Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2019, Vol. 30(8) 1218–1233

DOI: 10.1177/0956797619858969
www.psychologicalscience.org/PS
SAGE

What the Future Holds and When: A Description–Experience Gap in Intertemporal Choice

TC

Junyi Dai[1,2], Thorsten Pachur[2], Timothy J. Pleskac[2,3], and Ralph Hertwig[2]
[1]Department of Psychology and Behavioral Sciences, Zhejiang University; [2]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany; and [3]Department of Psychology, University of Kansas

Abstract
Uncertainty about the waiting time before obtaining an outcome is integral to intertemporal choice. Here, we showed that people express different time preferences depending on how they learn about this temporal uncertainty. In two studies, people chose between pairs of options: one with a single, sure delay and the other involving multiple, probabilistic delays (a lottery). The probability of each delay occurring either was explicitly described (*timing risk*) or could be learned through experiential sampling (*timing uncertainty*; the delay itself was not experienced). When the shorter delay was rare, people preferred the lottery more often when it was described than when it was experienced. When the longer delay was rare, this pattern was reversed. Modeling analyses suggested that underexperiencing rare delays and different patterns of probability weighting contribute to this description–experience gap. Our results challenge traditional models of intertemporal choice with temporal uncertainty as well as the generality of inverse-S-shaped probability weighting in such choice.

Keywords
intertemporal choice, description–experience gap, temporal uncertainty, rank-dependent discounted-utility model, hierarchical Bayesian modeling, open data, open materials, preregistered

Received 4/28/17; Revision accepted 5/12/19

Imagine that you want to improve future business trips by ordering noise-canceling headphones online. You narrow your options down to two. Will you order the pair that could be delivered tomorrow, or will you wait 2 weeks for your favorite pair to arrive? Standard economic models of such *intertemporal choices* suggest that you should choose as if the value of each option is discounted as a function of its delay. However, choices between options with delayed outcomes are rarely this straightforward. Seldom does one know precisely when the outcome will materialize. For instance, if your preferred headphones need to be shipped to Germany from the United States via DHL, the average delivery time is 4.9 days, but the actual waiting period varies.[1] There is an 86.3% chance that the package will arrive within a week, a 9.8% chance that it will take 7 to 15 days, and a 2% chance that it will take 16 to 30 days. This decision is thus made under *timing risk*, in which the probability of each possible delay is known (e.g., Chesson & Viscusi, 2003).

But even decisions involving timing risk are likely to represent only a small proportion of the intertemporal choices people face. Instead, in most choices—such as a decision to bide one's time in an unsatisfying job in the hope of getting promoted—the possible delays, let alone their probability of occurring, are likely to be vaguely known at best. To make these choices, people not only have to somehow consider the likelihoods of the delays—they must also learn about the possible

Corresponding Author:
Junyi Dai, Zhejiang University, Department of Psychology and Behavioral Sciences, No. 148 Tianmushan Rd., Hangzhou, 310028, China
E-mail: junyidai@zju.edu.cn

delays and their probabilities in the first place. These decisions are made under *timing uncertainty*.

How do people make intertemporal choices when facing timing uncertainty? In this pair of studies, we investigated this question by studying intertemporal choice from experience, in which people learned about the probabilities of different delays from experiential sampling (the delays themselves were not experienced). We first showed that intertemporal choice from description (involving timing risk with a stated probability distribution of possible delays) and from experience (involving timing uncertainty and thus requiring learning from sampled information) led to qualitatively different preferences. Second, we demonstrated that this new manifestation of a *description–experience gap* (Hertwig & Erev, 2009) was accompanied by distinct probability-weighting patterns in the description and experience conditions. Third, we showed that these results challenge assumptions in current theories of intertemporal choice under temporal uncertainty.[2] Note that in this research we examined stated, rather than experienced, delays under both timing risk and timing uncertainty.

Intertemporal Choice With Temporal Uncertainty

The normative economic model for intertemporal choices under temporal certainty is the discounted-utility model (Samuelson, 1937). Standard economic theory (von Neumann & Morgenstern, 1947) suggests that in the face of temporal uncertainty, this model is modified to the discounted-expected-utility (DEU) model (e.g., Andreoni & Sprenger, 2012b; Chesson & Viscusi, 2003). According to the DEU model, when people face options with different possible delays, they behave as if they maximize the weighted average discounted utility of the chosen option, with the probability of occurrence of the relevant delays serving as the weights. For instance, consider the choice between a *timing lottery* that offers €160 and a *sure-timing option* offering the same amount of money. The timing lottery will deliver the money either in 1 month with a probability of .8 or in 11 months with a probability of .2; in contrast, the sure-timing option will deliver the money in 3 months for sure. According to the DEU model, the value or weighted average discounted utility of the sure-timing (ST) option is simply its discounted utility,

$$V_{\mathrm{ST}} = u(160) \times d(3),$$

where $u(x)$ is a utility function specifying the subjective value of the possible reward x and $d(t)$ is a discount function that determines the degree of discounting produced by a delay t. The value of the timing lottery (TL) is its weighted average discounted utility,

$$V_{\mathrm{TL}} = .8 \times u(160)d(1) + .2 \times u(160)d(11).$$

According to the DEU model, one would prefer the sure-timing option if V_{ST} were greater than V_{TL} and would prefer the timing lottery if V_{ST} were less than V_{TL}. But is this model a good descriptive account of people's actual choices? A close inspection reveals a notable prediction of the DEU model: With a convex discount function (e.g., an exponential or a hyperbolic function), the model always predicts a preference for the timing lottery if the lottery has the same reward and expected delay as the sure-timing option (for a proof, see Onay & Öncüler, 2007; see also the Supplemental Material available online). Yet people do not always prefer timing risk. Instead, when the longer delay in the lottery option is less probable than the shorter delay, people tend to be risk averse, choosing the sure-timing option. In contrast, when the timing lottery is modified such that the longer delay is more probable than the shorter delay, people tend to prefer the timing lottery and thus are risk seeking (Onay & Öncüler, 2007).

To account for this choice pattern, Onay and Öncüler (2007) proposed a rank-dependent discounted-utility (RDDU) model. According to this model, people make intertemporal choices involving timing risk as if they weighted the discounted utility at each delay such that rare delays were overweighted, giving them more impact than they deserve (relative to the delays' stated probability). The RDDU model thus closely aligns with cumulative-prospect theory's account of risky choice (Tversky & Kahneman, 1992) and its assumed probability-weighting pattern. As we explain next, this commonality raises the key question of whether timing risk and timing uncertainty will prompt systematically different intertemporal choices akin to the description–experience gap in risky choice.

A Description–Experience Gap in Intertemporal Choice?

Risky choice involves a trade-off between two or more options with uncertain payoffs. In studies on risky choice, payoffs and probabilities have traditionally been explicitly stated, thus requiring people to make decisions from these descriptions. In decisions from experience, in contrast, these properties must be learned—for instance, by sampling from the payoff distributions. Numerous studies of risky choice have revealed a systematic gap between decisions from description and decisions from experience (for reviews, see Hertwig, 2015; Rakow & Newell, 2010). The description–experience gap refers to the finding that, for example, when offered a choice between a sure option (e.g., €3) and a risky option involving a rare gain (e.g., winning €32 with a probability of .1 and otherwise nothing), people are more likely to choose the sure option in the experience than in the description condition. This difference between description and experience, obtained in choices involving a risky and a safe option, reverses when the risky option involves a likely gain.

The description–experience gap results partly from sampling error caused by limited exploration in decisions from experience (see Wulff, Mergenthaler-Canseco, & Hertwig, 2018, for this and other contributors). Consequently, payoffs' objective probabilities might differ systematically from the relative frequencies that people actually experienced, especially for rare events, which tend to occur even less frequently in small samples. Furthermore, modeling analyses suggest that description and experience evoke distinct patterns of probability weighting: Rare events tend to be weighted differently between description and experience, no matter whether the weight analysis in experience is premised on the objective probabilities (Hertwig & Pleskac, 2018; Regenwetter & Robinson, 2017) or on the experienced (sampled) relative frequencies (Kellen, Pachur, & Hertwig, 2016).

