randomization probabilities could be used as an incentivized quantitative measure of decision confidence. Subjects' randomization probabilities were strongly and positively correlated with both confidence statements and probabilistic confidence (median Spearman correlation between 0.86 to 0.89). Consistent with our hypotheses, the two exogenous manipulations affected decision confidence. Increasing the complexity of the lottery leads to a decrease in self-reported decision confidence, while increasing experience with lotteries leads to an increase in self-reported decision confidence. These exogenous changes in self-reported decision confidence are met with corresponding changes in randomization probabilities. As a result, the correlations between randomization probabilities and self-reported decision confidence measures are robust to the manipulations of the complexity of and experience with the lottery. Further analyses show that indifference, random errors, or utility differences alone cannot account for the randomization behavior in our experiment, and that decision confidence has important behavioral implications: across subjects and decisions, lower decision confidence in choosing an option corresponds with a lower proportion of choosing that option over the other option in binary choices.

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 (Rubinstein, 2002; Qiu and Ong, 2017; Dwenger et al., 2018; Feldman and Rehbeck, 2022; Miao and Zhong, 2018; Agranov and Ortoleva, 2020). 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 study that is closest to ours is Agranov and Ortoleva (2020), who also allowed subjects to choose randomization probabilities when deciding between two options. They found that subjects often randomized and did so over large ranges, and they related these ranges to certainty bias and non-monotonic choices.

Given the prevalence of preferences for randomization, it is critical to understand why they occur. 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, decision confidence, and randomization behavior. The only other study explicitly relating decision confidence to randomization probabilities that we are aware of is Agranov and Ortoleva (2020). 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, 2020, Appendix A.6). Our finding of a systematic relationship between the randomization and decision confidence suggests that a lack of decision confidence may be a psychological foundation for randomization behavior. The systematic changes of randomization probabilities and self-reported confidence due to the exogenous manipulation of decision confidence further supports Fudenberg et al.'s (2015) interpretation that randomization is deliberate and could arise from preference uncertainty.

We also contribute to the literature on stochastic choices, 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 choices, for example, Mosteller and Nogee (1951, Figure 2) and Loomes and Pogrebna (2017, Table 1).[2] The similarity

[2]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 face the lottery and a sure payment once and chose

between the choice probability 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 decisions, higher decision confidence in an option corresponds 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 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.

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. Section 5 presents a discussion of some practical considerations. Finally, Section 6 concludes the paper.

2 Experimental design

Besides eliciting randomization probabilities, we obtained two popular measures of self-reported decision confidence: qualitative confidence statements (Dubourg et al., 1994; Butler and Loomes, 2007, 2011) and quantitative probabilistic confidence (Enke and Graeber, 2021b). We went beyond establishing simple correlations between these measures by exogenously manipulating subjects' decision confidence and examining how the three measures of decision confidence responded to these manipulations. Our manipulations built on earlier findings suggesting decision confidence could be affected by the complexity of the decision problem (Enke and Graeber, 2021a,b) as well as people's experience with the decision (Myagkov and Plott, 1997; Plott and Zeiler, 2005; van de Kuilen and Wakker, 2006; Kuilen, 2009). We conducted two experiments with the same general structure but

the randomization probability of receiving the lottery.

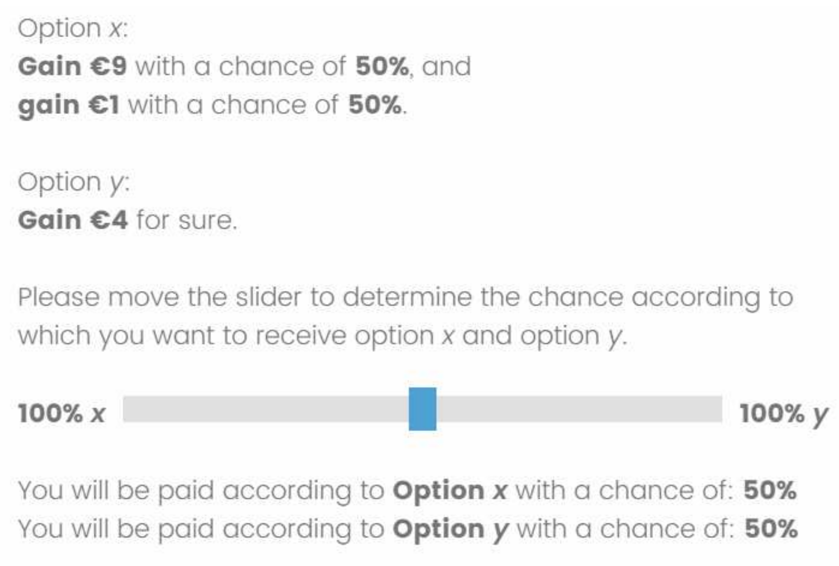


Figure 1: An example of the decision screen, where option x is a lottery with a 50% chance of gaining 9 euro and a 50% chance of gaining 1 euro. Option y is a sure payment and varies across choices. Subjects had to move the slider to determine the randomization probability. The randomization probability changed at an increment of 1%. Changes in the randomization probability were reflected in the descriptions below the slider.

with different treatments and experimental procedures. In the following, we 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: lottery x and sure payment y. Each lottery was paired with thirteen values of sure payments (0, 2, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, and 10 euros), resulting in 13 pairs of options. Subjects faced the 13 pairs of options in a random sequence for each type of decision. Each decision was made on a separate screen, and subjects were not allowed to review or change their decisions once they were made.[3] Each subject made three types of choices: binary choices, self-reported decision confidence, and randomized choices.

[3] Displaying the decisions from a choice list on separate computer screens helps to preserve isolation (subjects treat each decision as an independent decision from others), which is a sufficient condition for the incentive compatibility of the random incentive mechanism implemented in our experiments (see e.g., Brown and Healy, 2018; Freeman et al., 2019).

Binary choices

The binary choices required subjects to choose either lottery x or sure payment y for each of the 13 different values of sure payment y, in a random order. If the lottery x was chosen, the computer would draw a random number to determine the lottery's outcome. For example, for a lottery 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).[4] 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. Probabilistic confidence was not elicited in Experiment 1. Instead of making a direct binary choice, subjects had to choose how confident they felt about choosing lottery x versus sure payment 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 the subject chose "60% x, 40% y," she was considered to have chosen x over y in that decision. Her payoff would then be based on the lottery x. To use the probabilistic confidence as a measure of decision confidence as well as an indicator of a subject's choice between lottery x and sure payment y, we omitted the option of "50% x, 50% y" in the probabilistic confidence measure. Hence, probabilistic confidence has a threshold effect on choices, but selecting different values below (or above) the threshold made no difference in terms of the payoffs. Probabilistic confidence was also used in Enke and Graeber (2021b)

[4]Butler and Loomes (2007) and Butler and Loomes (2011) did not include the unsure option because they used the change of the confidence statements from Probably x to Probably y as the switching point to determine the payment.

as part of a two-step approach to measure cognitive uncertainty.

Randomized choices

To make randomized choices, subjects had to choose a randomization probability λ, based on which they would receive lottery x (and hence with a probability of $1-\lambda$ they would receive sure payment y). For example, a value of $\lambda = 0.40$ means subjects would receive lottery x with a chance of 40% and receive sure payment 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.[5] An important difference between randomized choices and probabilistic confidence is that every increment in the randomization probability affects the payoff, while increments in probabilistic confidence make no difference in the payoff except when probabilistic confidence increases from "40% x, 60% y" to "60% x, 40% y." Figure 1 shows the decision screen for the randomized choice. 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. Thus, unlike the self-reported decision confidence measures, the randomized choices were incentivized. As the subjects could perceive randomized choices as difficult, they were given two examples to illustrate the payoff mechanism of the randomized choice before making the decision. Additionally, we provided a concrete illustration in the lower part of the subjects' decision screen to remind them of the above incentive scheme (see Figure C.4 and C.5 in Appendix D).

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 and a detailed explanation of the lottery before making their decisions. We manipulated the subjects' decision confidence in two ways: varying the complexity of the lottery and the subjects' experience with the lottery.

[5] In experiment 1 the slider was set in the middle at the start. To avoid anchoring, 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.

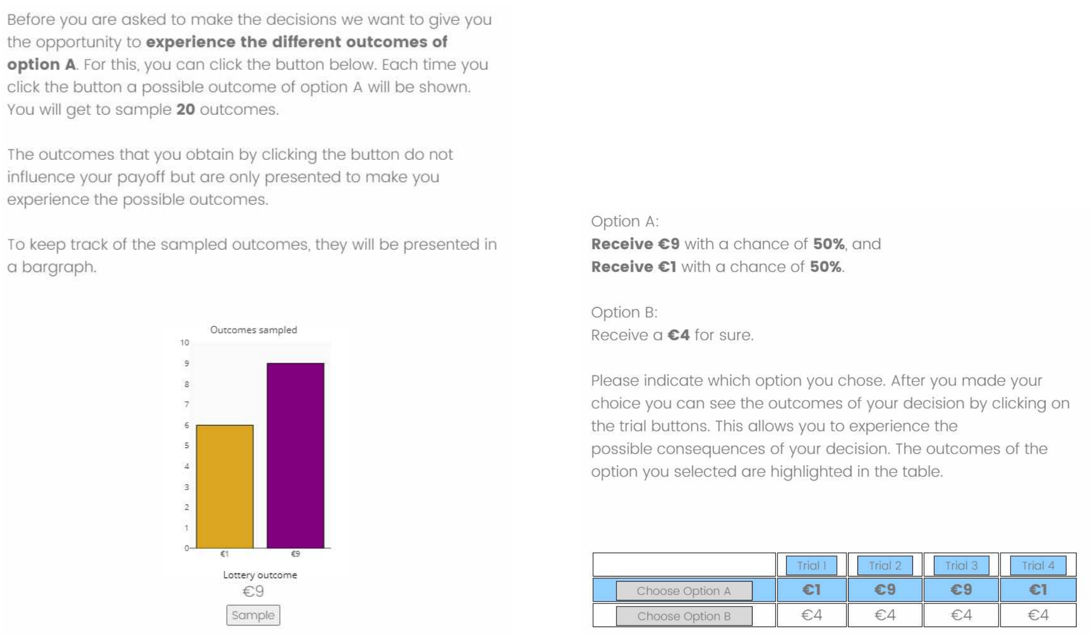
Varying the complexity of the lottery

In both experiments, apart from the simple lottery, the subjects also had to make decisions involving a complex lottery with four outcomes, but with the same expected payment as the simple lottery. The complex lottery offered the subjects 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%. Unlike studies documenting complexity seeking behaviors as well as complexity aversion (see e.g., Abdellaoui et al., 2020, and the references therein), our focus was on decision confidence, such as the range of sure payments that subjects do not have full confidence, rather than the average valuation of the lottery. All subjects made decisions involving the simple lottery and the complex lottery. The order of the lotteries was randomized: some subjects proceeded from the simple lottery to the complex lottery, while others completed the decisions in the reverse order.[6]

Varying subject's experience with the lottery

While we provided a complete description of the lottery, past studies have shown that decision experience helps individuals to better understand their preferences (Myagkov and Plott, 1997; Plott and Zeiler, 2005; van de Kuilen and Wakker, 2006; Kuilen, 2009), which could in turn improve their decision confidence. Following this literature, we also manipulated decision confidence by allowing some subjects to gain experience with the lottery before making actual choices. We varied the subjects' experience in two ways. The subjects were randomly assigned to the baseline treatment or the experience treatment of the respective experiment. In Experiment 1, we implemented a partial-experience treatment, where subjects had to click and view 20 draws of the lottery after they read the full description of the lottery, prior to proceeding to making actual decisions. While this allowed the subjects to experience the different outcomes of the lottery to gain more insights about their preferences between the lottery and sure payment amounts, this treatment only provided a "partial experience," as the subjects did not make any active choices or experience the consequences of their decisions. As the subjects viewed each lottery draw, an accom-

[6]The subjects in Experiment 1 additionally faced a loss lottery and a mixed lottery. We included these lotteries because we believe that they might lead to larger preference uncertainty due to the additional uncertainty in attitudes toward loss. 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 we currently rely on. Nevertheless, the results of these lotteries are consistent with our hypotheses and can be provided upon request.



(a) The partial-experience treatment

(b) The full-experience treatment

Figure 2: Panel (a) is an example of the partial-experience treatment, in which subjects could generate and experience the outcomes of the lottery 20 times. The final bar graph of the experienced outcomes was displayed on the subsequent decision screens. Panel (b) is an example of the full-experience treatment, in which subjects faced five levels of sure payments (3, 4, 5, 6, and 7) in a random sequence for the lottery. They made a hypothetical decision for each pair and could then click the trial button to experience the potential consequences of their choices four times.

panying bar chart, which recorded each lottery outcome, was updated. Figure 2(a) shows an example of the partial-experience treatment. The outcome distribution that the subjects saw after 20 draws was the same as the probability distribution of the lottery. Our treatment differs from Hertwig et al.'s (2004) decision from experience in two ways. First, our subjects saw the full description of the probabilities of the lottery, while the Hertwig et al. (2004)'s experience group did not. Second, experience in our experiment aims to help subjects to better understand their preferences between the lottery and sure payment amounts, while experience in Hertwig et al. (2004) aims to help subjects to understand the outcome distribution of the lottery.

In Experiment 2, we implemented a full-experience treatment, where subjects made hypothetical choices between the lottery and the sure payment and could observe possible payoffs of their choices as well as their counterfactuals. To avoid experimental fatigue, the

Experiment 1 (N = 205)		
Lotteries	Experience	Order
The simple lottery (50%, 9 ; 50%, 1)	Baseline no-experience	In the same order in one session:
		- Binary choices and confidence statements
The complex lottery (20%, 9.75 ; 30%, 7.50 ; 30%, 2.50 ; 20%, 0.25)	Partial-experience	- Randomized choices (free)
Experiment 2 (N = 293)		
Lotteries	Experience	Order
The simple lottery (50%, 9 ; 50%, 1)	Baseline no-experience	In random order across three sessions (at least 7 days apart):
The complex lottery		- Binary choices and confidence statements
(20%, 9.75 ; 30%, 7.50 ;	Full-experience	- Probabilistic confidence
30%, 2.50 ; 20%, 0.25)		- Randomized choices (a cost of 0.10 euro)

Table 1: Summary of the treatments and experimental procedure in Experiment 1 and 2. In Experiment 1, the subjects made the binary and confidence statements decision as well as the randomized choice sequentially. In Experiment 2, the subjects additionally made probabilistic confidence choices and answered the three types of questions in a random order across three sessions. They also needed to pay a fixed cost of 0.10 euros to randomize strictly. In the baseline no-experience treatment, subjects learned the full description of the lotteries. In the partial-experience treatment (in Experiment 1) the subjects could generate and observe the potential outcomes of the lotteries 20 times. In the full-experience treatment (in Experiment 2) the subjects made hypothetical choices and then generated and experienced the consequences of their hypothetical choices.

hypothetical decisions were limited to the sure payments of 3, 4, 5, 6 and 7 euros. For each decision, subjects were shown the payoffs of both the chosen and unchosen options in four separate trials. For example, if they had chosen the sure payment of 4 euros over the lottery, their hypothetical payoffs would be 4 euros for all four trials. Conversely, if they had chosen the lottery, the computer would make four random draws and display the outcome of the lottery in each trial. Figure 2(b) shows the decision screen viewed by a subject who chose the lottery over the sure payment of 4 euros in the hypothetical decision. A payoff table provided complete information about the hypothetical payoffs associated with each choice over the four trials as well as their counterfactuals. The row that is not highlighted shows what the subject could have earned if she had chosen the sure payment rather than the lottery. Likewise, a subject who chose the sure payment would see what she could have earned if she had chosen the lottery.

