Supplementary Materials for

Disentangling the Effects of Counterfactual Feedback on Maximization and Risk Preference across Gains and Losses

Ali Shiravand*, Maëlle Gueguen, Sophie Bavard, Dirk U. Wulff, Julien Bastin, Stefano Palminteri

***Corresponding author. Email: ali.shiravand@ens.fr**

This PDF file includes:

Supporting text
Figs. S1 to S6
SI References

Supplementary Text

Utility function estimation. To assess and compare utility curvature between experience-based and description-based choices, we fitted a simple utility function with an exponent parameter to the participants' behavior using a Bayesian maximum a posteriori (MAP) method for fitting (see the following section for modeling details). This utility function helps us quantify the relative weight given by participants to gains and losses. As shown in Equation 1, each participant has two exponents for the utility $U(x)$, one for the gain domain (λ_+) and one for the loss domain (λ_-). The exponent λ determines the curvature of the utility function, reflecting how subjectively sensitive participants are to gains and losses:

$$U(x) = \begin{cases} x^{\lambda_+} & if \quad x \geq 0 \\ -|x|^{\lambda_-} & if \quad x < 0 \end{cases} \qquad [1]$$

In our modeling approach, we do not considered to include a probability weighting function (1), because our experiments do not feature extreme probability values. The schematic utility functions of all participants, based on the median fitted exponent values, are depicted in Figure S1a. These functions visually represent the average attitude of participants toward gains and losses. Although participants generally followed similar utility functions in the gain domain across both experience-based and description-based choices, the pattern differed significantly in the loss domain. This suggests that how participants perceive and react to losses changes depending on whether they are making decisions based on experience or descriptions.

To further investigate this difference, the fitted values per participant are shown in Figure S1b. In description-based decisions, consistent with prospect theory (2, 3), λ_- is significantly higher than λ_+ (Mann–Whitney U test, $U = 48944$, $P < 0.001$, N=400). This means that when decisions were based on descriptions, participants showed a stronger aversion to losses compared to their attraction to gains.

On the other hand, in both partial and complete feedback experience-based choices, λ_- is significantly lower than λ_+ (Mann–Whitney U test, N=200, partial feedback condition: $U = 22370$, $P = 0.040$; complete feedback condition: $U = 33189$, $P = 0 < 0.001$). This indicates that when decisions were based on past outcomes, participants were less sensitive to losses and more responsive to gains—opposite to what was observed in the description-based decisions. Furthermore, we note that the difference between experience- and description-based loss parameters was even more pronounced in the complete feedback experiments.