Might there also be a description–experience gap in intertemporal choice under temporal uncertainty? If so, is it caused by mechanisms similar to those in risky choice? Although some research suggests commonalities between intertemporal and risky choices (e.g., Halevy, 2008; Luckman, Donkin, & Newell, 2018; Prelec & Loewenstein, 1991; Sozou, 1998; Takahashi, Ikeda, & Hasegawa, 2007; Walther, 2010), they also differ in important respects. For instance, Green, Myerson, and Ostaszewski (1999) found that reward size had opposite effects on the apparent discounting of delayed versus probabilistic (risky) outcomes, suggesting distinct decision mechanisms. Other economic analyses suggest distinct utility functions for risk and time preferences (Andreoni & Sprenger, 2012a, 2012b; Miao & Zhong, 2015). In addition, intertemporal choice under temporal uncertainty is more complex than risky choice in that the former consists of three elements—payoffs, delays, and their probabilities—rather than only payoffs and probabilities. Therefore, it is currently unclear whether a description–experience gap would emerge in intertemporal choice with temporal uncertainty and, if it did, what its causes would be.

Study 1

In Study 1, we sought to establish the existence of a description–experience gap in intertemporal choice. To this end, we adapted Onay and Öncüler's (2007) design, which focused on intertemporal choice with timing risk, and expanded it to timing uncertainty. This new design also allowed us to directly contrast choices from description (risk) and from experience (uncertainty) and to examine the extent to which sampling error and distinct patterns of probability weighting contributed to a possible gap. Finally, we varied the reward size to investigate its impact on choice in each condition.

Method

Participants. We recruited 124 adults (76 women; mean age = 25.31 years, *SD* = 3.90 years) from a subject pool of the Max Planck Institute for Human Development in Berlin, Germany. Half of the participants ($n = 62$) were randomly assigned to the description condition (40 women; mean age = 25.06 years, *SD* = 3.63 years) and the other half to the experience condition (36 women; mean age = 25.56 years, *SD* = 4.17 years). Because we aimed to extend previous findings of Onay and Öncüler (2007), we chose a slightly larger sample size—60 for each condition, for a total of 120 participants compared with their 50 participants for a single description condition. The recruitment process stopped when the planned sample size was reached and when all participants present on the final day of testing had completed the study. Participants signed a consent form approved by the institute's ethical review board before starting the study and were paid €7 for participation.

Materials and procedure. In the *description condition*, we implemented the same setup as in Onay and Öncüler (2007). Figure 1a presents a screenshot from the choice task (translated from German). At the beginning of the study, participants were asked to imagine that they had won some prize money and had to choose between different payout schedules. On each trial, participants were asked to choose between a sure-timing option and a timing lottery, with the two possible delays for the latter being set at 1 and 11 months, respectively. The probability of the longer delay was set at .1, .5, or .8, and the delay of the sure-timing option was set to be equal to the expected delay of the timing lottery (i.e., 2, 6, or 9 months). Participants were told that if they chose the timing lottery, the temporal uncertainty would be resolved immediately after they had made their choices. In addition, we used three reward sizes—€60, €300, and €1,500—to investigate the impact of reward magnitude on choices. Both options in each choice problem offered the same fixed (hypothetical) reward. Overall, each participant was asked to choose between nine pairs of options in a 3 (expected delay) × 3 (reward size) within-subjects design. The positions of the sure-timing option and the timing lottery on the screen

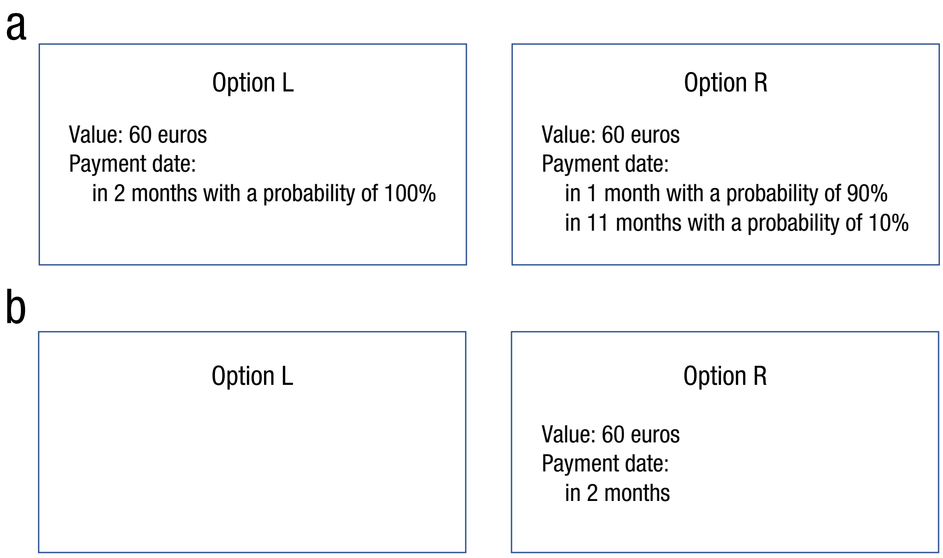

Fig. 1. Screenshots from the choice task in Study 1 (translated from German). In the description condition (a), participants were asked to imagine that they had won some prize money and to choose between two payout schedules (Option L or Option R). In the experience condition (b), the reward sizes and probabilities of delays were the same as in the description condition, but participants had to learn about them via experiential sampling. In the example shown here, the participant has clicked the right box, so information only for Option R is shown.

were randomized across trials. Finally, we elicited participants' present values and certain timing equivalents of the various options (these responses were not considered further; see the Supplemental Material for a detailed explanation). Participants familiarized themselves with the choice task in a practice session with three example questions. The reward sizes and probabilities in the practice trials differed from those in the test trials. The order of test trials was randomized for each participant.

In the *experience condition*, the choice options were the same as in the description condition in terms of reward sizes and objective probabilities, but participants had to learn about the reward size and possible delays of each option, as well as the probabilities of the delays, via experiential sampling (Hertwig & Erev, 2009). Before making a choice, participants were instructed to sample from each option as often as they wanted. Each sample revealed the reward and a single possible delay, which was drawn randomly using the objective probabilities. Each trial started with a fixation cross displayed at the center of the screen for 1 s before two blank boxes (with option labels) were shown. The participant could then sample from either the left or the right option by pressing the "Q" or "P" key, respectively. Each sample was visible for 1 s, after which another sample could be drawn. Figure 1b presents a screenshot from the sampling phase in this condition. The positions of the sure-timing option and the timing lottery on the screen were randomized across trials. In a practice phase, participants had to sample at least five times from each option in order to familiarize themselves with the probabilistic nature of the sampling process. In the subsequent test trials, this constraint was relaxed, and each option had to be sampled at least once. When participants felt confident enough to make a final choice, they pressed the space bar to proceed to a choice interface, where they indicated their selection. As in the description condition, we also elicited present values and certain timing equivalents of the various options (which we again did not consider further).

Data analysis. To test for a description–experience gap in intertemporal choice, we first ran a mixed-effects logistic regression predicting the choice of the sure-timing option, with learning mode (description vs. experience) as a fixed effect and participant as a random effect.[3] We further included as fixed effects the main effect of the probability of the longer delay as well as the interaction between this factor and learning mode, in order to test whether the size and direction of a potential choice gap between description and experience would depend on the type of rare delay (i.e., longer or shorter) in the timing lottery. Furthermore, we included as fixed effects the terms for the main effect of reward size and the corresponding interaction terms with learning mode.[4]

Because a description–experience gap could be produced (at least to some extent) by sampling error in the experience condition (e.g., because participants undersample the rare delay), we tested for such a possible effect in our data. To control for the impact of sampling error in the logistic regression, we conducted an additional analysis involving only (a) description trials with rare delays and (b) experience trials with rare delays and with relative frequencies that matched the objective probabilities. Because of the low number of experience trials for each combination of reward size and expected delay, trials with different reward sizes and expected delays were pooled; in this analysis, we used the choice of the option favored by the overweighting of rare delays as a dependent variable to compare choices in the two learning-mode conditions.

All analyses were conducted using a Bayesian approach, and most of them were implemented using the R package *rstanarm* (Gabry & Goodrich, 2016). In Bayesian analysis, when a 95% credible interval (CI) for a parameter excludes zero, the null hypothesis that the respective effect is absent is interpreted as not credible (see the Supplemental Material for details of the Bayesian analysis).

Finally, to examine whether a possible description–experience gap in choice was accompanied by distinct patterns of probability weighting, we applied a modified version of the RDDU model proposed by Onay and Öncüler (2007), implemented within a hierarchical Bayesian approach (e.g., Lee & Wagenmakers, 2013; Scheibehenne & Pachur, 2015). The modified version assumes a power-utility function, $u(x) = x^{\alpha}$ (Tversky & Kahneman, 1992); an additive combination of utility and discount functions for discounted utility, $du(x,t) = u(x) + d(t) = u(x) - kt$ (Killeen, 2009); and a two-parameter probability-weighting function, $w(p) = \delta p^{\gamma}/(\delta p^{\gamma} + (1-p)^{\gamma}$ (Goldstein & Einhorn, 1987). The parameter γ of the probability-weighting function mainly controls the curvature of the function (or sensitivity to probability differences), and the parameter δ mainly controls the elevation of the function (or optimism to risk and uncertainty).