2.3 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. In both experiments, about half of the subjects were male (48% in Experiment 1, 45% in Experiment 2). The mean age of the subjects was 23 years. In Experiment 1, 49% of the subjects were randomly assigned to the partial-experience treatment, while the rest were assigned to the baseline treatment. In Experiment 2, 51% of the subjects were randomly assigned to the full-experience treatment, while the rest were assigned to the baseline treatment.

To demonstrate the correspondence between self-reported decision confidence and randomization probabilities for each subject, we elicited both types of decisions for each lottery x and sure payment y pair. A potential concern of this within-subject design is that experimenter demand effects may unintentionally influence subjects to give similar answers to the self-reported decision confidence measures and the randomized choices, resulting in a systematic relationship between the two. We used several methods to make it more obscure and costly for subjects to connect self-reported decision confidence and randomization probabilities in response to experimenter demand effects (Zizzo, 2010), such as spreading the decisions over three sessions (at least seven days apart), including a cost for randomizing, and randomizing the decision order. The key features of the two experiments are summarized in Table 1, and further details can be found in Appendix D.1.

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. The experimental instructions and decision screens are presented in online Appendix D. 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. If a subject's payment was based on a randomized choice, a random draw was used to determine whether the subject's payment would be based on the lottery or the sure payment. If the lottery was chosen for payment, a second random draw determined the outcome of the lottery. The average additional payment was 6.28 euros. We made the payment via bank transfers.

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 were asked about in that session. For example, after subjects completed the session on binary choices and confidence statements, they were asked to express how they understood the confidence statements by assigning the minimum and maximum probabilistic confidence level to each statement. Likewise, after they had completed the session on probabilistic confidence, the subjects had to choose which confidence statement best represents 100%, 90%, 80%, 70%, and 60% probabilistic confidence in choosing an option. Finally, after the session on randomized choices, the subjects were either asked to explain qualitatively why they chose to randomize or why they chose not to randomize, depending on which choice they made.

3 Theoretical analysis

Under the expected utility theory (EUT) which ignores decision confidence and assumes that a unique utility function (subject to positive affine transformation) captures an individual's preferences, 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. This prediction is inconsistent with a growing body of experimental evidence suggesting that, when facing two options, individuals may prefer to randomize strictly between rather than selecting a particular option (Qiu and Ong, 2017; Dwenger et al., 2018; Feldman and Rehbeck, 2022; Miao and Zhong, 2018; Cettolin and Riedl, 2019; Agranov and Ortoleva, 2020). Our theoretical analyses are based on this body of literature on preference convexity in probabilities.

As the main properties of decision confidence are not well-established in the literature, we focus on three properties that seem fairly intuitive in the context of our experiment. First, when two options are similar, individuals may feel less confident about which one they prefer. Second, when a lottery is harder to evaluate, such as a lottery with a more complex payoff structure, individuals may be less confident in making decisions over a wider range of sure payments. Finally, when individuals develop clearer preference about the lottery, they may become more confident in making decisions over a wider range of sure payments. If randomization probabilities capture decision confidence, they would also exhibit these properties.

We conducted two theoretical analyses of our experiments, which showed that randomization probabilities exhibit the properties of decision confidence. Appendix A.1 presents a theoretical framework based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015). To accommodate the possibility that a decision-maker might not be fully confident about her choices, we assume an individual has preference uncertainty. She has multiple utility functions that we call multiple selves, with each self representing one particular way of making the tradeoff between conflicting objectives in choices. With multiple selves, the individual is not fully confident about her choices when some selves choose one option while others choose other options. The individual randomizes to reconcile the disagreement among the different selves. This approach for capturing the lack of decision confidence from unsureness about preferences is closely related to but 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). In Appendix A.2 we extend Fudenberg et al.'s (2015) model to link randomization probabilities to decision confidence. 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 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.

To develop our experimental hypotheses, we note that subjects may perceive more preference uncertainty with the complex lottery than with the simple lottery, and experience with the lottery may reduce preference uncertainty regarding the lottery. As we demonstrate in Appendix A, a) subjects choose randomization probabilities close to 0.5 when two options have similar decision utility; 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 and correspond in absolute levels to probabilistic confidence and the confidence statements.*

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) Increasing the complexity of the lottery decreases subjects' 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) Gaining experience with a lottery increases subjects' decision confidence when making decisions about the lottery, 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 of our experiments in two steps. We begin by showing the systematic link between randomization probabilities and the two measures of self-reported confidence (confidence statements and probabilistic 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 intended 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 (hypothesis 3 and 4).

In online Appendix C.1, we discuss the different orders in which subjects completed the decision tasks in Experiment 2. While we observe some order effects, the main results hold across the different orders. For this reason, we report the pooled results for all the subjects of Experiment 2 here.

4.1 Randomization probabilities and self-reported confidence

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

Result 1. *Randomization probabilities were significantly correlated and corresponded in absolute levels with probabilistic confidence and confidence statements.*

To obtain the correlation between randomization probabilities and confidence statements, we transformed the confidence statements "Surely y," "Probably y," "Unsure," "Probably x," and "Surely x" 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 the decision confidence in choosing the lottery x. We computed the nonparametric Spearman correlation between confidence statements and randomization probabilities for each subject in Experiment 1 and 2. We also computed the nonparametric Spearman correlation between subjects' probabilistic confidence and their randomization probabilities for each subject in Experiment 2. The results are summarized

	Correlation between randomization probabilities and		
	Confidence statements		Prob. confidence
	Experiment 1	Experiment 2	Experiment 2
10th percentile	0.60	0.71	0.73
median	0.91	0.89	0.90
90th percentile	0.97	0.96	0.98

Table 2: Nonparametric Spearman correlations between randomization probabilities and confidence statements as well as probabilistic confidence at the 10th percentile, 50th percentile, and 90th percentile in the two experiments in the baseline no-experience treatment for decisions about the simple lottery.

in Table 2. Consistent with Hypothesis 1, the first two columns of Table 2 show that confidence statements and randomization probabilities have a high and positive correlation. Moderate to strong correlations of 0.60 in Experiment 1 and 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 least seven days, the similarities between the correlations found in Experiments 1 and 2 suggest that confidence statements and randomization probabilities are associated in ways beyond the order effects. In Experiment 2, we also found high correlation levels between self-reported probabilistic confidence and randomization probabilities which are reported in the third column: the correlation is 0.73 at the 10th percentile, 0.90 at the median, and 0.98 at the 90th percentile.

As correlations do not describe the correspondence between randomization probabilities and self-reported decision confidence at absolute levels, we also computed the average randomization probability at each self-reported probabilistic confidence level in Table 3. 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. The results suggest randomization probabilities can be used as a direct proxy for probabilistic confidence.

There may be concerns that subjects gave similar answers to probabilistic confidence and randomization probabilities because the two measures resemble each other closely. To reduce such concerns, we used the probabilistic confidence associated with each confidence

	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
Rand.	0.98	0.86	0.82	0.73	0.58	0.36	0.23	0.18	0.09	0.02
prob.	(0.008)	(0.021)	(0.022)	(0.022)	(0.027)	(0.024)	(0.023)	(0.022)	(0.018)	(0.004)

Table 3: The mean randomization probabilities at each self-reported probabilistic confidence level in Experiment 2 in the baseline no-experience treatment for decisions about the simple lottery. The standard errors of the mean are reported in parentheses. We compute the mean randomization probability at each level of probabilistic confidence for each subject before taking its mean across subjects.

statement reported in Vanberg (2008) as references - 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 (Vanberg, 2008, Footnote 10, p.1472). We then computed the mean, minimum, and maximum randomization probabilities associated with each confidence statement from our experiments and compare these values to the cutoff values from Vanberg (2008). Consistent with Vanberg (2008), 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. More details can be found in Figure B.1 in Appendix B.

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. *The subjects reported lower decision confidence around the switching choices based on the self-reported confidence measures. Meanwhile, they 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 consistently preferred x over y, and $\bar{y_b}$ denote the lowest sure payment amount at and above which the subject consistently 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).[7]

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 4 shows the median randomization probabilities, probabilistic confidence, 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.

Having established that self-reported confidence is lower in the switching range, we turn to the randomization probabilities within this range. We find that 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 was 0.5. These results are consistent with Hypothesis 2 that subjects are more likely to choose randomization probabilities close to 0.5 for choices that they find difficult to compare.

4.2 Manipulating decision confidence

In this section, we examine whether our exogenous manipulations of the decision situation affect self-reported decision confidence in the expected direction and whether randomiza-

[7] 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	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 4: Behavior around the switching choices for the baseline no-experience treatment for decisions about the simple lottery. We let $\underline{y_b}$ denote the highest sure payment amount at and below which subjects consistently preferred x over y, and $\bar{y_b}$ denote the lowest sure payment amount at and above which subjects consistently chose y over x. Here, we show the median randomization probabilities, confidence statements and probabilistic confidence levels subjects reported at $\underline{y_b}$ and $\bar{y_b}$.

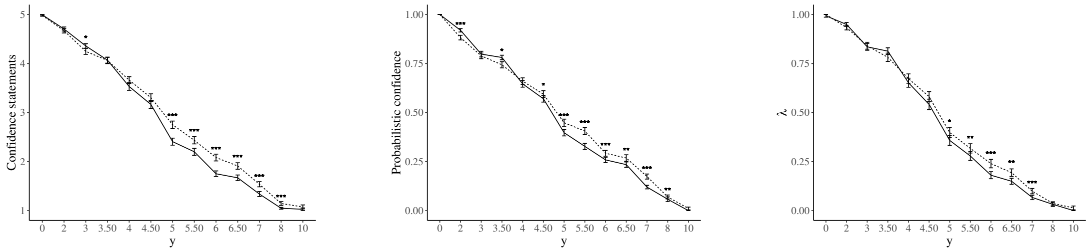
tion probabilities are affected in similar ways to maintain a systematic relationship with the self-reported confidence measures. We discuss the results for the two treatments separately.

4.2.1 The complex lottery versus the simple lottery

Hypothesis 3 states that subjects have lower decision confidence when making decisions about the complex lottery compared to the simple lottery, and this is reflected in the differences in self-reported decision confidence measures as well as randomization probabilities across lotteries. Result 3 summarizes our findings.

Result 3. *Compared to decisions about the simple lottery, the subjects did not report full decision confidence over a wider range of sure payments and reported decision confidence more compressed toward "Unsure" or (50% x, 50% y) for decisions about the complex lottery. 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.*

Comparing the range of sure payments over which the subjects chose confidence statements "Unsure" or "Probably," we find that the range size was larger for decisions about the complex lottery than for decisions about the simple lottery in both Experiment 1 and 2, and it was 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 $p < 0.01$). We also computed the range of sure payments for which subjects were not fully



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

Figure 3: 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 difference between the simple and the complex lotteries for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

confident about their choices using their probabilistic confidence reports in Experiment 2. The range is the difference between the minimum and maximum sure payment for which subjects did not indicate probabilistic confidence of (100% x, 0% y) or (0% x, 100% y). Similar to the results obtained from the confidence statements, 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).

We detail how confidence statements and probabilistic confidence varied with different sure payment amounts across the two lotteries in Experiment 2 in Panel (a) and Panel (b) of Figure 3. The results for Experiment 1 were similar, albeit weaker, and are presented in Figure B.2 of Appendix B. 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 less often statistically significant for lower sure payment amounts.

Having established that the complexity of the lottery affected the subjects' self-reported decision confidence in the intended way, we proceed to examine the randomization probabilities chosen for each lottery. 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 $p < 0.10$). A summary of the range sizes for each lottery found in Experiment 1 and Experiment 2 can be found in Table B.4 in Appendix B.

Like self-reported decision confidence, Panel (c) of Figure 3 shows that the randomization probabilities were also more compressed towards 0.5 when the subjects faced the complex lottery compared to the simple lottery. The difference in randomization probabilities across the two lotteries was statistically significant between sure payments of 5 and 7 euros, coinciding with the range obtained from probabilistic confidence. Like self-reported decision confidence, Panel (c) of Figure 3 show 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 is consistent with the theoretical analysis. This asymmetry arises because the treatment manipulation affects both preference uncertainty and the average valuation of the lotteries.

Next, we find that the correlations between the two decision confidence measures and randomization probabilities remain similar after our exogenous manipulation. 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. Further results on the associations between self-reported decision confidence measures and randomization probabilities, similar to those reported in subsection 4.1, can be found in Table B.1, B.2, and B.3 in Appendix B.

4.2.2 Experience and no experience

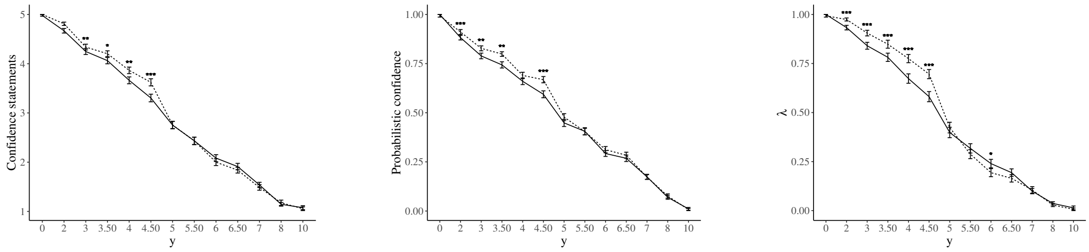
Hypothesis 4 states that, compared to the baseline no-experience treatment, gaining experience with the lotteries increases decision confidence. We find that the subjects in the full-experience treatment had higher decision confidence than those in the no-experience treatment when the subjects had to make decisions about the complex lottery. Result 4 summarizes our findings. As the partial-experience treatment and decisions about the simple lottery did not exhibit significant treatment effects, we report the results of these treatments in Figure B.3 and Figure B.4 and Table B.5 in Appendix B.