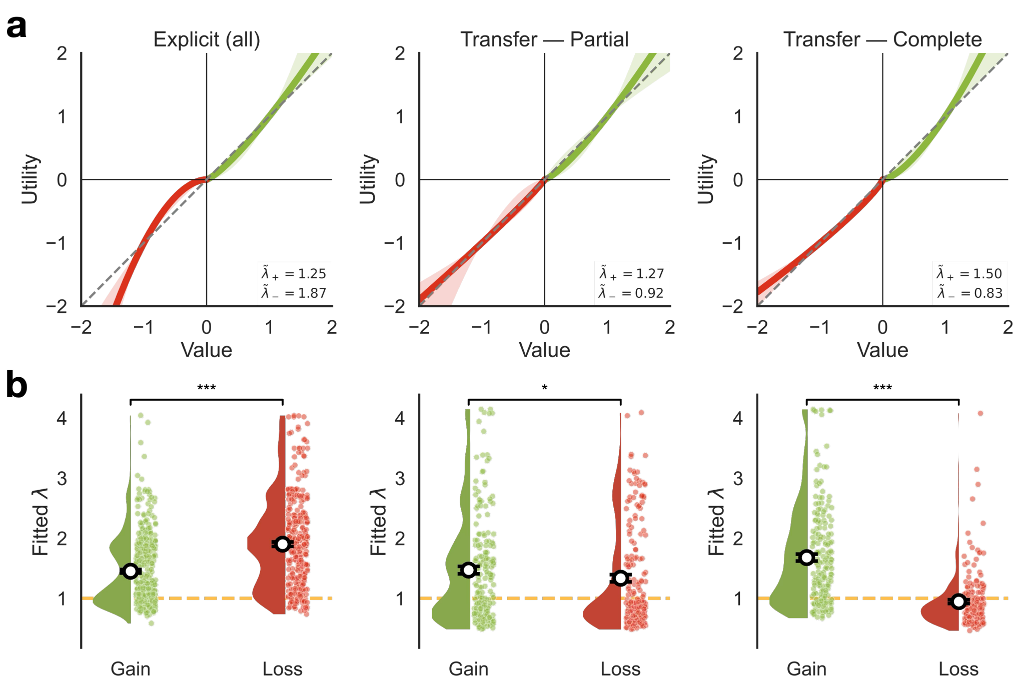

Fig. S1. Fitted exponent parameter of the utility function (λ). **a**, The utility function in gains and losses based on the median fitted λ to participants' decisions in description-based decisions (left) and experience-based decisions (separated by the feedback type: middle: partial, right: complete). **b**, The average of the fitted λ in gains and losses. ***P<0.001, *P<0.05 Mann–Whitney U test (N_{EXP}=200, N_{DES}=400). Error bars and shaded areas show s.e.m.

These results highlight two important findings. First, as previously suggested (4, 5), the description–experience gap extends beyond differences in probability weighting to include differences in the subjective valuation of gains and losses. Second, our results show that the gap is not reduced, and if anything, is amplified, under complete feedback conditions. This suggests that, contrary to previous assumptions, the gap does not primarily stem from undersampling of unchosen outcomes (4, 6, 7).

Utility modeling details. We quantified individual risk preferences in the transfer and lottery phases using a power–utility model with a lapse process and no inverse-temperature. For each participant we estimated two curvature parameters (one for gains and one for losses) denoted $\lambda_+ > 0$ and $\lambda_- > 0$, and a lapse rate $\varepsilon \in (0, 0.5)$. The instantaneous utility of an outcome x was $U(x)$ (Eq.1) and the (additively separable) EV of each option was $EU = p \cdot U(x)$. On trial t, with left (L) and right (R) options, the decision variable was

$$dV_t = p_t^{(L)}\, U\big(x_t^{(L)}\big) - p_t^{(R)}\, U\big(x_t^{(R)}\big) \qquad [2]$$

and the choice rule mixed a logistic choice with a uniform lapse:

$$P(\text{choose L} \mid dV_t) = (1 - 2\varepsilon)\, \sigma(dV_t) + \varepsilon, \quad \sigma(z) = \frac{1}{1+e^{-z}} \qquad [3]$$

Thus the slope of the logistic is implicitly fixed to 1 (no temperature), and ε captures stimulus-independent noise and inattention.

Estimation. We used per-subject maximum a posteriori (MAP) estimation on an unconstrained reparameterization

$$\theta = \big(\log \lambda_+,\, \log \lambda_-,\, \zeta\big), \qquad \varepsilon = \tfrac{1}{2}\,\text{logistic}(\zeta) \qquad [4]$$

with weakly informative Gaussian priors

$$\log \lambda_+,\, \log \lambda_- \sim \mathcal{N}(0, \sigma_\lambda^2), \quad \zeta \sim \mathcal{N}(\mu_\zeta, \sigma_\zeta^2) \qquad [5]$$

using $\sigma_\lambda = 0.6$, $\mu_\zeta = -3$ (centering $\varepsilon \approx 0.02$), and $\sigma_\zeta = 1.5$. For each participant we minimized the negative log posterior (negative log-likelihood plus negative log-prior) with L-BFGS-B from multiple random starts, which prevents boundary sticking and stabilizes fits when evidence is weak. The resulting MAP estimates $\hat{\lambda}_+, \hat{\lambda}_-, \hat{\varepsilon}$ summarize each participant's curvature in gains and losses and lapse rate. These parameters are then visualized with utility curves and compared across gain/loss and phase using half-violins with mean±SEM overlays and nonparametric tests as reported in the figure captions.

Choice repetition dynamics. We examined the probability of repeating a choice as a function of the previous trial's choice (whether it was risky or safe) in Figure S2. In partial feedback experiments, repetition rates were significantly influenced by both the type of previous choice and the valence of the decision. Participants were less likely to repeat a risky choice than repeating a safe choice (Wilcoxon signed-rank test; $Z = 26336.0$; $P < 0.001$). Additionally, they were more likely to repeat the same choice in the gain domain than in the loss domain (Wilcoxon signed-rank test; $Z = 24869.0$; $P < 0.001$). This pattern aligns with our macroscopic analyses, where the lowest repetition rates were observed for risky choices in the loss domain (0.457 ± 0.011), while the highest repetition rates were found after safe choices in the gain domain (0.64 ± 0.017). In contrast, under the complete feedback condition, we observed a different pattern: in contrast to the partial condition, repetition rates were higher after risky choices than after safe ones, which is consistent with the overall increase in risk propensity in these experiments (Wilcoxon signed-rank test; $Z = 25397.5$; $P < 0.001$). In addition, as shown in Figure S2, the repetition rates were significantly higher in gains compared to losses in complete condition (Wilcoxon signed-rank test; $Z = 33908.5$; $P = 0.0116$).