We adopted an additive rather than the more traditional multiplicative combination rule for the discounted utility for two main reasons. First, the multiplicative combination rule, together with convex discount functions such as exponential or hyperbolic functions, predicts a preference for timing lotteries and is thus inconsistent with the empirical results of timing-risk aversion found in Onay and Öncüler (2007). Consequently, the estimated probability-weighting function might be distorted. The additive combination rule, by contrast, is neutral regarding its prediction on the preference for temporal uncertainty, making a distorted estimation of the probability-weighting function less likely. The fact that the additive combination rule does not necessarily predict a preference for the timing lottery means that this function provides a more appropriate tool for testing whether the description–experience gap is accompanied by differences in probability weighting. Second, the additive combination rule outperformed the multiplicative rule in a hierarchical, Bayesian model-comparison analysis. It therefore appears appropriate to adopt the additive combination rule for the current analysis. Hereafter, we refer to the resulting model as the *additive RDDU model*. See the Supplemental Material for details of the model comparison and parameter estimation.

According to the additive RDDU model, the valuation of a sure-timing option providing a reward of amount x at a delay of t is given by

$$V(x,t) = u(x) + d(t), \tag{1}$$

and the valuation of a timing lottery providing a reward of amount x at either a shorter delay of s with a probability (or experienced relative frequency) of p or a longer delay of l with a probability (or experienced relative frequency) of q (= $1 - p$) is given by

$$V(x,s,p;x,l,1-p) = u(x) + \begin{bmatrix} w(p) \times d(s) + \\ (1-w(p)) \times d(l) \end{bmatrix}. \tag{2}$$

Finally, a logistic choice function (with a sensitivity parameter ϕ) was used to map the difference in valuation onto predicted choice probabilities. The probability of choosing option A over option B is given by

$$P(A|\{A,B\}) = \frac{1}{1+\exp\left[\varphi\left(V(B)-V(A)\right)\right]}.$$

In total, the additive RDDU model has five parameters: α for reward sensitivity, k for discount rate, γ for the shape of the probability-weighting function, δ for the elevation of the probability-weighting function, and φ for choice sensitivity. However, because each pair of options involved in our studies had the same payoff, the utility function would be canceled out in the expression of difference in valuation and thus the predicted choice probabilities. Therefore, only four parameters of the additive RDDU model needed to be estimated: k, δ, γ, and φ. A hierarchical Bayesian analysis was conducted to estimate the model parameters at both the individual and group levels. A graphical illustration of the model is shown in Figure 2 (see the Supplemental Material for details). See https://osf.io/k5b7a/ for raw data and analysis code for this study.

Results

Figure 3 shows the proportion of choices of the sure-timing option, separately for the description and experience conditions and types of rare delay. As can be seen, there was a pronounced description–experience gap in choice problems involving a rare delay. In the description

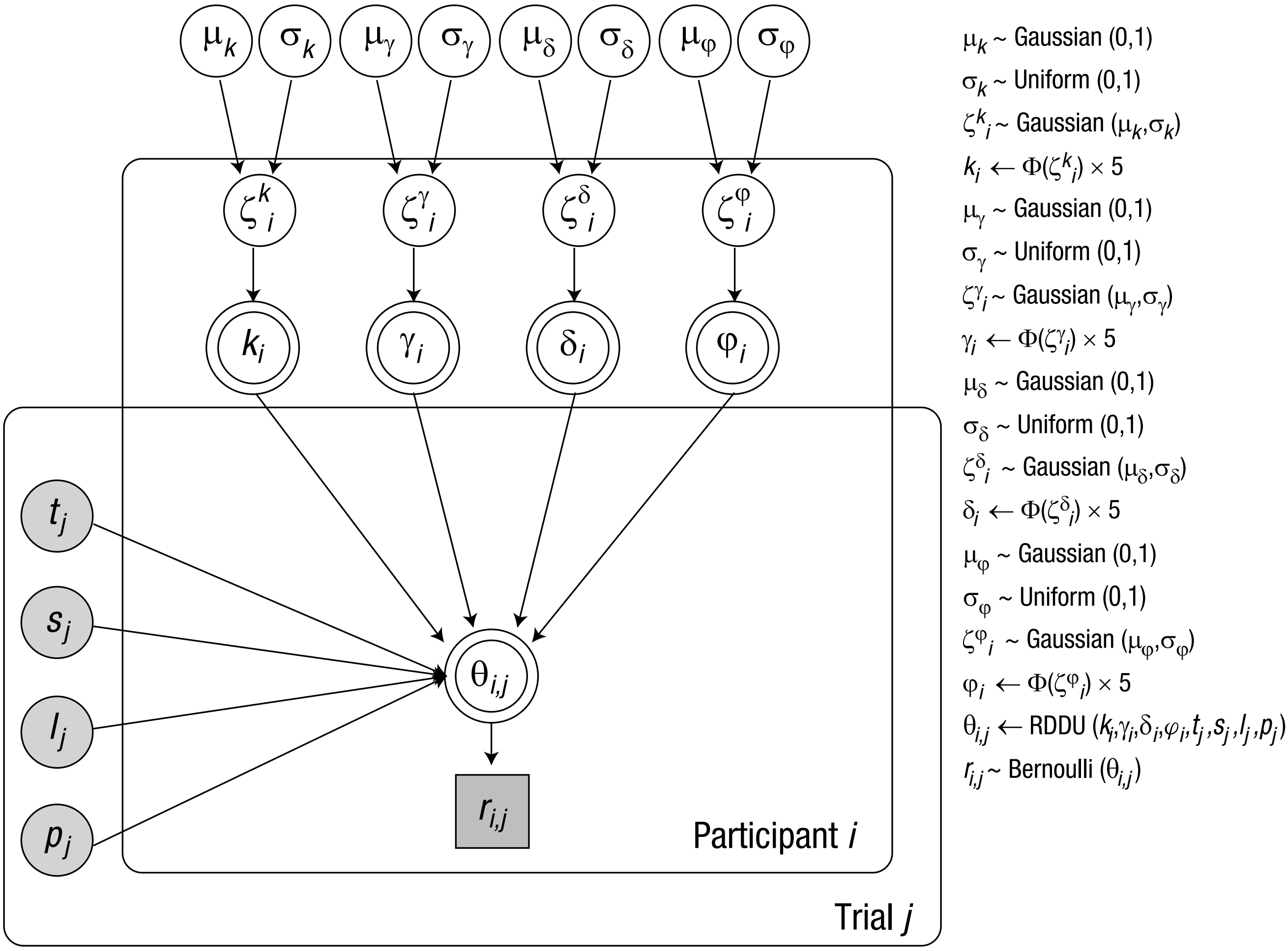

Fig. 2. Graphical illustration of the Bayesian implementation of the additive rank-dependent discounted-utility (RDDU) model. In the model, choice of the timing lottery ($r_{i,j}$) is a Bernoulli random variable governed by the probability of choosing the timing lottery $\theta_{i,j}$. This probability is partly determined by the delay of the sure-timing option (t_j), the shorter delay of the timing lottery (s_j), the longer delay of the timing lottery (l_j), and the probability of the shorter delay (p_j) in the timing lottery (objective in the description condition and relative frequency in the experience condition). The choice probability $\theta_{i,j}$ is also determined by the discount function, the weighting function, and the choice rule. Each of these is governed by their respective parameters, which are ultimately drawn from group-level distributions. Φ is the standard normal cumulative distribution function.

condition, participants chose the sure-timing option more frequently when the longer delay was rare than when the shorter delay was rare. In the experience condition, this pattern was reversed: The sure-timing option was chosen more frequently when the shorter delay, rather than the longer delay, was rare (Table 1). A mixed-effects logistic regression corroborated an interaction between learning mode and the probability of the longer delay (and thus type of rare delay) and the simple effects of learning mode given particular types of rare delay. The regression analysis also revealed an interaction between reward size and learning mode (Table 1): Participants in the description condition chose the sure-timing option more frequently when the reward size was larger than when it was smaller. In the experience condition, reward size had no credible impact on the choices.