Result 4. *Compared to the subjects in the no-experience treatment, the subjects in the full-experience treatment revealed less than full decision confidence over a narrower range of sure payments for decisions about the complex lottery and their self-reported decision confidence were stretched further away from "Unsure" or (50% x, 50% y). Likewise, the subjects in the full-experience treatment randomized over a narrower range of sure payments and chose randomization probabilities further away from 0.5 for the complex lottery compared to the subjects in the no-experience treatment.*

Examining the range of sure payments over which subjects reported confidence statements of "Probably" or "Unsure", we find a significantly narrower range in the full-experience treatment compared to the no-experience treatment (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 significantly different 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 4 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.

Having shown the treatment effects of the full experience on the self-reported decision confidence, we next examine these treatment effects on randomization probabilities. Like decision confidence, we find that the range of sure payments over which subjects chose randomization 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 4 shows that the difference in mean randomization probabilities across sure payments between the full-



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

Figure 4: 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 difference between the simple and complex lottery for each value of y: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 due to allowing subjects to gain 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 in the full-experience treatment was broadly similar to the median correlation in the no-experience treatment. More details can be found in Table B.1, B.2, and B.3 in Appendix B.

5 Discussion: Practical considerations

So far, we have demonstrated that randomization probabilities could be a good proxy for decision confidence. Here, we discuss some practical considerations of using randomization probabilities as a proxy for decision confidence. First, randomization probabilities, or any decision confidence measure, have little practical relevance if they merely reflect indifference, random errors, or utility difference. If randomization is only driven by indifference, randomization probability values do not provide additional information, as subjects can choose any values of randomization probabilities when they are indifferent. If random errors are responsible for randomization, randomization probabilities provide little information other than potentially revealing inattention during the experiment and thus those data should be ignored in the analysis. Finally, if utility difference alone drives randomization and preference uncertainty is irrelevant, randomization probabilities would not have significant behavioral implications; subjects will always choose the option with higher utility, regardless of how small the utility difference is. In Appendix B.1 we show that indifference, random errors, or utility difference alone cannot be the driving force behind subjects' randomization behavior.

More importantly, there is evidence that randomization probabilities contain information about how subjects made binary choices. Figure 5 reports the proportions of decisions in which lottery x was chosen in the binary choice at each randomization probabilities and probabilistic confidence level.[8] Across subjects and decisions, higher randomization probability was associated with a higher likelihood that the option would be chosen in the binary choices. Figure 5 also shows this association with the probabilistic confidence. Interestingly, while probabilistic confidence had similar behavioral implications as randomization probabilities, it corresponded less closely to binary choices, as reflected by the Fisher's exact tests in Figure 5. Further analysis also suggested that randomization probabilities were more sensitive to variations in sure payments around the switching choices than probabilistic confidence, where subjects are least likely to have strict preferences. This means that changes in decision confidence around the switching choices are more likely to be detected when randomization probabilities are used as the proxy for decision confidence rather than probabilistic confidence. Detailed analysis can be found in Appendix B.2.

[8] We do not consider confidence statements in this analysis, as binary decisions and confidence statements were made on the same decision screen.

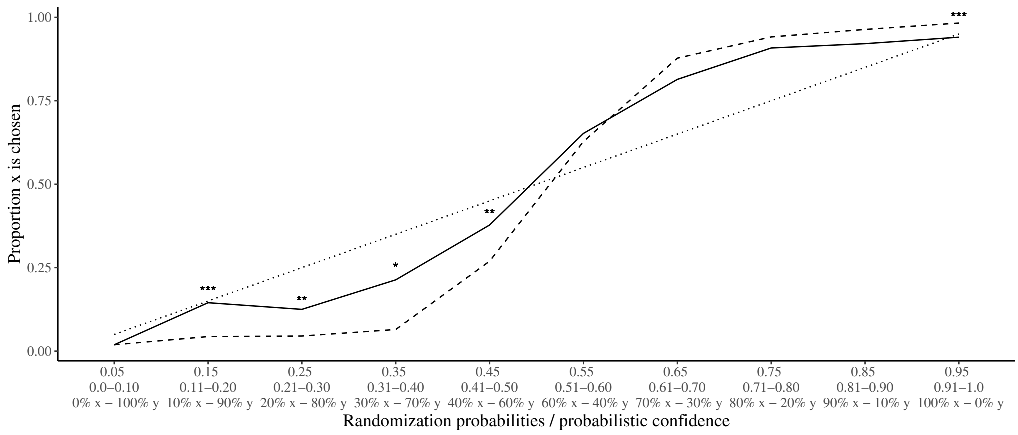


Figure 5: Correspondence of randomization probabilities (solid line) and probabilistic confidence (dashed line) with the proportion of times lottery x is chosen (y-axis) for the simple lottery in the baseline no-experience treatment. The dotted line represents a 45-degree line. Fisher's exact tests were performed to test the difference in choice proportions between the randomization probabilities and probabilistic confidence: ** $p < 0.05$, *** $p < 0.01$.

6 Conclusion

We have shown in this study that letting individuals assign randomization probabilities according to which they receive two options can be a viable and incentivized way to elicit decision confidence. We showed the link between randomization probabilities and decision confidence theoretically through the frameworks based on Klibanoff et al. (2005) and Cerreia-Vioglio et al. (2015) as well as Fudenberg et al. (2015) and demonstrated this relationship empirically through two experiments.

Our experimental results provide strong evidence suggesting that the randomization probability for an option could be interpreted as the probabilistic confidence related to choosing that option. In this study, the majority of subjects randomized frequently, and the randomization pattern was consistent with our theoretical analysis. We further found that randomization probabilities were highly correlated and varied systematically with self-reported decision confidence, with high randomization probabilities for options associated with high self-reported decision confidence. Our further examination of alternative interpretations of randomization suggested that indifference, random errors, and differences in utility alone were 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. (2020). Ranges of preferences and randomization. Working paper, California Institute of Technology.

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.

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

Brown, A. L. and Healy, P. J. (2018). Separated decisions. *European Economic Review*, 101(C):20–34.

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(4):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.

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

Cerreia-Vioglio, S., Dillenberger, D., Ortoleva, P., and Riella, G. (2019). Deliberately stochastic. *American Economic Review*, 109(7):2425–45.

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

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

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):1065–92.

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., Navarro-Martinez, D., and Starmer, C. (2015). On preference imprecision. *Journal of Risk and Uncertainty*, 50(1):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(2):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 Series 4740, CESifo Group Munich.

Dwenger, N., Kübler, D., and Weizsäcker, G. (2018). Flipping a coin: Evidence from university applications. *Journal of Public Economics*, 167(C):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. (2021a). Cognitive uncertainty. Working paper, Harvard University.

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

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

Freeman, D. J., Halevy, Y., and Kneeland, T. (2019). Eliciting risk preferences using choice lists. *Quantitative Economics*, 10(1):217–237.

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.

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.

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

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. rdaa016.

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.

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.

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.

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

Qiu, J. and Ong, Q. (2017). Indifference or indecisiveness: a strict discrimination. MPRA Paper 81790, University Library of Munich, Germany.

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(4):297–323.

van de Kuilen, G. and Wakker, P. (2006). Learning in the Allais paradox. *Journal of Risk and Uncertainty*, 33(3):155–164.

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

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 $l \in L$.[9] 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

[9]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 $U_\tau\left[\lambda x + (1-\lambda)y\right] = \lambda U_\tau(x) + (1-\lambda)U_\tau(y)$. The individual's decision is then to maximize her utility by choosing the optimal randomization probability $0 \leq \lambda \leq 1$:

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

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:[10]

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

where $\sigma_x^2 = E_\pi\left[U_\tau(x) - E_\pi(U_\tau(x))\right]^2$ 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

[10] More precisely, since $0 \leq \lambda \leq 1$, $\lambda^* \approx min\left\{max\left\{0, \frac{1}{-\frac{\phi''[u(y)]}{\phi'[u(y)]}} \times \frac{\Delta_u}{\sigma_x^2}\right\}, 1\right\}$. 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

δ_x decreases). These lead to the following hypotheses:

Hypotheses.

a) Subjects choose randomization probabilities close to 0.5 around sure payments, where subjects switch between the lottery and the sure payments in binary choices.

b) Compared to the simple lottery, subjects randomize strictly over a wider range of sure payments when making decisions about the complex lottery.

c) Compared to the no-experience treatment, subjects in the experience treatments randomize strictly over a smaller range of sure payments.

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^*

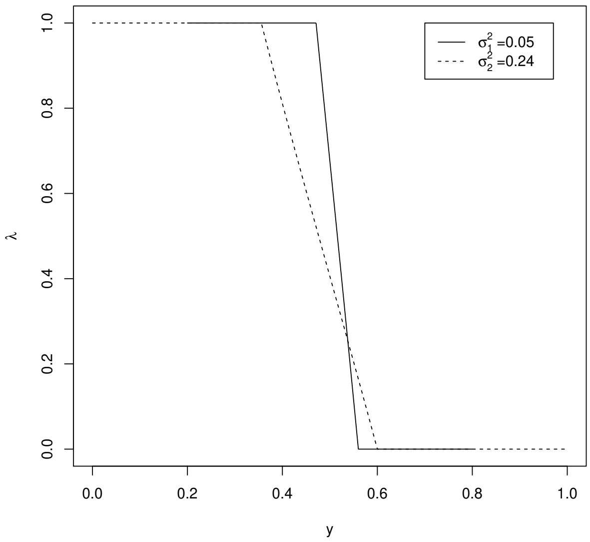


Figure A.1: The relationship between the randomization probability λ and the sure payment y. The figure is produced by assuming $\phi(U_\tau) = 1 - e^{-U_\tau}$, $\pi(u_1) = 0.6$, $\pi(u_2) = 0.4$, $U_1(x_1) = 0.8$ and $U_2(x_1) = 0.2$, $U_1(x_2) = 1.0$ and $U_2(x_2) = 0$, and $U_1(y) = U_2(y) = y$.

Taking the first order derivative of the optimization equation gives:[11]

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

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

[11]The second-order derivative is

$$\frac{d^2V\left[\lambda x + (1-\lambda y)\right]}{d\lambda^2} = \int_{\mathcal{T}} \phi''\left[\lambda U_\tau(x) + (1-\lambda)u(y)\right] \times \left[U_\tau(x) - u(y)\right]^2 d\pi.$$

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, $\frac{d^2V[\lambda x + (1-\lambda y)]}{d\lambda^2} = \int_{\mathcal{T}} \phi''\left[\lambda U_\tau(x) + (1-\lambda)u(y)\right] \times \left[U_\tau(x) - u(y)\right]^2 d\pi < 0$. This ensures we are indeed seeking for the maximum.

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 $U_\tau(x)$.

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.[12] 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:[13]

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

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

[12] 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.

[13] 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 $\Delta = c'(0.9) - c'(0.1) > 0$.[14] 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

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

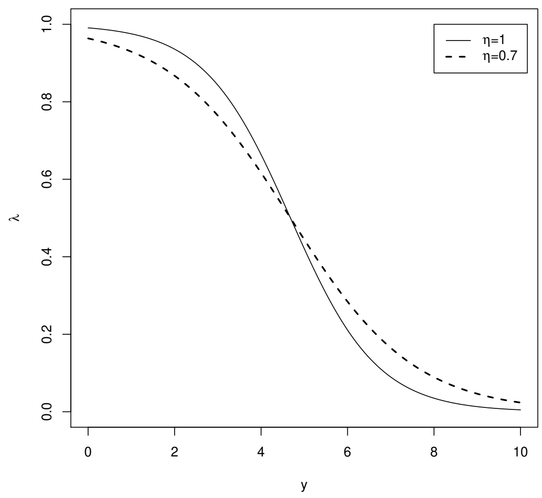
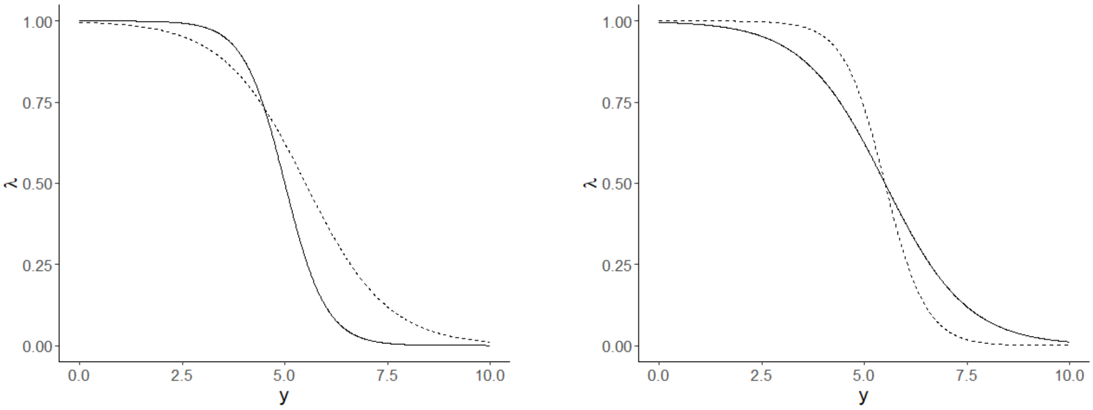


Figure A.2: The relationship between the optimal randomization probability λ^* and the sure payments y. The figure is produced according to the logit/logistic choice rule $\lambda^* = \frac{e^{U(x)/\eta}}{e^{U(x)/\eta}+e^{u(y)/\eta}}$. The parameter η captures the preference uncertainty over lottery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

These hypotheses are the same as Appendix A.1.