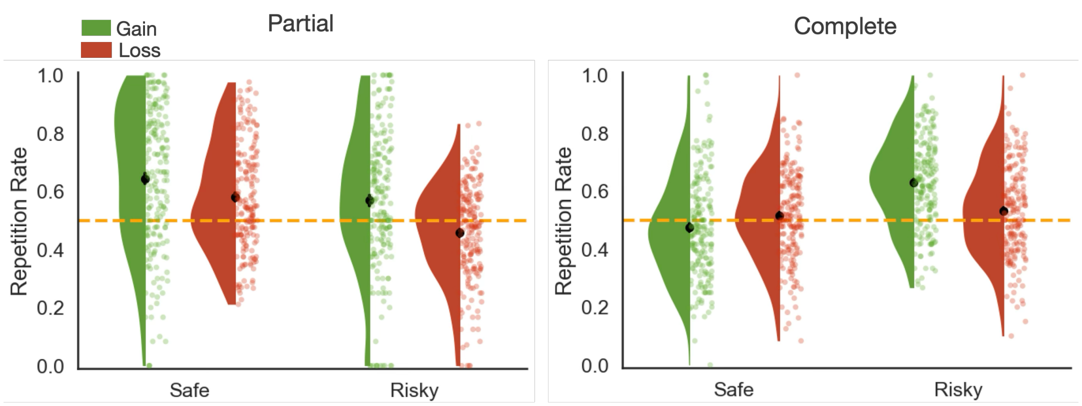

Fig. S2. Repetition rates of the safe and risky option, irrespective of EV maximization. Probability of repeating the previously chosen safe or risky option within the same context, shown for gain and loss contexts and displayed separately for partial (left) and complete (right) feedback experiments. All decision contexts are included; therefore, repetition rates reflect choice persistence independent of whether the repeated option is EV-maximizing. N=200, error bars show s.e.m.