Participants in the experience condition sampled, on average (per option, across choice problems), 6.27 times from the delay distributions before making a choice (*Mdn* = 6, *SD* = 3.69). It is no surprise, then, that sampling errors emerged and that the experienced relative frequencies of the rare delays tended to be lower than the objective probabilities. To illustrate, for an objective probability of .1 (regarding the longer delay of 11 months), the median experienced relative frequency was .061, and for an objective probability of .2 (regarding the shorter delay of 1 month), the median experienced relative frequency was .167. Overall, the rare delay was more likely to be underexperienced than overexperienced (95% CI of the probability of underexperiencing = [0.55, 0.66]). A mixed-effects Poisson regression predicting sample size with option type (sure-timing option vs. timing lottery), reward size, and

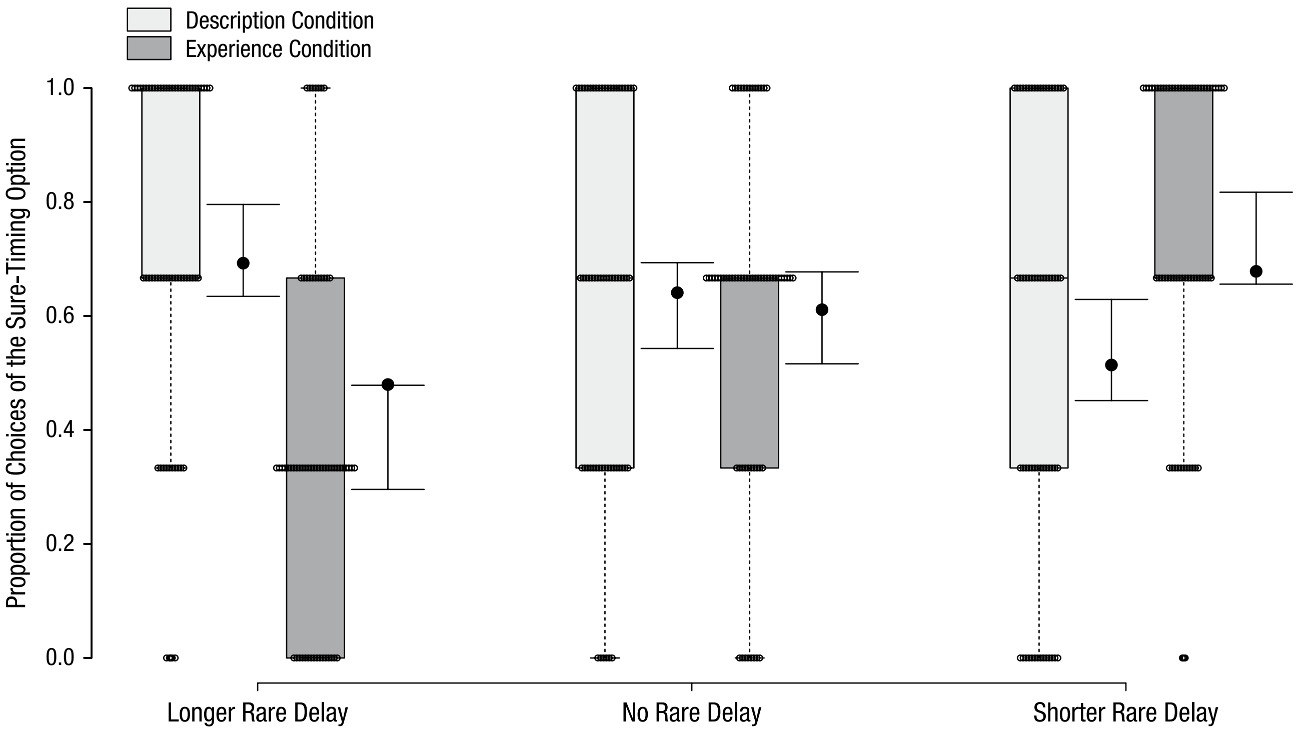

Fig. 3. Box-and-whisker plots showing the distribution (across participants) of the proportion of choices of the sure-timing option in Study 1, separately for the two learning-mode conditions at each type of rare delay. Open circles indicate individual data points. For each plot, the error bar (to the right) shows the 95% credible interval derived from the posterior prediction of the mixed-effects logistic regression. The solid dot to the right of each plot shows the posterior predictive check (i.e., mean choice probability of the sure-timing option derived from the posterior distribution) of the additive rank-dependent discounted-utility (RDDU) model for that distribution.

the probability of the longer delay as fixed effects also revealed credible effects of option type and reward size. Specifically, participants sampled more frequently from the timing lotteries than from the sure-timing options (95% CI of the slope for the dummy variable regarding option type = [–0.097, –0.017]), and they also sampled more frequently from options with larger rewards than from options with smaller rewards (95% CI of the slope for reward size = [0.000020, 0.000083]). See the Supplemental Material for more details.

What is the role of sampling error in the observed description–experience gap? Figure 4 shows the proportion of choices of the option favored by an overweighting of the rare delay, both with and without controlling for sampling error. The figure compares trials in the description condition that had rare delays with trials in the experience condition that had rare delays whose experienced frequencies matched the objective probabilities (thus controlling for sampling error), as well as to all experience trials with rare

Table 1. Results of the Mixed-Effects Logistic Regression in Study 1: Interaction and Simple Effects of Learning Mode, Reward Size, and Expected Delay on Choices of the Sure-Timing Option

Effect	Posterior mean	95% credible interval
Probability of the Longer Delay × Learning Mode	3.69	[2.75, 4.65]
Reward Size × Learning Mode	–0.00050	[–0.00093, –0.000080]
Probability of the longer delay in description condition	–1.37	[–2.05, –0.69]
Probability of the longer delay in experience condition	2.33	[1.68, 3.00]
Reward size in description condition	0.00063	[0.00032, 0.00096]
Reward size in experience condition	0.00011	[–0.00017, 0.00039]

Note: When a 95% credible interval for a parameter excludes zero, the null hypothesis that it is zero is interpreted as being not credible.

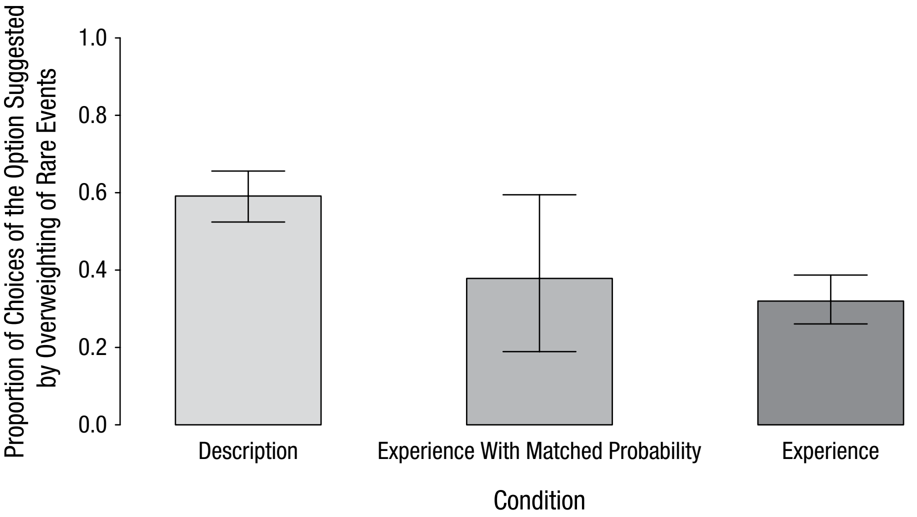

Fig. 4. Observed proportion of choices of the option suggested by overweighting of rare events in Study 1. Results are shown separately for the description condition, the experience condition when sampling error was taken into account (middle bar), and the experience condition when sampling error was not taken into account (right-hand bar). The error bar for the latter condition shows the 95% credible interval derived from the mixed-effects logistic regression. The error bars for the other two conditions show 95% credible intervals derived from the pooled logistic regression with learning mode as the only predictor.

delays. As can be seen, there was a description–experience gap even when sampling error was taken into account. Specifically, pooled across the three reward sizes and two types of rare delay, the data showed that participants chose the option suggested by an overweighting of rare events more often in the description than in the experience condition (95% CI for the term of learning mode in the corresponding logistic regression = [−1.86, −0.10]).

Finally, to examine the role of probability weighting for the observed description–experience gap, we fitted the additive RDDU model to our data. As input for estimating the probability-weighting function, we used the stated probabilities in the description condition and the experienced relative frequencies in the experience condition, thus controlling for sampling error. The additive RDDU model was implemented using a Bayesian hierarchical approach.