(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)'s approach

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), $u(\bar{y_x}) = E_\pi [U_\tau(x)]$ and $\underline{y} = E_\pi [U_\tau(x)] - \frac{-\phi''[u(y)]}{\phi'[u(y)]}\sigma_x^2$. The changes in the upper and lower bounds depend on both $E_\pi [U_\tau(x)]$ 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 $E_\pi [U_\tau(x)] - \frac{-\phi''(E_\pi[U_\tau(x)])}{2}\sigma_x^2$, this implies an increase in $E_\pi [U_\tau(x)]$ for the complex lottery. The increase in $E_\pi [U_\tau(x)]$ 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 $E_\pi [U_\tau(x)]$ 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

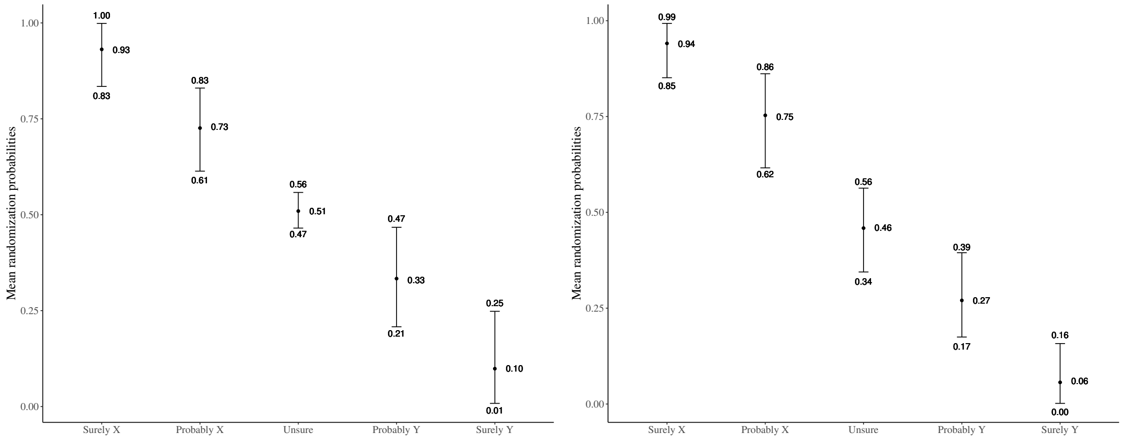
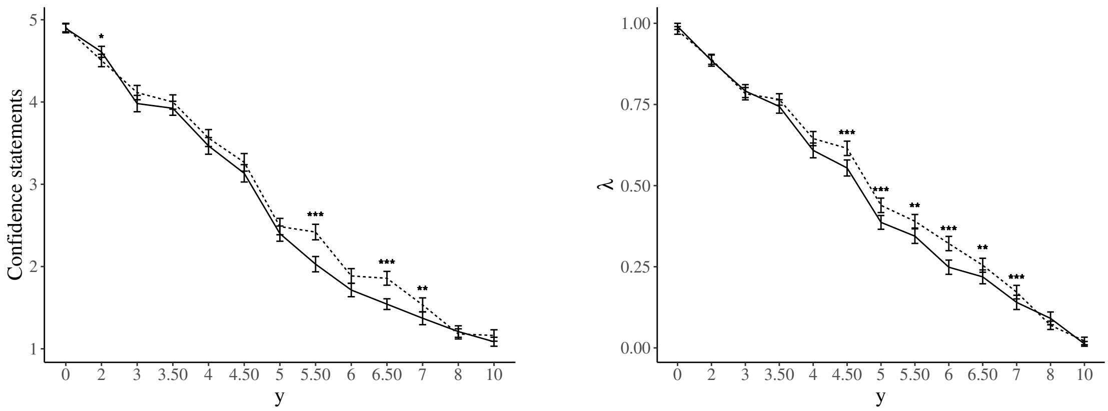
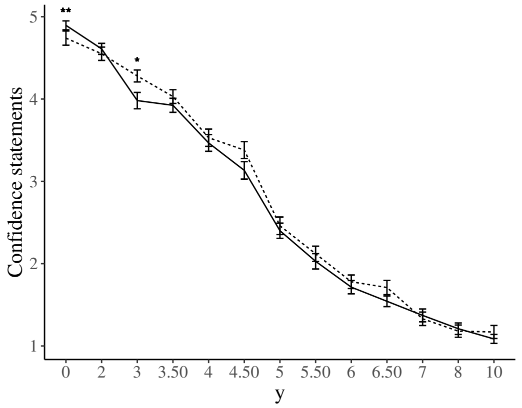


Figure B.1: The mean randomization probabilities at each confidence statement. The bars show the average minimum and maximum values. The values show the aggregate values for the baseline treatment – simple lottery, no-experience – in Experiment 1 (left) and Experiment 2 (right). The mean, minimum, and maximum values for the separate treatments in each of the experiments can be found in Table B.1 in Appendix B.

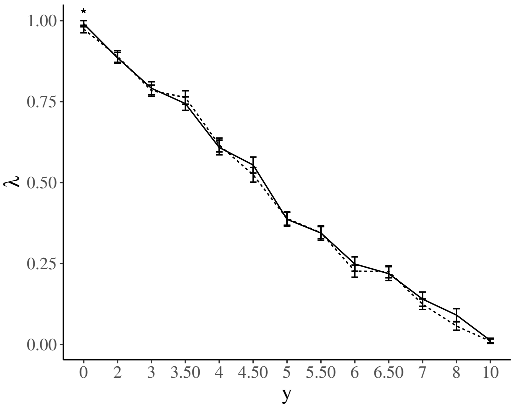


(a) Confidence statements (b) Randomization probabilities

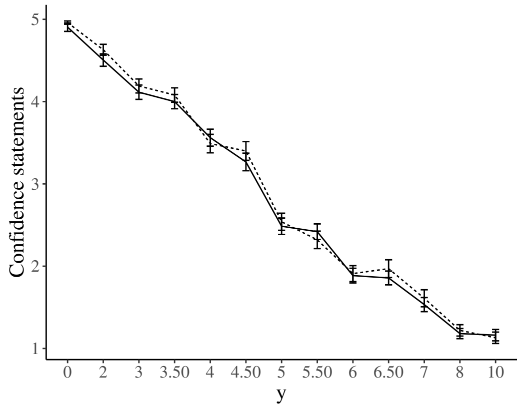
Figure B.2: The mean self-reported decision confidence and randomization probabilities for each value of y obtained from decisions about the simple lottery (solid line) and decisions about the complex lottery (dashed line) in Experiment 1. Wilcoxon signed-rank tests were performed to test the difference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



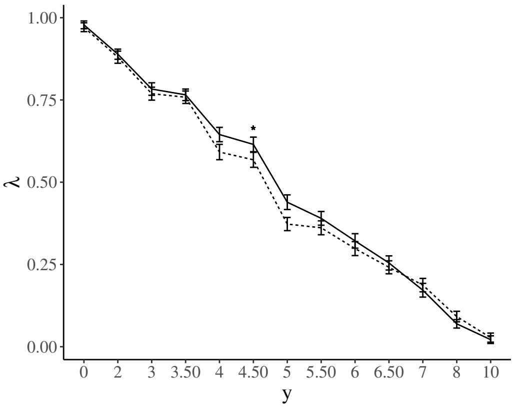
(a) Confidence statements, simple lottery



(b) Randomization probabilities, simple lottery

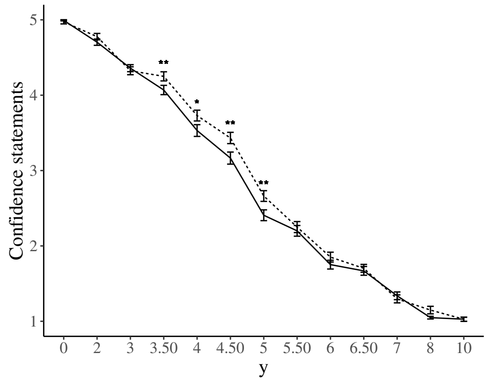


(c) Confidence statements, complex lottery

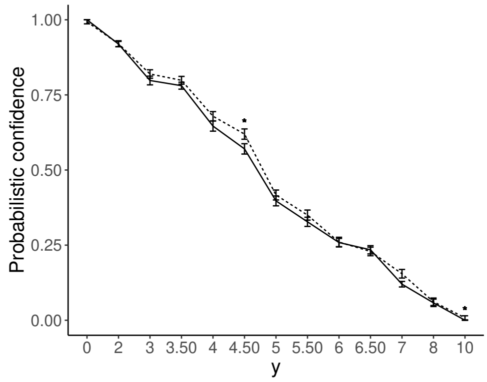


(d) Randomization probabilities, complex lottery

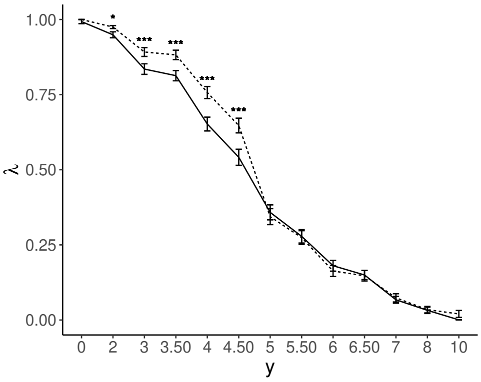
Figure B.3: The mean self-reported decision confidence and randomization probabilities for each value of y in the no-experience treatment (solid line) and the partial-experience treatment (dashed line) in Experiment 1. Wilcoxon rank-sum tests were performed to test the difference between the partial-experience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



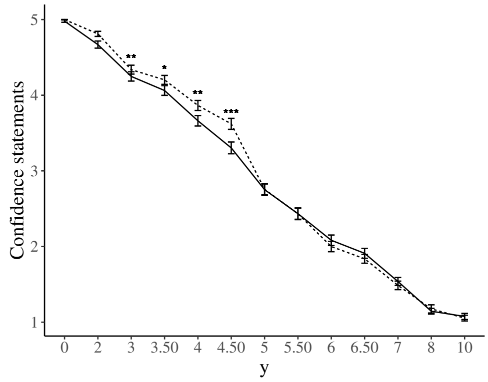
(a) Confidence statements, simple lottery



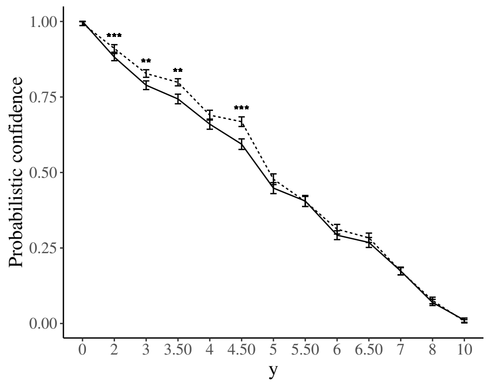
(b) Probabilistic confidence, simple lottery



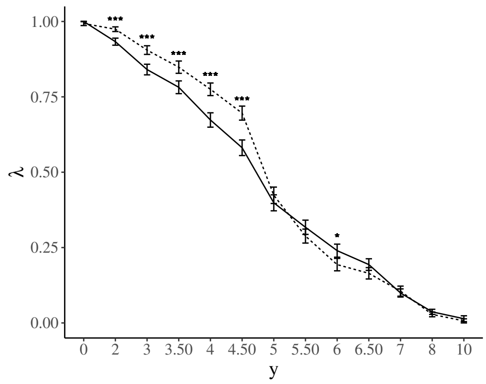
(c) Randomization probabilities, simple lottery



(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: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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

Table B.1: 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	Treatment		Correlation between randomization probabilities and confidence statements Experiment 1	Correlation between randomization probabilities and confidence statements Experiment 2	Correlation between randomization probabilities and prob. confidence Experiment 2
Simple	No experience	10th percentile	0.60	0.71	0.73
		median	0.91	0.89	0.90
		90th percentile	0.97	0.96	0.98
	Experience	10th percentile	0.60	0.78	0.77
		median	0.92	0.90	0.91
		90th percentile	0.97	0.97	0.99
Complex	No experience	10th percentile	0.69	0.67	0.64
		median	0.90	0.88	0.89
		90th percentile	0.97	0.96	0.97
	Experience	10th percentile	0.62	0.69	0.77
		median	0.88	0.90	0.90
		90th percentile	0.96	0.97	0.97

Table B.2: Nonparametric Spearman correlation at the 10th percentile, median, and 90th percentile in the two experiments for each lottery and experience treatment group.

	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.3: 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	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 ***	6.57 ***	7.21 ***
Range size	Simple	3.03	3.15	4.37
	Complex	3.19 *	3.58 ***	4.63 ***

Table B.4: 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: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 Other interpretations of randomization probabilities

We have interpreted randomization behavior as a lack of decision confidence in the face of preference uncertainty. Our theoretical analysis provides an explicit link between randomization probabilities and decision confidence, and our experimental results show a systematic relationship between randomization probabilities and two measures of self-reported decision confidence. Like Agranov and Ortoleva (2020), we also found that many subjects explicitly mentioned unsureness, complexity, difficulty, and hedging as reasons for randomization in the post-experiment questionnaire (for more details, see online Appendix C.2). Nevertheless, the subjects may randomize for reasons other than decision confidence. While it is not possible to eliminate all alternative interpretations, we show in this section that indifference, random errors, or utility difference alone cannot be the driving force behind subjects' randomization behavior.

First, we rule out the explanation that indifference drives randomization because randomization from indifference could occur at most once, but Table B.6 shows that the majority of the subjects randomized at least two times. Second, randomization is also unlikely due to random errors because, as Panel (c) in Figure 3 shows, despite the random sequence of y, the randomization probabilities of choosing x decreased monotonically with the value of y. This result is consistent with Equation 1 and Figure A.1. The systematic response of randomization probabilities to our manipulation of decision confidence also suggests that subjects randomize deliberately.

A third explanation is that randomization captures only utility differences instead of decision confidence. Butler et al. (2014) call this 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 (comparing the simple lottery to the complex lottery or the no-experience treatment to the full-experience

Randomization Interval	The number of subjects who chose randomization			
	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.6: 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.

treatment), (1) the randomization probability will increase for each value of sure payment; and (2) the lower bound and the upper bound of randomization will increase. Our results clearly reject the first prediction, and the results were mixed for the second. 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 stretched away from 0.5 in the full-experience treatment compared to the no-experience treatment. In addition, while $\bar{y}$ for the complex lottery was significantly higher than that 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 that for the simple lottery in Experiment 1 at 10% significance level (2.84 vs 2.99, Wilcoxon signed-rank test $p < 0.10$), and 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, we observe the opposite effects when we compare the full-experience treatment with the no-experience treatment for decisions about the same lottery. We find that $\underline{y}$ were significantly lower for subjects in the full-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 for $\bar{y}$ (Simple lottery: 6.18 vs 6.18, Wilcoxon rank-sum test $p = 0.789$; Complex lottery: 6.27 vs 6.38, Wilcoxon rank-sum test $p = 0.369$). In general, randomization probabilities were larger at low sure payments but smaller at high sure payments comparing the full-experience treatment with the no-experience treatment. These results highlight the central role of preference uncertainty in decision confidence.