Figure 5 shows the estimated individual- and group-level probability-weighting functions, using the means of the posterior distributions of the parameters obtained from the Bayesian modeling analysis. In the description

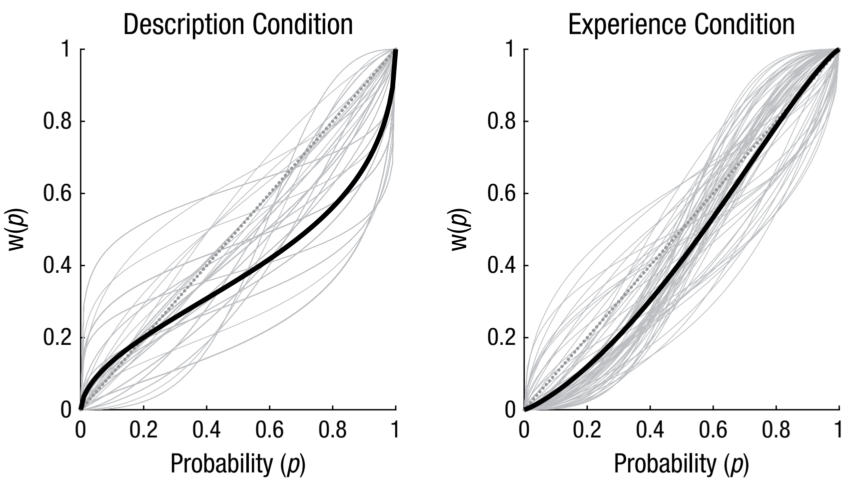

Fig. 5. Probability-weighting, or w(p), functions based on the individual- and group-level parameters (solid lines in gray and black, respectively) in Study 1. The dashed line marks the diagonal.

Table 2. Posterior Group-Level Means for the Rank-Dependent Discounted-Utility (RDDU) Parameters and the Differences Between the Learning-Mode Conditions in Study 1

RDDU parameter	Description condition	Experience condition	Difference (experience − description)
k (delay discounting)	1.31 [0.09, 4.55]	1.51 [0.27, 4.49]	0.20 [−3.58, 3.71]
δ (elevation)	0.57 [0.30, 0.84]	0.71 [0.51, 0.93]	0.14 [−0.20, 0.48]
γ (curvature)	0.59 [0.24, 0.93]	1.19 [0.84, 1.63]	0.60 [0.10, 1.15]
φ (choice sensitivity)	1.67 [0.11, 4.71]	1.62 [0.27, 4.48]	−0.05 [−3.81, 3.56]

Note: Values in brackets are 95% credible intervals.

condition, the obtained parameters indicate an inverse-S-shaped probability-weighting function, consistent with an overweighting of rare events (as is commonly observed in risky choice). In the experience condition, in contrast, the obtained parameters indicate an S-shaped probability-weighting function, consistent with an underweighting of rare events. Overall, rare delays were assigned a lower decision weight in the experience than in the description condition. Table 2 reports the group-level means and 95% CIs for each parameter of the additive RDDU model from the hierarchical Bayesian analysis, separately for the two learning-mode conditions. The group-level curvature parameter (i.e., γ) of the probability-weighting function for the description condition was credibly lower than 1; in the experience condition, the same parameter was slightly higher than 1, though not credibly so. The difference between the two conditions was credibly larger than zero, suggesting distinct patterns of probability weighting. Finally, the dots in Figure 4 show the posterior prediction of the additive RDDU model; as can be seen, the model captures the description–experience gap rather well, verifying the validity of the model.

Summary

Study 1 established a fourfold pattern of preferences in intertemporal choices with temporal uncertainty. When the shorter delay in the lottery was rare, people preferred a timing lottery more often in the description condition (risk) than in the experience condition (uncertainty). When the longer delay in the lottery was rare, this pattern was reversed: People preferred a timing lottery more often in the experience than in the description condition. Neither the influence of the type of delay nor that of learning mode is predicted by the standard DEU model. This new description–experience gap in intertemporal choice can be partly attributed to an underexperiencing of rare delays. Computational modeling that controlled for sampling error also established that the gap in people's choices was accompanied by distinct patterns of probability weighting. In Study 2, we sought to replicate this description–experience gap over a larger space of options and to examine it in a new condition that controlled for sampling error experimentally (rather than statistically, as in Study 1).

Study 2

In Study 1, we adopted Onay and Öncüler's (2007) intertemporal-choice stimuli to contrast timing risk and timing uncertainty. One potential drawback of the stimuli, however, was that they involved only two extreme probabilities, one small and one large, for the shorter delay in the timing lotteries. This is not ideal for reliably estimating the shape of the probability-weighting function. The modeling analysis was further constrained by the fact that shorter and longer delays in the timing lotteries were fixed across options and that the certain delays in the sure-timing options always equaled the expected delays of the timing lotteries. In Study 2, we addressed these shortcomings by implementing a larger number of probability levels and possible delays for the timing lotteries, each of which was paired with several sure-timing options with different delays. We also introduced a fixed-sampling condition to investigate the description–experience gap while experimentally removing any mismatch between objective probabilities and experienced relative frequencies. This study was preregistered as a replication and extension of Study 1 (see https://osf.io/k5b7a/).

Method

Participants. We recruited 180 adults (102 women; mean age = 25.82 years, *SD* = 4.38 years) from a subject pool of the Max Planck Institute for Human Development in Berlin, Germany. A third of the participants (n = 60) were randomly assigned to the description condition (34 women; mean age = 26.05 years, *SD* = 4.68 years); another third to a *sampling condition*, which corresponded to the experience condition in Study 1 (36 women; mean age =

Table 3. Design Matrix of Study 2 Regarding Possible Delays and Associated Probabilities

Shorter delay in timing lottery	Probability of shorter delay	Longer delay in timing lottery	Probability of longer delay	Expected delay of timing lottery	Certain delay of sure-timing option
1	.9	30	.1	3.9	2
1	.9	30	.1	3.9	3
1	.9	30	.1	3.9	4
1	.9	30	.1	3.9	5
1	.9	30	.1	3.9	6
3	.8	28	.2	8	4
3	.8	28	.2	8	6
3	.8	28	.2	8	8
3	.8	28	.2	8	10
3	.8	28	.2	8	12
4	.2	27	.8	22.4	18
4	.2	27	.8	22.4	20
4	.2	27	.8	22.4	22
4	.2	27	.8	22.4	24
4	.2	27	.8	22.4	26
2	.1	29	.9	26.3	24
2	.1	29	.9	26.3	25
2	.1	29	.9	26.3	26
2	.1	29	.9	26.3	27
2	.1	29	.9	26.3	28

Note: Values in the delay columns are given in months.

25.47 years, *SD* = 3.88 years); and the remaining third to a *fixed-sampling condition* (32 women; mean age = 25.93 years, *SD* = 4.58 years). We aimed for the same sample size (i.e., 60) for each condition in Study 2 as in Study 1, because the latter demonstrated a description–experience gap with this sample size. Each participant signed a consent form approved by the institute's ethical review board before starting the study and was paid €15 for participation.

Materials and procedure. Three conditions were tested in this study: description, sampling, and fixed sampling. The description condition was implemented in the same way as in Study 1, but with a more informative design matrix regarding possible delays and associated probabilities for estimating the probability-weighting function (Table 3). First, we adopted two small probabilities (.1 and .2) and two large probabilities (.8 and .9) for the shorter delay in the timing lottery so that more points on the probability-weighting function were involved in the estimation. Each of the probabilities was coupled with a slightly different pair of shorter and longer delays to generate a delay distribution for the timing lottery. Second, for each timing lottery, we created five sure-timing options with distinct certain delays that varied around the expected delay of the timing lottery in order to render possible a more precise estimation of the probability-weighting function. Finally, for each pair of delay distributions (one sure and one risky), we used three possible hypothetical rewards (€60, €300, and €1,500) to generate three choice problems. In total, 60 choice problems (4 Probability × 5 Certain Delay × 3 Reward Amount) were generated for each participant. All other settings were the same as in Study 1, except that neither present values nor certain timing equivalents were elicited in Study 2.

In the sampling condition, the choice options were the same as in the description condition in terms of reward sizes, delay lengths, and objective probabilities, and the learning mode was the same as in the experience condition of Study 1. Neither current values nor certain timing equivalents were elicited. Finally, the fixed-sampling condition was identical to the sampling condition except that in both the practice and test trials of Study 2, participants were required to draw 10 samples from each option; the experienced relative frequencies of possible delays were fixed to be the same as the objective probabilities. Consequently, any gap between this and the description condition must be due to causes other than sampling error.

Data analysis. In order to test for a description–experience gap in intertemporal choice, we first conducted a

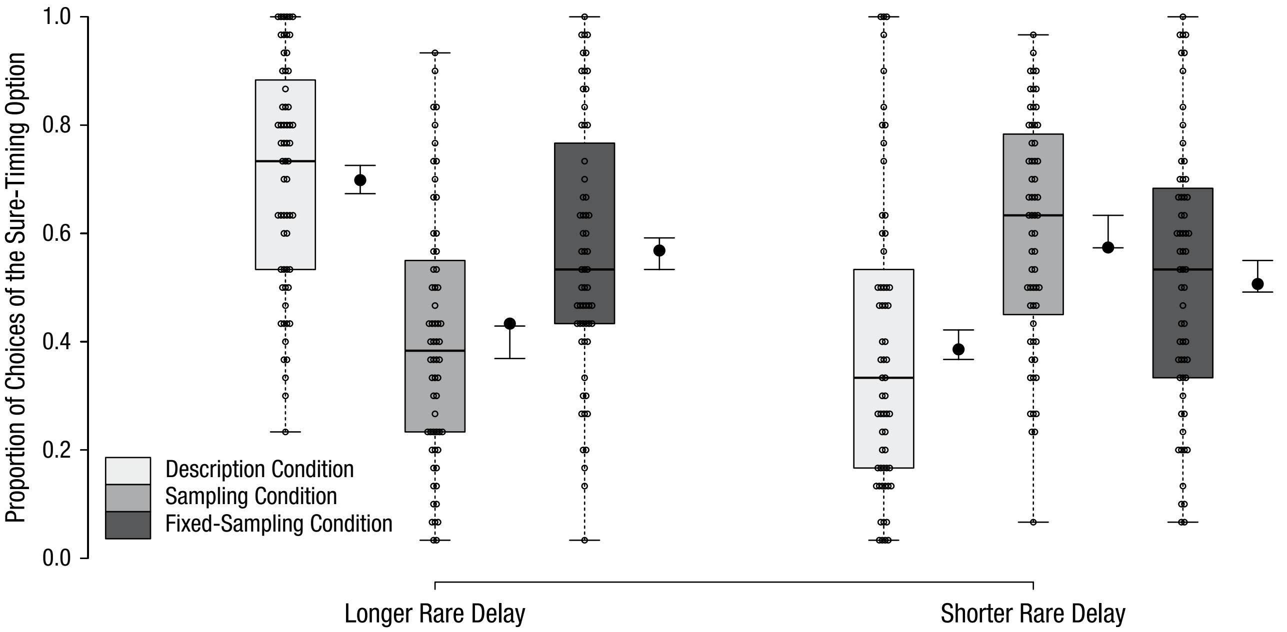

Fig. 6. Box-and-whisker plots showing the distribution (across participants) of the proportion of choices of the sure-timing option in Study 2, separately for the three learning-mode conditions at each type of rare delay. Open circles indicate individual data points. For each plot, the error bar (to the right) shows the 95% credible interval derived from the posterior prediction of the mixed-effects logistic regression. The solid dot to the right of each plot shows the posterior predictive check (i.e., mean choice probability of the sure-timing option derived from the posterior distribution) of the additive rank-dependent discounted-utility (RDDU) model for that distribution.

mixed-effects logistic regression predicting the choice of the sure-timing option with learning mode as a fixed effect (description vs. sampling vs. fixed sampling). Because the direction of the gap might depend on the type of rare delay (which was a function of the probability of the longer delay) in the timing lottery, we also included an interaction term between learning mode and probability of the longer delay (as well as a term for the main effect of the probability of the longer delay). Furthermore, to examine the potential impact of reward size on choice and its interaction with learning mode, we included terms for the main effect of reward size and corresponding interaction terms as fixed effects.[5] Finally, we also included a term for the difference in expected delay between the timing lottery and the sure-timing option as well as its interaction term with learning mode. The additional terms were required because, in contrast to Study 1, this difference now varied across problems.

As in Study 1, we tested for sampling error in the sampling condition. In addition, we ran a comparison between the description condition and the sampling condition while controlling for sampling error; we did this by including those experience trials with rare delays in which experienced relative frequencies matched the objective probabilities. As in Study 1, we compared choice proportions of the option that would be more attractive under overweighting of the rare delay between the description and experience conditions. Finally, we applied the additive RDDU model to examine differences in probability weighting among the three learning modes. See https://osf.io/k5b7a/ for raw data and analysis code for this study.

Results

Figure 6 shows the proportion of choices of the sure-timing option, separately for the different learning modes and types of rare delay. When a rare delay was relatively long (i.e., the probability of the longer delay was either .1 or .2), participants chose the sure-timing option over the timing lottery most frequently in the description condition, less frequently in the fixed-sampling condition, and least frequently in the sampling condition. When a rare delay was relatively short (i.e., the probability of the longer delay was either .8 or .9), participants instead chose the sure-timing option least frequently in the description condition, more frequently in the fixed-sampling condition, and most frequently in the sampling condition. The same patterns occurred when problems with different probabilities of the longer delay were analyzed separately. This interaction between learning mode and type of rare delay (as represented by the probability of the longer delay) and the simple effects of learning mode given different

Table 4. Results of the Mixed-Effects Logistic Regression in Study 2: Interactions and Simple Effects of Learning Mode, Reward Size, Type of Delay, and Difference in Expected Delay on Choices of the Sure-Timing Option

Effect	Posterior mean	95% credible interval
Type of Delay × Sampling Condition	3.81	[3.50, 4.13]
Type of Delay × Fixed-Sampling Condition	2.11	[1.79, 2.43]
Reward Size × Sampling Condition	−0.00029	[−0.00046, −0.00012]
Reward Size × Fixed-Sampling Condition	−0.00033	[−0.00050, −0.00016]
Difference in Expected Delay × Sampling Condition	−0.20	[−0.25, −0.15]
Difference in Expected Delay × Fixed-Sampling Condition	−0.041	[−0.094, 0.012]
Description condition vs. sampling condition at longer rare delay	−1.57	[−2.10, −1.02]
Description condition vs. fixed-sampling condition at longer rare delay	−0.61	[−1.15, −0.06]
Description condition vs. sampling condition at shorter rare delay	1.41	[0.81, 2.01]
Description condition vs. fixed-sampling condition at shorter rare delay	1.03	[0.42, 1.63]
Reward size in description condition	0.00036	[0.00023, 0.00048]
Reward size in sampling condition	0.00006	[−0.00005, 0.00018]
Reward size in fixed-sampling condition	0.00002	[−0.00009, 0.00014]
Difference in expected delay in description condition	0.40	[0.36, 0.44]
Difference in expected delay in sampling condition	0.18	[0.15, 0.21]
Difference in expected delay in fixed-sampling condition	0.35	[0.31, 0.39]

Note: When the 95% credible interval for a regression coefficient excludes zero, the null hypothesis that it is zero (i.e., that there is no effect) is interpreted as being not credible.

types of rare delay were corroborated by the results of the mixed-effects logistic regression (Table 4). Note that the description–experience gap emerged irrespective of whether sampling was fixed so that the experienced relative frequencies perfectly matched the objective probabilities. The regression analysis also revealed a similar interaction between reward size and learning mode on choices (Table 4): Participants in the description condition chose the sure-timing option more frequently when the reward size was larger, whereas reward size did not have a credible impact on choices in either the sampling or fixed-sampling conditions.

Participants in the sampling condition sampled, on average, 6.25 times from the uncertain option (Mdn = 5, SD = 4.72), almost exactly replicating the number of samples observed in Study 1. Again, there was sampling error, and the experienced relative frequencies of the rare delays tended to be lower than their objective probabilities: For an objective probability of .1, the median experienced relative frequency was 0, whereas for an objective probability of .2, the median experienced relative frequency was .167. For both probabilities, the rare delay was credibly more likely to be underexperienced than overexperienced (95% CIs for the probability of underexperiencing were [0.59, 0.64] and [0.54, 0.59], respectively). Finally, a mixed-effects Poisson regression on sample size with option type (sure-timing option vs. timing lottery) and reward size as fixed effects also revealed credible effects of the two predictors. Participants sampled more frequently from the timing lotteries than from the sure-timing options (95% CI of the slope for the dummy variable regarding option type = [−0.154, −0.116]), and they also sampled more frequently from options with larger rewards than from options with smaller rewards (95% CI of the slope for reward size = [0.0000023, 0.000032]). See the Supplemental Material for more details.