Interpretability of self-reported decision confidence measures vs randomization probabilities

There may be concern that self-reported decision confidence measures are easier to interpret than randomization probabilities because they ask about decision confidence explicitly. Below we show that this perception is not warranted: self-reported decision confidence measures are equally, if not more, difficult to interpret. We analyze the subjects' responses in the post-experiment questionnaire administered at the end of each part of Experiment

2. We asked the subjects which confidence statement best describes 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 the 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%.

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.7: 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.

Table B.7 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 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.

Despite the explicit linkage to probabilistic confidence in our experiment, there appear to be two interpretations of the confidence statement "Unsure". While the majority of the subjects (172) reported a probabilistic confidence level higher than 50% for the maximum of

"Unsure" and higher than 0% for the minimum of "Unsure", as we expected, there exist also a substantial number of subjects (84) who reported a probabilistic confidence level lower than 50% for the maximum of "Unsure" (and close 0% for the minimum of "Unsure"). To clearly demonstrate the heterogeneity in the association between probabilistic confidence and confidence statements, we performed the analysis for these two groups separately. The first group was selected on the criterion that the maximum level of probabilistic confidence for "Unsure" was equal or larger than 50% and the minimum level of probabilistic confidence for "Unsure" was larger than 0%. For the second group, the criterion was that the maximum level of probabilistic confidence for "Unsure" was smaller than 50%. As we can see from Table B.8, in both groups there remained substantial heterogeneity in the association between probabilistic confidence and confidence statements. For example, the maximum level of probabilistic confidence for the statement "Probably" ranged from 75% to 99% in the first group and from 60% to 95% in the second group.

Probabilistic confidence associated with each confidence statement						
For subjects: Unsure (max) $\geq$ 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.8: 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.

B.2 Differences between randomization probabilities and self-reported decision confidence

Correspondence between decision confidence measures and binary choices

Intuitively, when the subjects report lower confidence in choosing an option, they would be less likely to choose that option. When decision confidence is linked to choices perfectly, we would expect that reported decision confidence of p% in choosing option x corresponds to option x being chosen in p% of the binary choices.

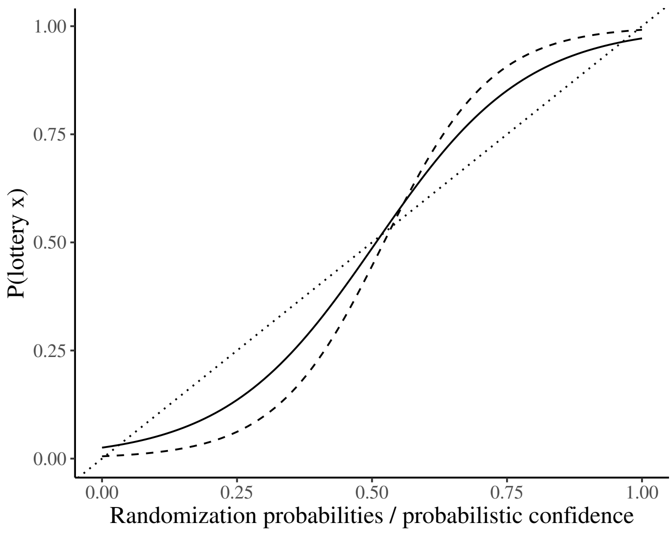


Figure B.5: The fitted line of the multilevel logistic regressions with the binary choice as the dependent variable and either the randomization probabilities or probabilistic confidence as the explanatory variable for the simple lottery in the baseline no-experience treatment. The dotted line represents perfect correspondence. The solid line is for randomization probabilities, and the dashed line is for probabilistic confidence.

Figure 5 suggests that both randomization probabilities and probabilistic confidence contain information about how the subjects made their binary choices. In addition, randomization probabilities appear to correspond more closely to binary choices, as suggested by the Fisher's exact tests in the figure. Multilevel logistic regression models with random intercepts for each participant to estimate the association show similar results. The regression results are presented in Table B.9 and are illustrated in Figure B.5.

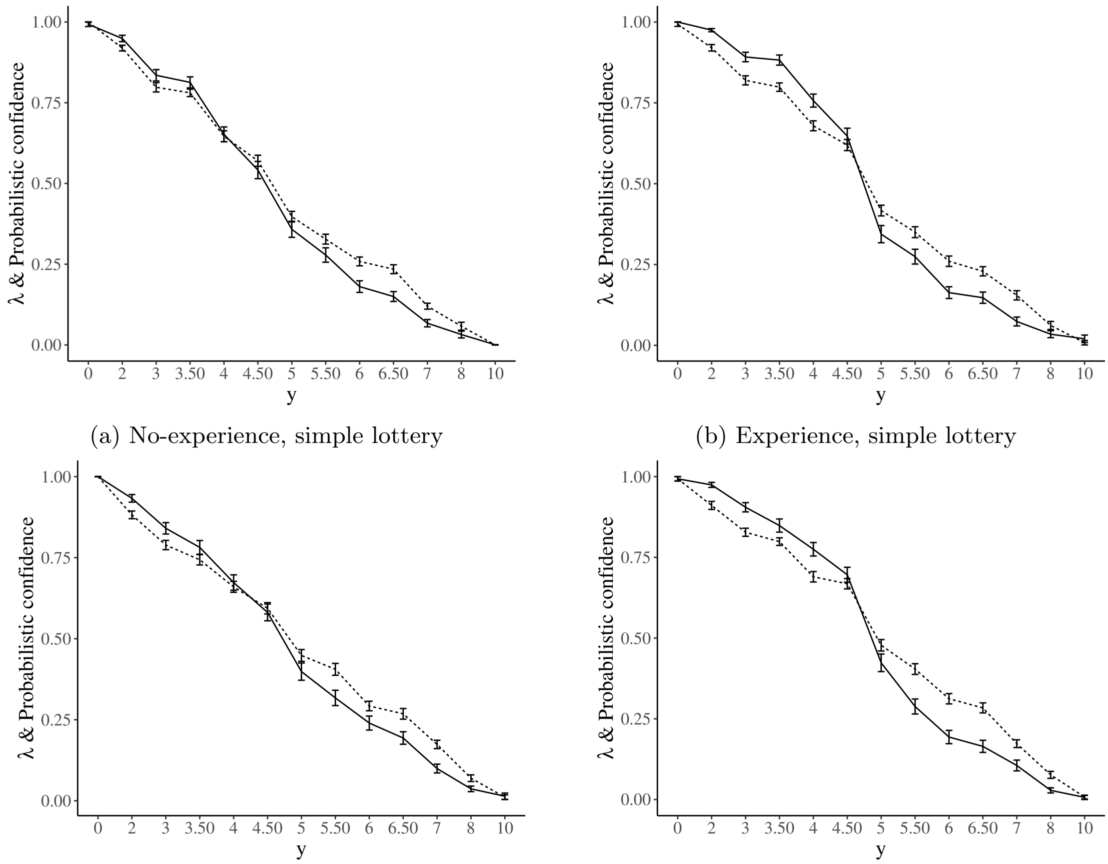
Relative sensitivity of randomization probabilities

Apart from its close correspondence with binary choices, we also find that randomization probabilities were more sensitive to variations in sure payments, including sure payments around the switching choices. Figure B.6 plots probabilistic confidence and randomization probabilities against sure payments y for each lottery and experience treatment. In all

	Randomization probabilities	Probabilistic confidence
Fixed Effects	Log odds	
Intercept	-3.648^{***}	-5.212^{***}
	(0.214)	(0.296)
Decision confidence	7.187^{***}	9.987^{***}
	(0.374)	(0.535)
Random Effects		
Subject intercept	0.797	0.658
	(0.893)	(0.811)

Table B.9: Results of the two multilevel logistic regressions with the binary choice as the dependent variable and either the randomization probabilities or probabilistic confidence as the explanatory variable. Including random-intercept at the subject level. The stars * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$. The values in the parentheses are the standard errors of the estimates.

four treatments, we find that probabilistic confidence was more compressed towards 50:50 compared to randomization probabilities. The gaps between probabilistic confidence and randomization probabilities were statistically significant for most of the sure payments, except in the simple lottery, no-experience treatment. We look at sure payments around the mean certainty equivalent of the lottery (between 4.5 and 5.07 euros for all lotteries and treatments), and we find a smaller difference in mean probabilistic confidence than difference in mean randomization probabilities between sure payments of 4.0 and 5.5 euros across all treatments. This was statistically significant for the full-experience treatment (Full-experience, simple lottery, 0.36 vs 0.48; Full-experience, complex lottery, 0.30 vs 0.34, Wilcoxon signed-rank tests: $p < 0.01$). This means that we are more likely to pick up changes in decision confidence around the certainty equivalent using randomization probabilities as the proxy for decision confidence compared to probabilistic confidence. One plausible explanation for why probabilistic confidence is less sensitive to changes in sure payment values is that there is only incentivizing effect on probabilistic confidence at the threshold of 50:50. In comparison, all values of randomization probabilities are incentivized, which may encourage subjects to reveal their precise decision confidence.



(a) No-experience, simple lottery

(b) Experience, simple lottery

(c) No-experience, complex lottery

(d) Experience, complex lottery

Figure B.6: The randomization probabilities, λ, (solid line) are more sensitive (steeper) around the switching choices than probabilistic confidence (dashed line). This implies with the same change in y, randomization probabilities change more than subjects' probabilistic confidence levels, making randomization probabilities better at capturing the small changes in decision confidence.

C Online appendix: Additional results

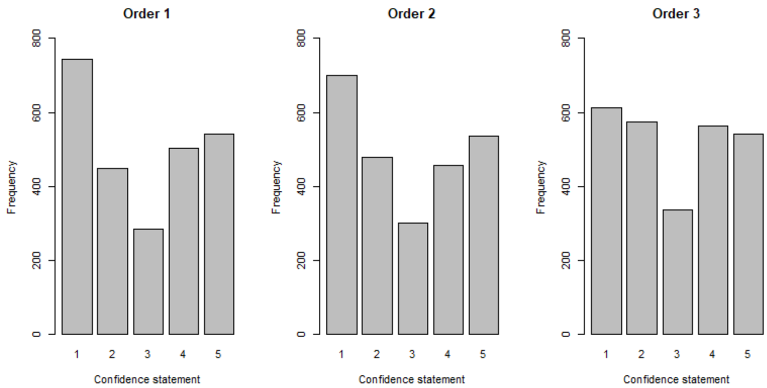
C.1 Order effects in experiment 2

To address the issue of experimenter demand effects due to the within-subject design, we elicited the experimental tasks in Experiment 2 with at least one week apart and in different orders. In each week, the subjects had to make only one type of decisions. The subjects were not informed about the order of their decisions; during each session, the subjects were only informed about the set of decisions they had to make for that session. Having the subjects make separate decisions each week reduces the likelihood that later decisions are made to cohere with earlier decisions since the subjects would have to remember many decisions in the earlier week(s) to do so.

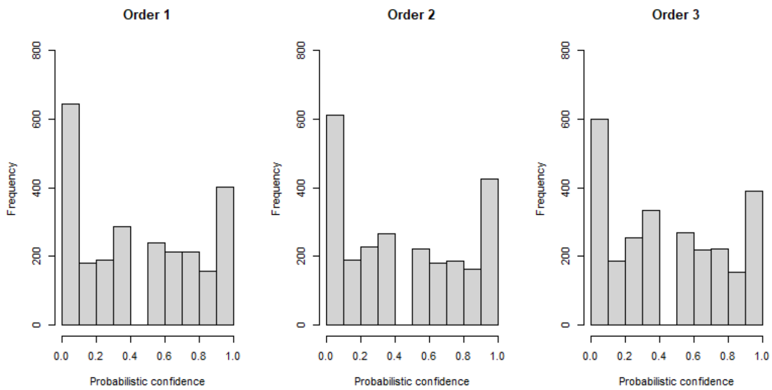
As a further safeguard against experimenter demand effects, we randomized the order in which the subjects completed the three decisions. To select the orders, we made sure that there was at least one order in which each of the three measures was elicited first, and, when possible, subjects did not have to report probabilistic confidence and randomization consecutively since the two measures bear close resemblance. We ended up with three orders: 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.

The main concern of order effects is whether the subjects' later decisions were affected by their earlier decisions. In particular, did the subjects randomize differently if they had to make randomization choices after confidence statements or probabilistic confidence because they were induced to think about decision confidence? To examine this possibility, we compare the proportion of decisions in which the subjects randomized when the randomized choices were made first versus when randomized choices were made after they completed other decision confidence measures.

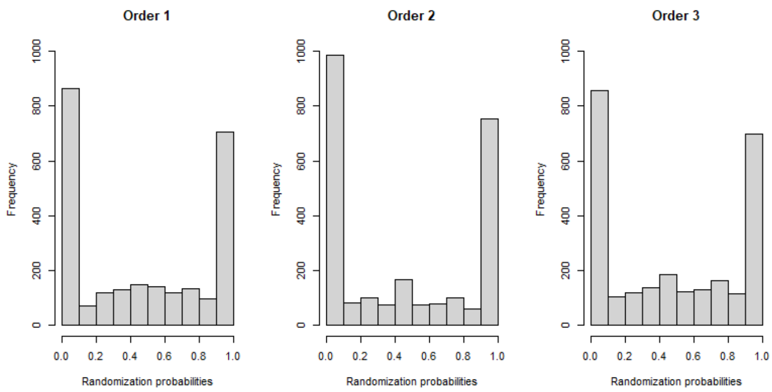
Figure C.1 reports the distribution of confidence statements, randomization probabilities



(a) Confidence statements



(b) Probabilistic confidence



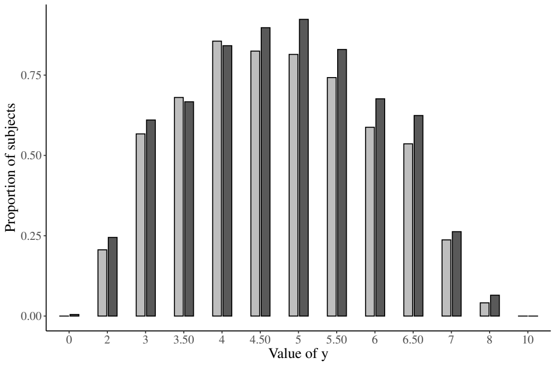
(c) Randomization probabilities

Figure C.1: The distribution of confidence statements (a), probabilistic confidence (b), and randomization probabilities (c) in each decision order. The three decision orders are: Order 1) binary choices and confidence statements $\rightarrow$ randomized choices $\rightarrow$ probabilistic confidence, Order 2) randomized choices $\rightarrow$ binary choices and confidence statements $\rightarrow$ probabilistic confidence, Order 3) probabilistic confidence $\rightarrow$ binary choices and confidence statements $\rightarrow$ randomized choices.

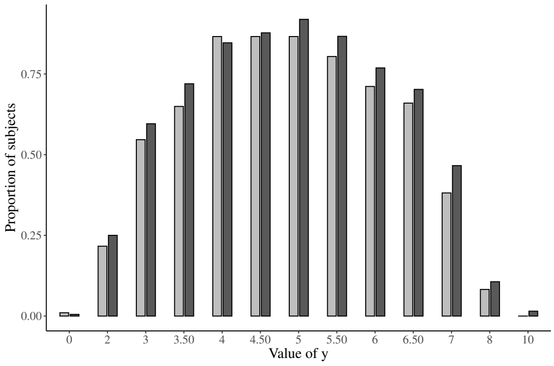
and probabilistic confidence across subjects for each order. Indeed, the subjects chose not to randomize in more decisions when randomization probabilities were asked first rather than later (Order 2: 65.8% of decisions with no randomization; Order 1 and 3: 57.0% of decisions with no randomization; Z-test of proportions: $p < 0.01$). Interestingly, we observe similar order effects for confidence statements in the sense that the subjects reported full confidence in more decisions (chose "Surely x" or "Surely y") when they reported their confidence statements first than later (Order 1: 51.0% of decisions with full confidence; Order 2 and 3: 46.9% of decisions with full confidence; Z-test of proportions: $p < 0.01$). However, we see the opposite effect for probabilistic confidence. The subjects had full confidence in fewer decisions when they reported their probabilistic confidence first rather than later (Order 3: 31.1% of decisions with full confidence; Order 1 and 2: 35.9% of decisions with full confidence; Z-test of proportions: $p < 0.01$). Figure C.2 illustrates the above comparisons at each value of y. Consistent with these results, on average the proportions of subjects who randomized or reported less than full confidence is lower when randomized choices or confidence statements were elicited first, while the opposite is true for probabilistic confidence. These differences were present across most sure payment values. Taken together, while Figure C.1 and Figure C.2 reveal some differences in the subjects' reports depending on whether a measure is elicited first or later, there is no consistent pattern of order effects across the three measures.

While the above analyses reveal some order effects, these order effects are unlikely to be the main driver of our results. Table C.1 shows the 10th percentile, the median, and the 90th percentile nonparametric Spearman correlations between the randomization probabilities and the self-reported decision confidence measures in each decision order for each lottery and treatment. While there were some differences (e.g., the 10th percentile correlations are sometimes lower in Order 2 than in other orders), median correlations were high (above 0.8) and comparable across all orders. This suggests that randomization probabilities were not chosen simply to cohere with the earlier reported decision confidence, and the high correlation between randomization probabilities and self-reported confidence measures is not purely driven by order effects.

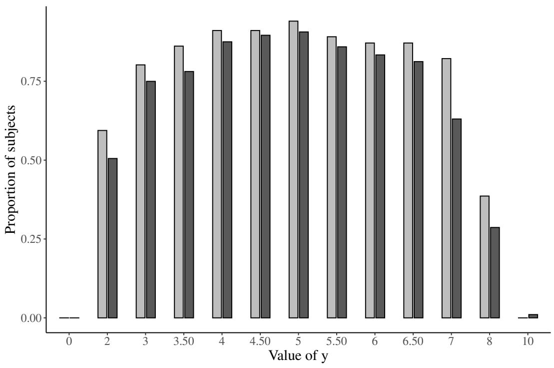
We further compared the correspondence between randomization probabilities and proba-



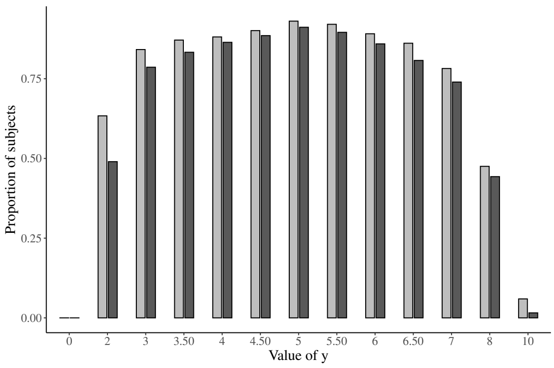
(a) Confidence statements: Simple lottery



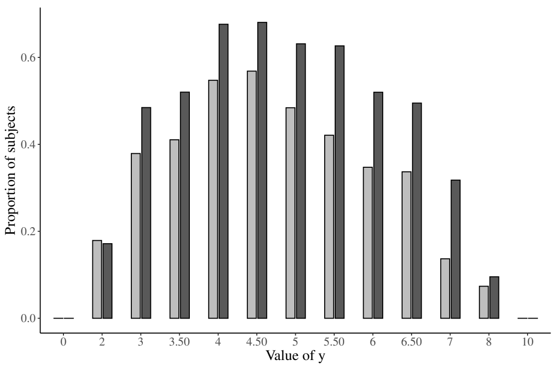
(b) Confidence statements: Complex lottery



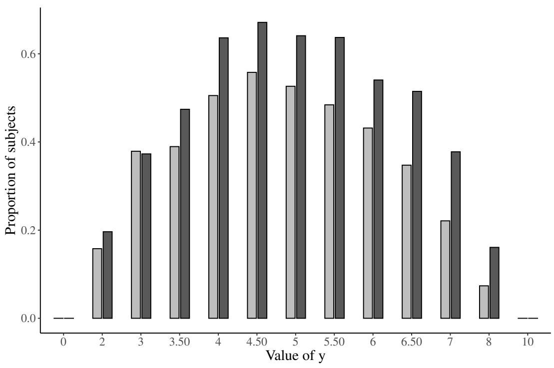
(c) Probabilistic confidence: Simple lottery



(d) Probabilistic confidence: Complex lottery



(e) Randomization probabilities: Simple lottery



(f) Randomization probabilities: Complex lottery

Figure C.2: Proportions of subjects who are not completely confident (confidence statements "probably x/y" or "unsure", probabilistic confidence and randomization probabilities between 0.1 and 0.9) for each value of y. The light (dark) gray bars represent the order in which a measure is elicited first (later, respectively).

bilistic confidence in absolute levels when they were asked in different orders. Table C.2 reports these comparisons. While the mean randomization probabilities were lower when they were reported first (in Order 2) than later across a wide range of self-reported probabilistic confidence, mean randomization probability for x remained broadly similar to the probabilistic confidence for x in absolute terms for most values of probabilistic confidence. These findings further support that our results are not mainly driven by order effects.

Finally, we show that randomization probabilities responded to the manipulations of decision confidence in similar ways regardless of the order they were asked. In Tables C.3 and C.4 we report the differences in the size of the ranges of sure payments over which the subjects indicated that they were not completely confident about their decision in the lottery and experience treatments. While there were some differences in the size of the treatment effects across different orders, the differences were mostly in the intended direction. Importantly, although we find that treatment effects on randomization probabilities are weaker in Order 1, the sequence of decisions in Order 1 (binary choices and confidence statements $\rightarrow$ randomized choices $\rightarrow$ probabilistic confidence) was similar to the decision order in Experiment 1 (binary choices and confidence statements $\rightarrow$ randomized choices), where we find strong support that randomization probabilities are good proxy for decision confidence. Taken together, we conclude that order effects are not the main driver behind our findings.

			Correlation between randomization probabilities and confidence statements			
Lottery	Treatment		Combined	Order 1	Order 2	Order 3
Simple	No experience	10th percentile	0.71	0.73	0.71	0.70
		Median	0.89	0.88	0.89	0.89
		90th percentile	0.96	0.96	0.96	0.95
	Experience	10th percentile	0.78	0.80	0.58	0.74
		Median	0.90	0.90	0.89	0.93
		90th percentile	0.97	0.97	0.96	0.97
Complex	No experience	10th percentile	0.67	0.73	0.57	0.72
		Median	0.88	0.88	0.88	0.89
		90th percentile	0.96	0.96	0.95	0.95
	Experience	10th percentile	0.69	0.73	0.59	0.79
		Median	0.90	0.90	0.87	0.92
		90th percentile	0.97	0.97	0.97	0.96
			Correlation between randomization probabilities and probabilistic confidence			
Lottery	Treatment		Combined	Order 1	Order 2	Order 3
Simple	No experience	10th percentile	0.73	0.78	0.75	0.72
		Median	0.90	0.94	0.87	0.90
		90th percentile	0.98	0.99	0.96	0.98
	Experience	10th percentile	0.77	0.79	0.77	0.81
		Median	0.91	0.91	0.89	0.92
		90th percentile	0.99	0.99	0.97	0.98
Complex	No experience	10th percentile	0.64	0.71	0.59	0.68
		Median	0.89	0.90	0.89	0.89
		90th percentile	0.97	0.99	0.97	0.97
	Experience	10th percentile	0.77	0.80	0.68	0.81
		Median	0.90	0.91	0.88	0.91
		90th percentile	0.97	0.98	0.97	0.97

Table C.1: Nonparametric Spearman correlations by decision order across lottery and experience treatments at the 10th percentile, 50th percentile, and 90th percentile in Experiment 2. The three orders are: 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.

	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
	Combined									
Rand. prob.	0.97 (0.005)	0.89 (0.013)	0.83 (0.016)	0.74 (0.017)	0.60 (0.019)	0.36 (0.017)	0.23 (0.017)	0.17 (0.015)	0.08 (0.012)	0.02 (0.005)
	Order 1									
Rand. prob.	0.97 (0.012)	0.89 (0.025)	0.83 (0.024)	0.73 (0.029)	0.64 (0.031)	0.37 (0.029)	0.28 (0.032)	0.18 (0.028)	0.12 (0.026)	0.02 (0.006)
	Order 2									
Rand. prob.	0.98 (0.005)	0.86 (0.026)	0.83 (0.032)	0.72 (0.035)	0.54 (0.040)	0.31 (0.035)	0.16 (0.027)	0.15 (0.028)	0.05 (0.014)	0.02 (0.007)
	Order 3									
Rand. prob.	0.97 (0.010)	0.91 (0.016)	0.83 (0.029)	0.78 (0.022)	0.62 (0.028)	0.39 (0.027)	0.25 (0.026)	0.17 (0.023)	0.08 (0.019)	0.02 (0.011)

Table C.2: The mean randomization probabilities at each self-reported probabilistic confidence level in Experiment 2 in the baseline no-experience treatment for decisions about the simple lottery on the aggregate and in each order separately. The standard errors of the mean are reported in the parentheses.

	Lottery	Combined	Order 1	Order 2	Order 3
Confidence Statements	Simple	3.15	2.96	2.77	3.68
	Complex	3.58***	3.25 *	3.56***	3.93
Probabilistic confidence	Simple	4.37	4.28	4.06	4.72
	Complex	4.63***	4.34	4.61***	4.93*
Randomization probabilities	Simple	3.09	3.17	2.67	3.14
	Complex	3.19*	3.09	2.89*	3.51**

Table C.3: Sizes of the ranges of sure payments over which subjects express that they are not fully confident for each of the confidence measures. The three orders are: Order 1) binary choices and confidence statements $\rightarrow$ randomized choices $\rightarrow$ probabilistic confidence, Order 2) randomized choices $\rightarrow$ binary choices and confidence statements $\rightarrow$ probabilistic confidence, Order 3) probabilistic confidence $\rightarrow$ binary choices and confidence statements $\rightarrow$ randomized choices. Wilcoxon signed-rank tests were performed to test the difference between the simple lottery and the complex lottery: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	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 C.4: The 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 and decision order. The three orders are: 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. Wilcoxon rank-sum tests were performed to test the difference between the no experience and experience treatment: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.2 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.

D Online appendix: Experimental materials

D.1 Reducing demand effects through decision order and cost for randomizing

We used different approaches in our two experiments to address the concern that experimenter demand effects drive a systematic relationship between the self-reported decision confidence measures and the randomized choices.

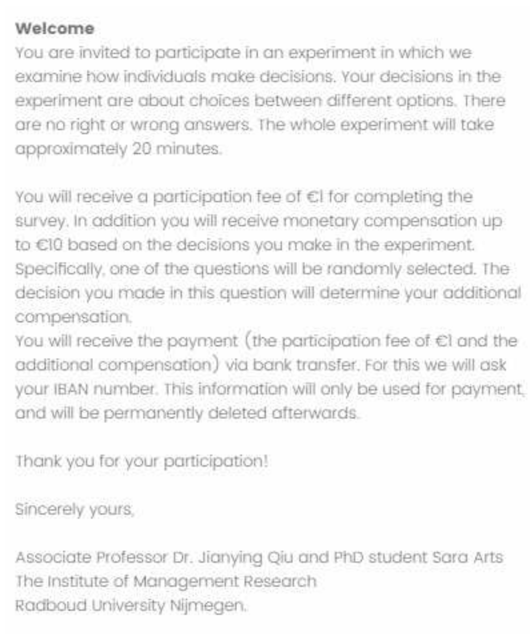
In Experiment 1, the subjects completed all their decisions within a single experimental session. The subjects first completed the binary choices and confidence statements for *all* pairs of the lottery and the sure payment before making the randomized choices. Although the order of the type of decisions (binary choice and confidence statements → randomized choices) was fixed in Experiment 1, the subjects made their decisions on a lottery and 13 different values of sure payment in a random sequence for each type of decisions. They were not allowed to make changes to the decisions they had already made. This was designed to make it difficult for the subjects to link their randomized choice to their earlier binary choice and confidence statement for each pair of options. We also randomized the order of the lottery treatments to reduce any order effects: some subjects proceeded from the simple lottery to the complex lottery, while others completed the decisions in the reverse order.

In Experiment 2, we made further attempts to disconnect decisions regarding self-reported confidence and randomization probabilities by spreading the decisions over three sessions (seven days apart) and by introducing a cost to randomize. In each session, the subjects only made one set of decisions, either 1) binary choices and confidence statements, 2) randomized choices, or 3) probabilistic confidence choices. The order of their decisions across the three sessions was determined randomly by the computer.[15] The subjects were not informed about the order of their three sets of decisions. During each session, the

[15] The order of the decisions was either 1) binary choices and confidence statements → randomized choices → probabilistic confidence, 2) randomized choices → binary choices and confidence statements → probabilistic confidence, 3) probabilistic confidence → binary choices and confidence statements → randomized choices.

subjects were only informed about the set of decisions they had to make for that session, and they made decisions on each lottery and 13 different values of sure payment in a random sequence. They were not allowed to refer to or change the choices they had made in the earlier session(s). Apart from decision order, we also introduced a cost for randomization in Experiment 2: the subjects in had to pay a fixed cost of 0.10 euros if they would like to choose a randomization probability other than 0% or 100%. In comparison, randomized choices were free of charge in Experiment 1. These new features in Experiment 2 made it more obscure and costly for the subjects to connect self-reported decision confidence and randomization probabilities in response to experimenter demand effects (Zizzo, 2010).