Finally, in order to examine the role of probability weighting in the description–experience gap, we fitted the additive RDDU model to the data separately for the three conditions. Figure 7 shows the individual- and group-level probability-weighting patterns for the timing information, using the means of the posterior distributions of the parameters obtained from the Bayesian modeling analysis. The probability-weighting patterns for the description and sampling conditions were similar to those in Study 1 (Fig. 5): an inverse-S-shaped function in the former and an S-shaped function in the latter. The pattern for the fixed-sampling condition was in between, assuming an approximately linear weighting function. This is consistent with the results on the observed choice proportions, where the fixed-sampling condition fell between the other two conditions. As in Study 1, rare delays were generally assigned a lower decision weight in the experience (sampling and fixed-sampling) than in the description conditions. Table 5 reports the group-level means and 95% CIs for each parameter from the hierarchical Bayesian analysis,

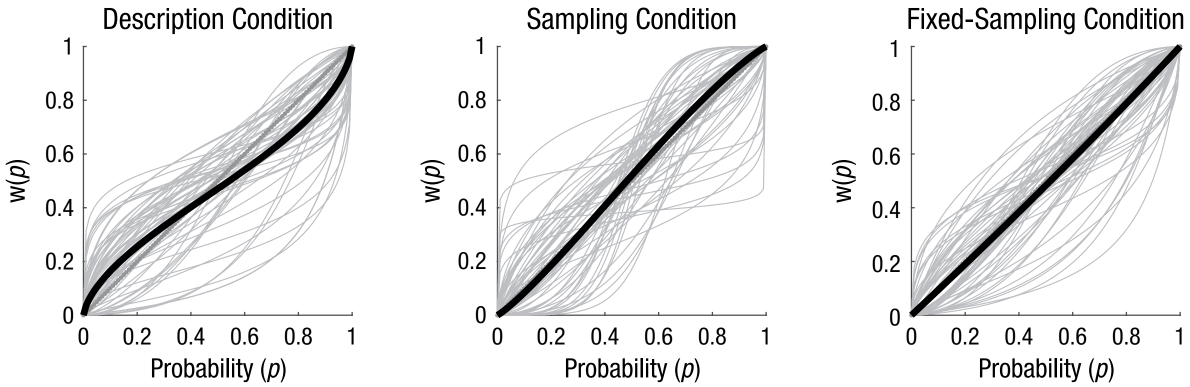

Fig. 7. Probability-weighting, or w(p), functions based on individual- and group-level parameters (solid lines in gray and black, respectively) in Study 2. The dashed line marks the diagonal.

separately for the three conditions. In the description condition, the curvature parameter (i.e., γ) of the probability-weighting function was credibly lower than 1. In the sampling condition, by contrast, this parameter was likely to be higher than 1, though not credibly so; the credible interval for the fixed-sampling condition was approximately symmetric around 1. Table 6 shows the means and credible intervals for the differences between the three conditions on the group-level parameters. For the curvature parameter, the differences between the description condition on the one hand and the two experience conditions on the other hand were credibly lower than zero, suggesting distinct patterns of probability weighting. By contrast, the difference between the two experience conditions was not credible. Finally, the solid dots in Figure 6 show the means of the posterior predictive distribution of the additive RDDU model. Again, the model captured the data rather well, verifying the validity of the model.

Summary

Replicating Study 1, Study 2 found a description–experience gap in intertemporal choice. When sampling error was not controlled for in the sampling condition, the underexperiencing of rare delays and the diverging patterns of probability weighting accompanied the gap. When sampling error was fully controlled for in the fixed-sampling condition, the gap persisted and was captured by different patterns in probability weighting.

General Discussion

It is difficult to think of intertemporal choices devoid of temporal uncertainty (McGuire & Kable, 2013). Will an online purchase be delivered on time? Will lunch be served quickly enough to get back to the meeting? Will the 4-year college degree take 4, 5, or 6 years? Nevertheless, many lab studies on intertemporal choice have removed not only timing risk but also, crucially, timing uncertainty from the stimuli.

In two studies, we compared intertemporal choice involving timing risk (known probabilities of delay) with intertemporal choice involving timing uncertainty (imprecise knowledge of the probabilities), leading us to draw three key conclusions. First, in contrast to the standard economic assumption embodied in the DEU model, people do not evaluate timing lotteries by taking the stated probabilities or experienced relative frequencies at face value. Instead, these values enter choices in a nonlinear fashion, consistent with Onay and Öncüler's (2007) findings about timing risk. Second, there is a

Table 5. Group-Level Means for the Posterior Distributions of the Parameters of the Additive Rank-Dependent Discounted-Utility (RDDU) Model

RDDU parameter	Description condition	Sampling condition	Fixed-sampling condition
k (delay discounting)	0.61 [0.11, 2.97]	2.35 [0.36, 4.88]	0.80 [0.11, 3.80]
δ (elevation)	0.89 [0.73, 1.06]	1.09 [0.87, 1.33]	0.93 [0.75, 1.14]
γ (curvature)	0.69 [0.58, 0.80]	1.12 [0.83, 1.45]	1.00 [0.88, 1.13]
φ (choice sensitivity)	1.80 [0.18, 4.72]	0.65 [0.16, 2.41]	1.50 [0.13, 4.34]

Note: Values in brackets are 95% credible intervals.

Table 6. Group-Level Means of the Differences Between the Learning-Mode Conditions in the Posterior Distributions of the Parameters of the Additive Rank-Dependent Discounted-Utility (RDDU) Model

RDDU parameter	Sampling – description	Fixed sampling – description	Fixed sampling – sampling
k (delay discounting)	1.74 [−0.60, 4.46]	0.19 [−2.56, 3.62]	−1.56 [−4.33, 2.86]
δ (elevation)	0.20 [−0.08, 0.49]	0.04 [−0.02, 0.30]	−0.15 [−0.46, 0.14]
γ (curvature)	0.44 [0.13, 0.78]	0.31 [0.14, 0.48]	−0.13 [−0.48, 0.19]
φ (choice sensitivity)	−1.15 [−4.21, 1.19]	−0.31 [−4.08, 3.35]	0.85 [−1.73, 3.82]

Note: Values in brackets are 95% credible intervals.

description-experience gap: Intertemporal choices under temporal uncertainty systematically depend on how people learn about the possible delays and their likelihoods. A timing lottery with a shorter rare delay was chosen more often against a sure-timing option when timing information was communicated via description rather than via experiential sampling (where timing information was more uncertain). This pattern was reversed when the longer delay in the timing lottery was rare, creating a fourfold pattern of preferences for intertemporal choices with temporal uncertainty. This brings us to our third key finding, which concerns the processes underlying this description–experience gap. Our analyses show that while the gap is in part due to sampling error in the experience conditions (Figs. 4 and 6), a gap remained even after we controlled for sampling error. Computational modeling indicated that this aspect of the gap is reflected in differences in the nonlinear weighing of stated probabilities versus learned relative frequencies. Specifically, rare delays receive greater subjective weight in description than in experience (Tables 2 and 6; see also Figs. 5 and 7), suggesting that people respond to timing risk and timing uncertainty in different ways.[6] In the sampling condition, the weighting function was even S-shaped. These findings raise questions about the adequacy of the conventional DEU approach to modeling intertemporal choice under temporal uncertainty; they also challenge the generality of the inverse-S-shaped probability-weighting pattern assumed in the RDDU model (Onay & Öncüler, 2007).

What mechanisms might lead to divergent probability weighting in the description and experience conditions? First, it is possible that this finding reflects differential attention policies (see also Pachur, Schulte-Mecklenbeck, Murphy, & Hertwig, 2018). In the description condition, comparable amounts of attention might be paid to different delays irrespective of their probability, leading to the inverse-S-shaped weighting pattern observed in this condition; in the experience condition, by contrast, people may instead treat the delays more categorically, classifying rare delays as impossible and more frequent delays as possible, leading to an S-shaped weighting pattern. Second, in the experience condition, memory limitations might reduce the effective influence of rarely experienced delays relative to what their experienced relative frequencies imply. Still another possibility is that decision makers rely on different heuristics in description and experience, which in risky choice has been shown to lead to different kinds of nonlinear probability weighting (Pachur & Hertwig, 2019; Pachur, Suter, & Hertwig, 2017).

Finally, we want to emphasize that in intertemporal choice, experiential learning can occur in various ways (e.g., Ashby & Gonzalez, 2017). The sampling approach implemented here captures only one form of experiential learning—for instance, asking friends how long they waited in line to get into a popular club or searching one's memory for waiting times at the general practitioner when pondering whether to find another doctor. In other intertemporal-choice situations, it is the delays themselves that are experienced before a decision is made (e.g., Jimura, Myerson, Hilgard, Braver, & Green, 2009). For example, one may decide not to wait in line at a favorite restaurant because the previous experience of a long wait was very unpleasant. In the former case, experiential learning involves the experience of stated delays, whereas in the latter case, the person is dealing with experienced delays. In real-world situations, both types of experiential learning might occur simultaneously (for a discussion, see Dai, Pachur, Pleskac, & Hertwig, 2019). In future work, researchers could implement and investigate different types of experience to obtain a better understanding of realistic intertemporal choices under temporal uncertainty.

Action Editor

D. Stephen Lindsay served as action editor for this article.

Author Contributions

All of the authors developed the study concept. J. Dai, T. Pachur, and T. J. Pleskac designed the study. J. Dai analyzed the data, and T. Pachur and T. J. Pleskac assisted in the data analysis. All of the authors contributed to writing the manuscript.

ORCID iD

Junyi Dai https://orcid.org/0000-0001-7052-2865

Acknowledgments

We thank Susannah Goss and Deborah Ain for editing the manuscript, Sarah Ottenstetter for editing the figures, and Dania Esch for editing the references.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was supported by a National Nature Science Foundation of China grant to Junyi Dai (Grant No. 31872780) and a Fundamental Research Funds for the Central Universities of China grant to Junyi Dai (Grant No. 2018QNA3014).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619858969

Open Practices

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/k5b7a/. The design and analysis plans for the studies were preregistered at https://osf.io/k5b7a/. There was one change to the preregistered analysis plan (see the Open Practices Disclosure for details). The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619858969. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

Notes

1. The delivery statistics given here were obtained from www.trackingmore.com.
2. We use the terms *temporal uncertainty* (as in McGuire & Kable, 2012, 2013) to describe the general problem that people often do not know how long it will take for a future outcome to materialize, *timing risk* (as in Onay & Öncüler, 2007) to refer to situations in which possible delays and their probabilities are known, and *timing uncertainty* to refer to situations in which possible delays and their probabilities are only vaguely known or unknown.
3. By-participants random slopes could, in principle, also be included in the mixed-effects regression. In this research, however, we adopted a different approach (i.e., hierarchical Bayesian analysis with the RDDU model reported later in this section) to examine potential individual differences in the effects of relevant predictors on choice.
4. We also tested the full factorial model with all interaction terms. It emerged that neither the three-way interaction among learning mode, reward size, and the probability of the longer delay nor the two-way interaction between learning mode and reward size was credible. We also performed the same analyses with the probability of the longer delay and reward size as ordinal instead of interval variables. The results were virtually the same. In the Results sections, we therefore report the results of the simpler model.
5. As in Study 1, more complex models with higher-order interactions or ordinal predictors were tested; again, we obtained the same results.
6. Mapping results from risky choice on intertemporal choice would suggest that the direction of the gap and the underlying difference in probability weighting may change when the safe option is also replaced by a timing lottery (Glöckner, Hilbig, Henninger, & Fiedler, 2016; Kellen, Pachur, & Hertwig, 2016).

References

Andreoni, J., & Sprenger, C. (2012a). Estimating time preferences from convex budgets. *The American Economic Review*, *102*, 3333–3356. doi:10.3386/w16347

Andreoni, J., & Sprenger, C. (2012b). Risk preferences are not time preferences. *The American Economic Review*, *102*, 3357–3376. doi:10.1257/aer.102.7.3357

Ashby, N. J. S., & Gonzalez, C. (2017). The influence of time estimation and time-saving preferences on learning to make temporally dependent decisions from experience. *Journal of Behavioral Decision Making*, *30*, 807–818. doi:10.1002/bdm.2006

Chesson, H. W., & Viscusi, W. K. (2003). Commonalities in time and ambiguity aversion for long-term risks. *Theory and Decision*, *54*, 57–71. doi:10.1023/A:1025095318208

Dai, J., Pachur, T., Pleskac, T. J., & Hertwig, R. (2019). Tomorrow never knows: Why and how uncertainty matters in intertemporal choice. In R. Hertwig, T. J. Pleskac, T. Pachur, & the Center for Adaptive Rationality (Eds.), *Taming uncertainty* (pp. 175–190). Cambridge, MA: MIT Press.

Gabry, J., & Goodrich, B. (2016). rstanarm: Bayesian applied regression modeling via Stan (R Package Version 2.13.1) [Computer software]. Retrieved from https://CRAN.R-project.org/package=rstanarm

Glöckner, A., Hilbig, B. E., Henninger, F., & Fiedler, S. (2016). The reversed description-experience gap: Disentangling sources of presentation format effects in risky choice. *Journal of Experimental Psychology: General*, *145*, 486–508. doi:10.1037/a0040103

Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomenon. *Psychological Review*, *94*, 236–254. doi:10.1037/0033-295X.94.2.236

Green, L., Myerson, J., & Ostaszewski, P. (1999). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 418–427. doi:10.1037/0278-7393.25.2.418

Halevy, Y. (2008). Strotz meets Allais: Diminishing impatience and the certainty effect. *The American Economic Review*, *98*, 1145–1162. doi:10.1257/aer.98.3.1145

Hertwig, R. (2015). Decisions from experience. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making* (Vol. 1, pp. 240–267). Chichester, England: John Wiley & Sons.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*, 517–523. doi:10.1016/j.tics.2009.09.004

Hertwig, R., & Pleskac, T. J. (2018). The construct–behavior gap and the description–experience gap: Comment on Regenwetter and Robinson (2017). *Psychological Review*, *125*, 844–849. doi:10.1037/rev0000121

Jimura, K., Myerson, J., Hilgard, J., Braver, T. S., & Green, L. (2009). Are people really more patient than other animals? Evidence from human discounting of real liquid rewards. *Psychonomic Bulletin & Review*, *16*, 1071–1075. doi:10.3758/PBR.16.6.1071

Kellen, D., Pachur, T., & Hertwig, R. (2016). How (in)variant are subjective representations of described and experienced risk and rewards? *Cognition*, *157*, 126–138. doi:10.1016/j.cognition.2016.08.020

Killeen, P. R. (2009). An additive-utility model of delay discounting. *Psychological Review*, *116*, 602–619.

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge, England: Cambridge University Press.

Luckman, A., Donkin, C., & Newell, B. R. (2018). Can a single model account for both risky choices and inter-temporal choices? Testing the assumptions underlying models of risky intertemporal choice. *Psychonomic Bulletin & Review*, *25*, 785–792. doi:10.3758/s13423-017-1330-8

McGuire, J. T., & Kable, J. W. (2012). Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition*, *124*, 216–226. doi:10.1016/j.cognition.2012.03.008

McGuire, J. T., & Kable, J. W. (2013). Rational temporal predictions can underlie apparent failures to delay gratification. *Psychological Review*, *120*, 395–410. doi:10.1037/a0031910

Miao, B., & Zhong, S. (2015). Comment on "risk preferences are not time preferences": Separating risk and time preference. *The American Economic Review*, *105*, 2272–2286. doi:10.1257/aer.20131183

Onay, S., & Öncüler, A. (2007). Intertemporal choice under timing risk: An experimental approach. *Journal of Risk and Uncertainty*, *34*, 99–121. doi:10.1007/s11166-007-9005-x

Pachur, T., & Hertwig, R. (2019). The weight of uncertain events. In R. Hertwig, T. J. Pleskac, T. Pachur, & the Center for Adaptive Rationality (Eds.), *Taming uncertainty* (pp. 153–174). Cambridge, MA: MIT Press.

Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, *147*, 147–169.

Pachur, T., Suter, R. S., & Hertwig, R. (2017). How the twain can meet: Prospect theory and models of heuristics in risky choice. *Cognitive Psychology*, *93*, 44–73.

Prelec, D., & Loewenstein, G. (1991). Decision making over time and under uncertainty: A common approach. *Management Science*, *37*, 770–786. doi:10.1287/mnsc.37.7.770

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14. doi:10.1002/bdm.681

Regenwetter, M., & Robinson, M. M. (2017). The construct–behavior gap in behavioral decision research: A challenge beyond replicability. *Psychological Review*, *124*, 533–550. doi:10.1037/rev0000067

Samuelson, P. A. (1937). A note on measurement of utility. *The Review of Economic Studies*, *4*, 155–161. doi:10.2307/2967612

Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin & Review*, *22*, 391–407. doi:10.3758/s13423-014-0684-4

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society B: Biological Sciences*, *265*, 2015–2020. doi:10.1098/rspb.1998.0534

Takahashi, T., Ikeda, K., & Hasegawa, T. (2007). A hyperbolic decay of subjective probability of obtaining delayed rewards. *Behavioral and Brain Functions*, *3*, Article 52. doi:10.1186/1744-9081-3-52

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representations of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. doi:10.1007/bf00122574

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Walther, H. (2010). Anomalies in intertemporal choice, time-dependent uncertainty and expected utility – a common approach. *Journal of Economic Psychology*, *31*, 114–130. doi:10.1016/j.joep.2009.11.006

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*, 140–176. doi:10.1037/bul0000115