D.2 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.

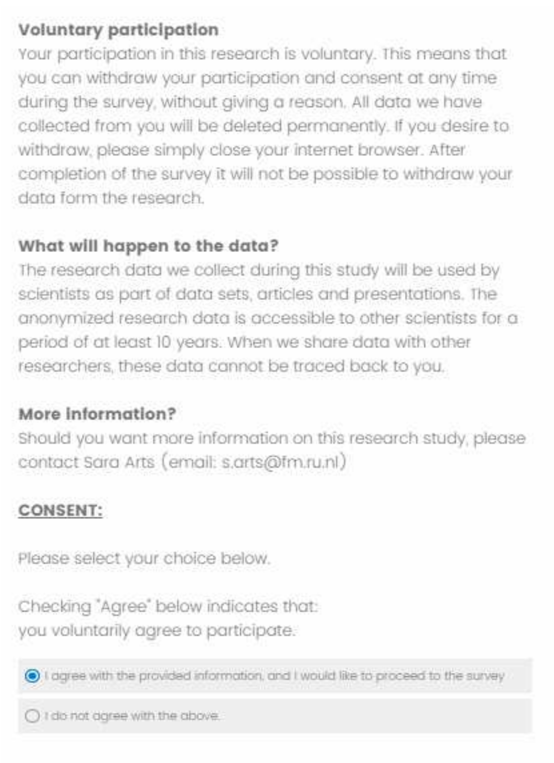
You will receive a participation fee of €1 for completing the survey. In addition you will receive monetary compensation up to €10 based on the decisions you make in the experiment. Specifically, one of the questions will be randomly selected. The decision you made in this question will determine your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

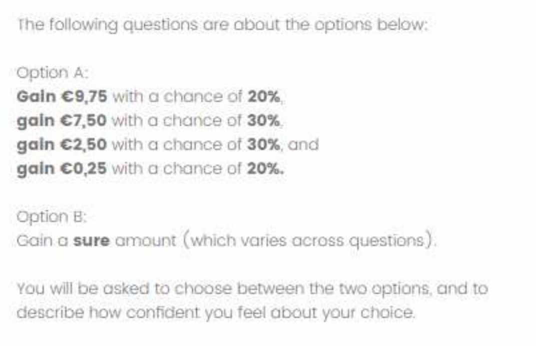
Checking "Agree" below indicates that:
you voluntarily agree to participate.

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

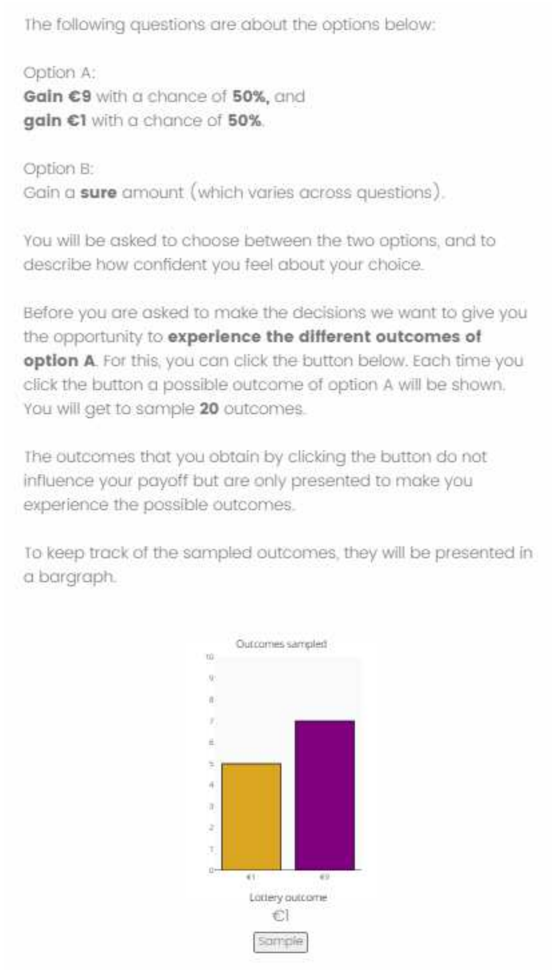
I do not agree with the above.

(b)

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

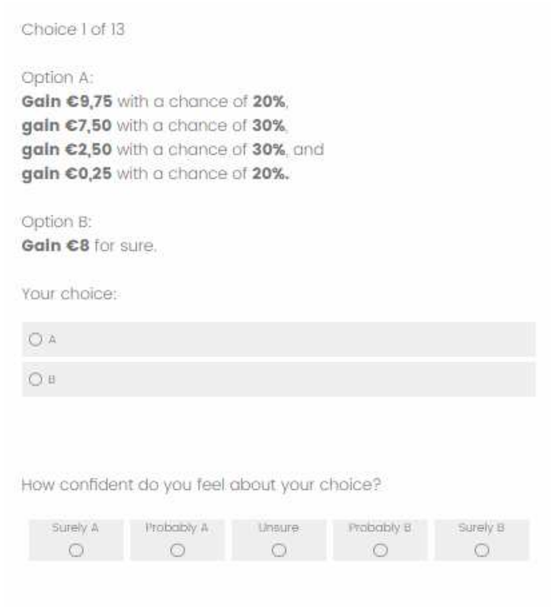


(a)

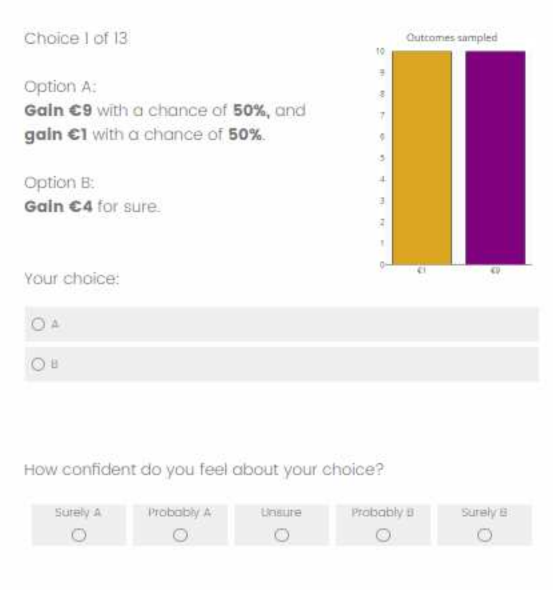


(b)

Figure D.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 D.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).

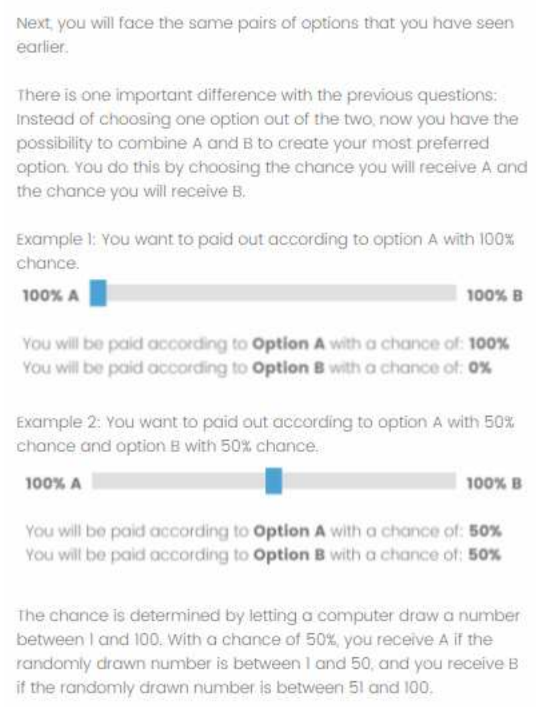
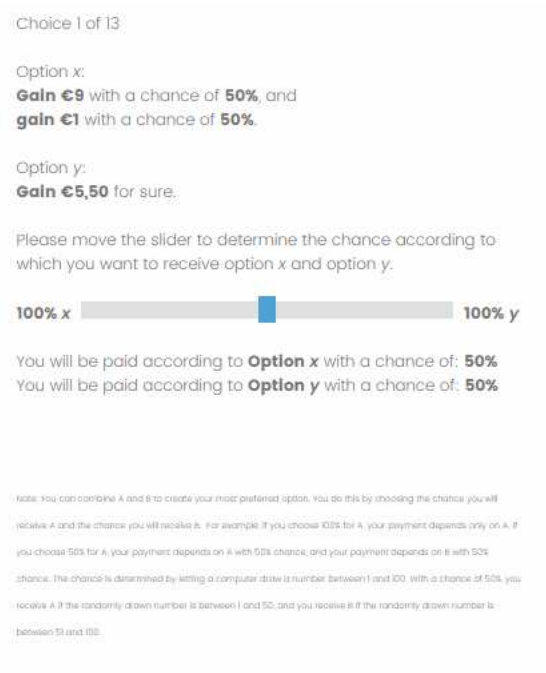


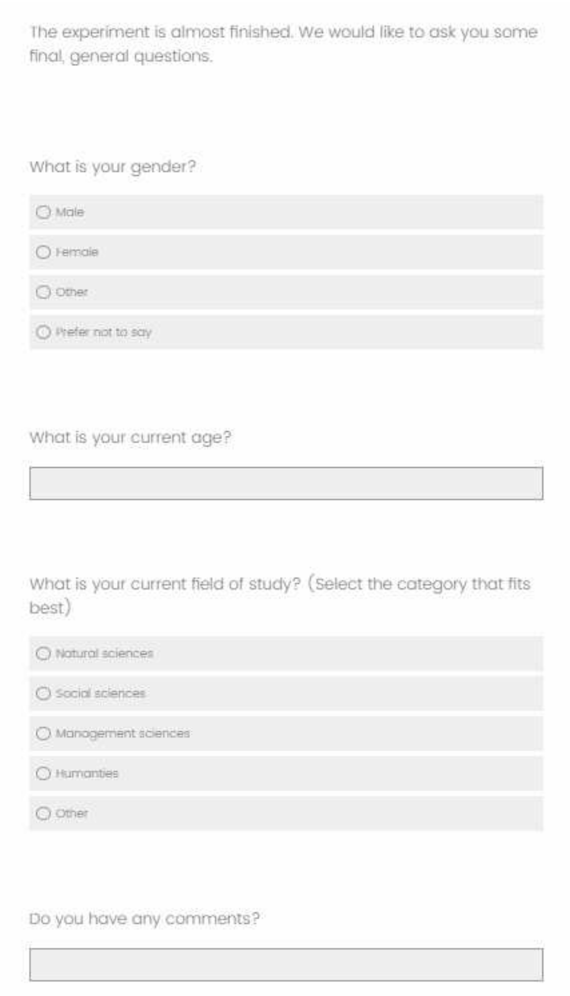
Figure D.4: Explanation of the randomized choices.



(a)

(b)

Figure D.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?

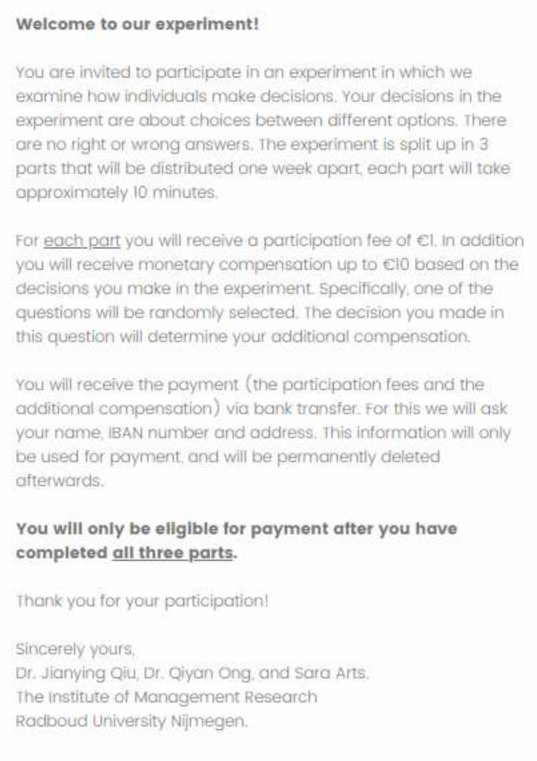
What is your current field of study? (Select the category that fits best)

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

Do you have any comments?

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

D.3 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.

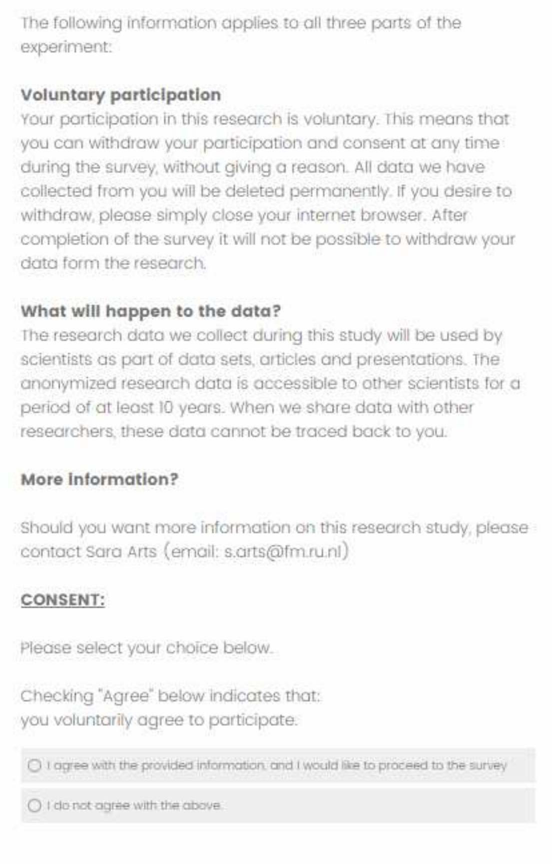
You will receive the payment (the participation fees and the additional compensation) via bank transfer. For this we will ask your name, IBAN number and address. This information will only be used for payment, and will be permanently deleted afterwards.

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.

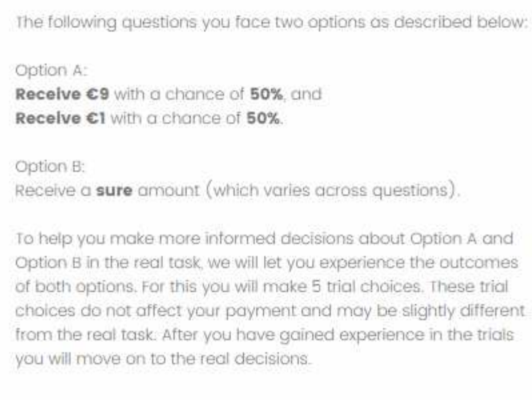
Checking "Agree" below indicates that:
you voluntarily agree to participate.

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

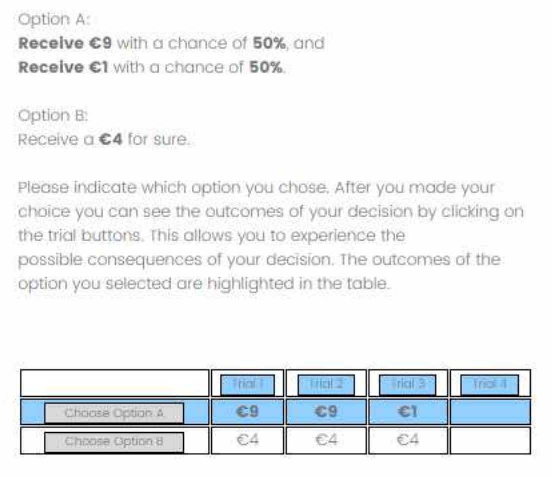
○ I do not agree with the above.

(b)

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

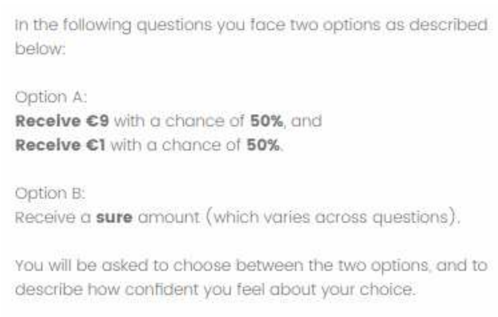


(a)

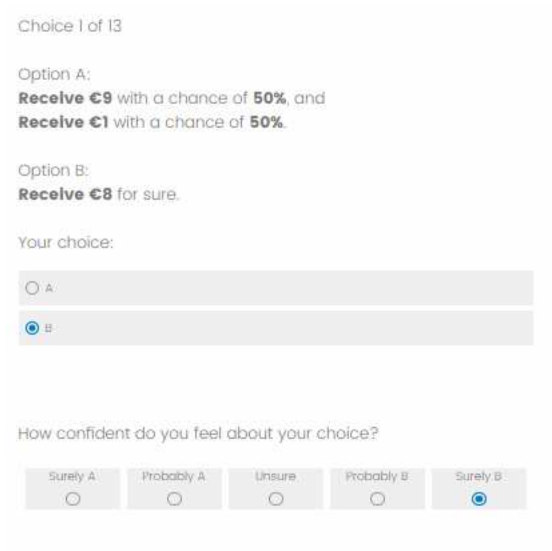


(b)

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

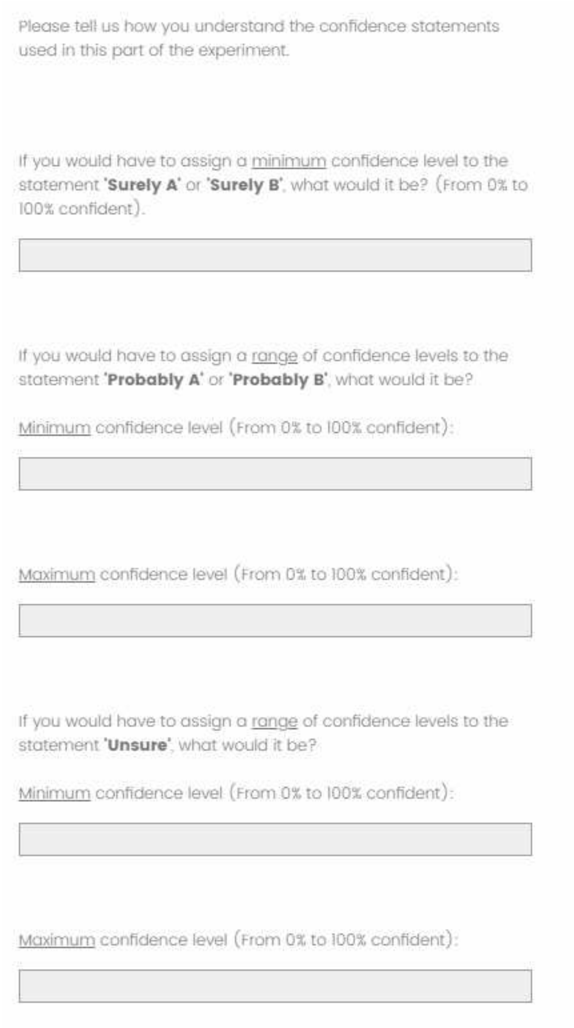


(a)



(b)

Figure D.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):

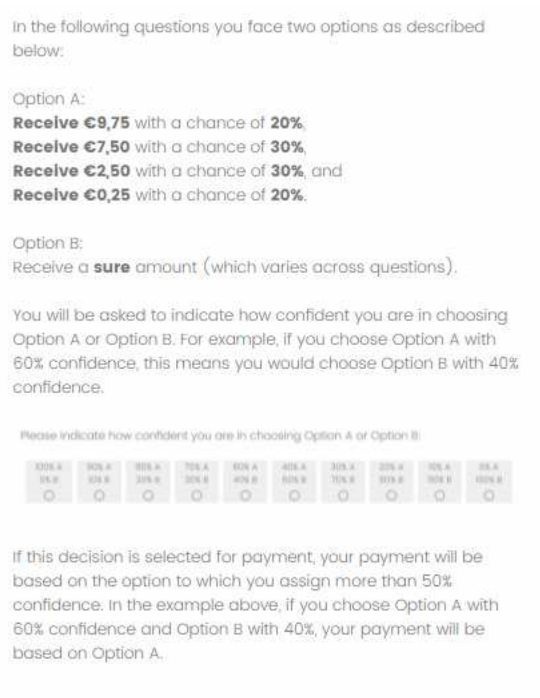
Maximum confidence level (From 0% to 100% confident):

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



(a)

(b)

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

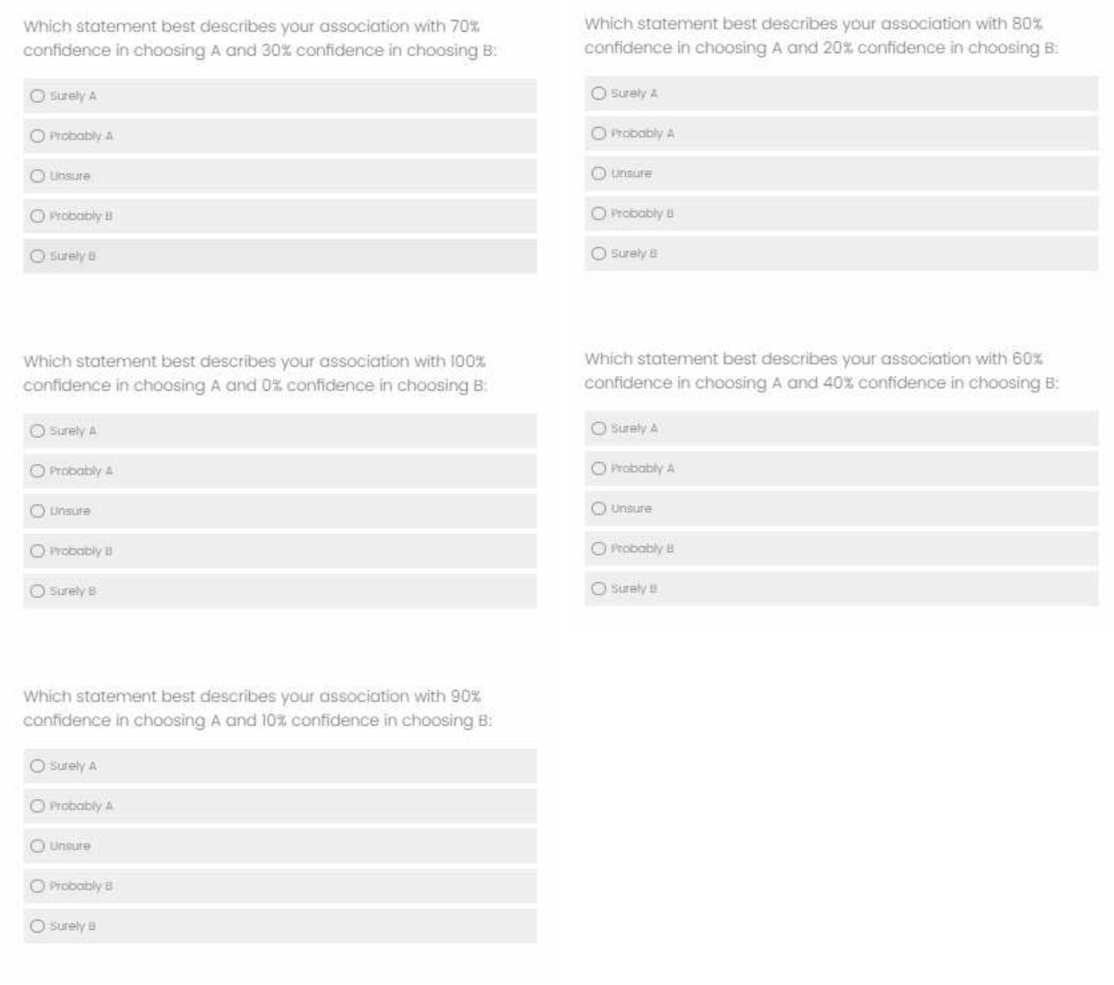
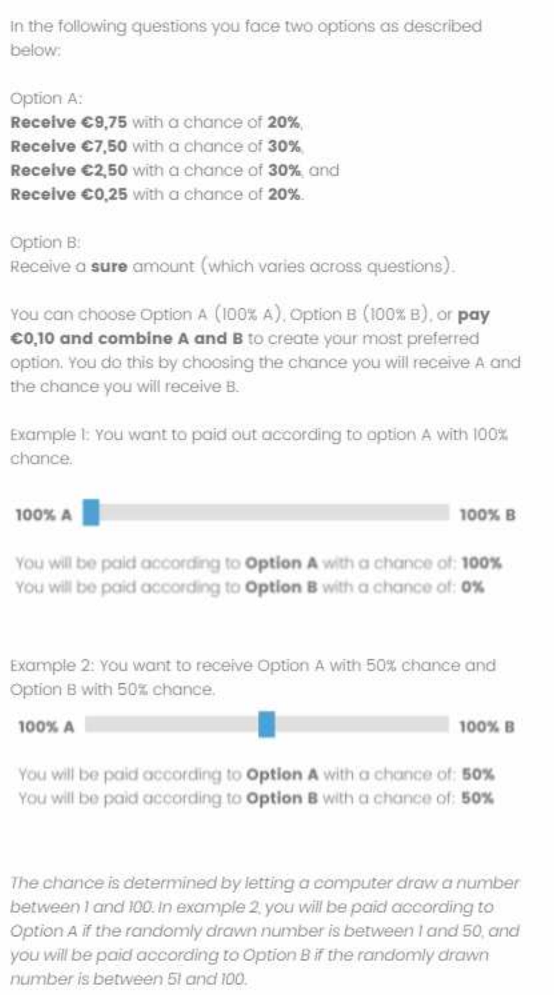


Figure D.12: post-experiment questionnaire after the probabilistic confidence choices.



(a)

(b)

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

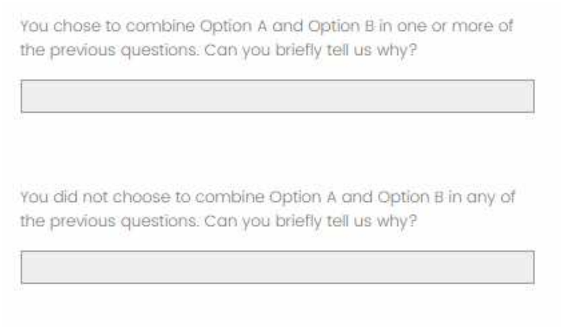


Figure D.14: post-experiment questionnaire after the randomized choices. The first question was asked if a subject chose randomization probabilities other than 0 or 1 in at least 1 choice. The second question was asked if a subject only chose randomization probabilities of 0 or 1.

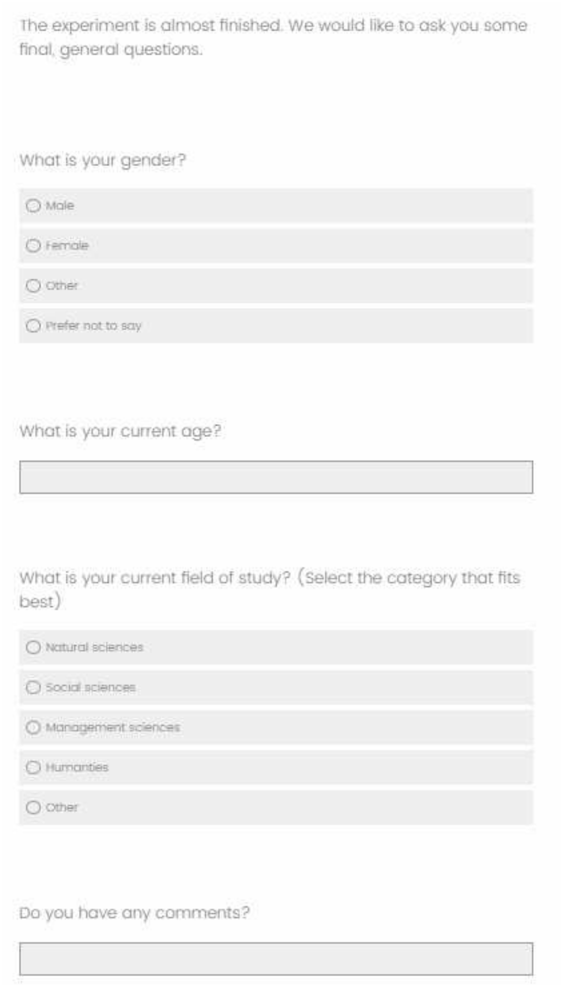


Figure D.15: Demographic questions asked at the end of the experiment.