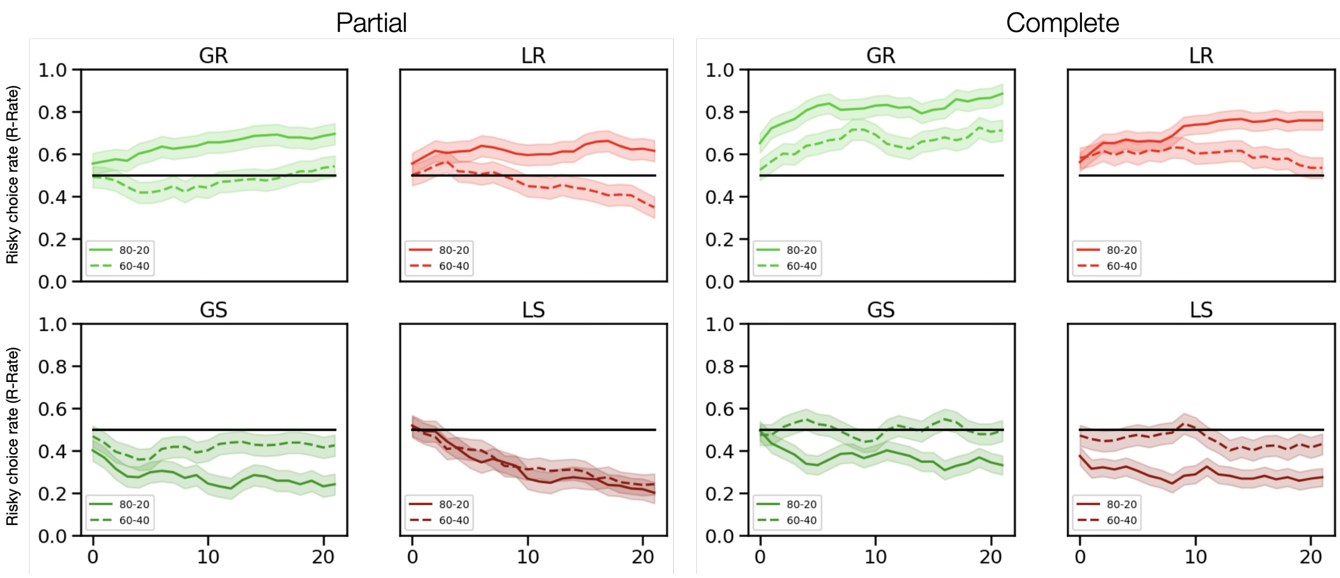

Fig. S3. Dynamics of risky choice rate within each choice context. Mean probability of choosing the risky option (R-rate) over successive presentations of the same context, shown separately for partial (left) and complete (right) feedback experiments. Rows correspond to the four contexts: GR and GS (gain contexts, green) and LR and LS (loss contexts, red). Within each panel, the solid line depicts the 80/20 condition and the dashed line the 60/40 condition; shaded bands are ±SEM across participants. The horizontal black line marks chance (0.5). The x-axis is trial number within a context. Note that in GR and LR the risky option is EV-maximizing, whereas in GS and LS the safe option is EV-maximizing; the figure therefore illustrates how feedback regime (partial vs. complete) and probability structure (80/20 vs. 60/40) jointly shape the dynamics of risk preference within each context. N=100, shaded areas show s.e.m.

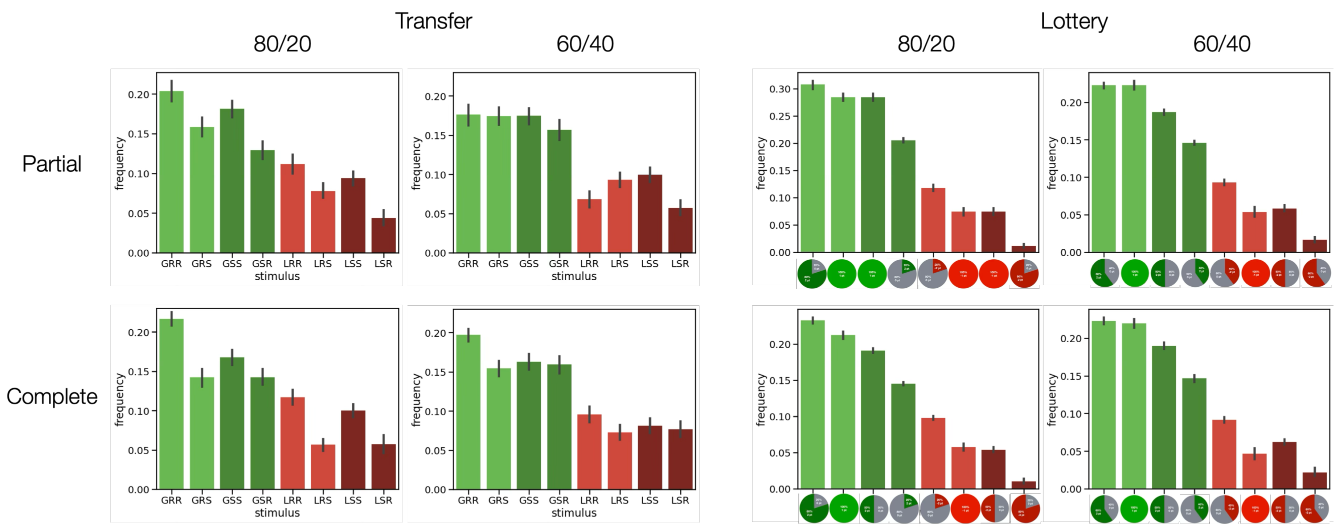

Fig. S4. Normalized frequency of choosing each stimulus in the transfer and lottery phases. Mean ($\pm$ SEM) choice frequency for each of the eight options, plotted separately for Transfer (two left columns) and Lottery/explicit (two right columns) phases, and for partial (top row) and complete (bottom row) feedback experiments. Within each phase, columns split the probability structures (80/20 and 60/40). In the Transfer panels, x-axis labels use a three-letter code in which the first two letters indicate the context (GR, GS, LR, LS) and the third letter indicates the option type (R = risky, S = safe); for example, LRS is the safe option from the LR context. In the Lottery panels, the x-axis shows the lotteries themselves; these match the payoffs and probabilities of the transfer stimuli, plus two additional 50/50 lotteries (gain: 2 or 0; loss: -2 or 0). Bars are normalized within participant and condition so that frequencies sum to 1 across the eight options. Green hues denote gain contexts and red hues denote loss contexts. Sample size N=100 per condition; error bars show the SEM across participants. Note that the lottery phase comprised 56 trials in all experiments except the 80/20 partial condition, which had 30 trials without the 50/50 lotteries, hence, the sixth and seventh bar is identical in the lottery of 80/20 partial experiment.

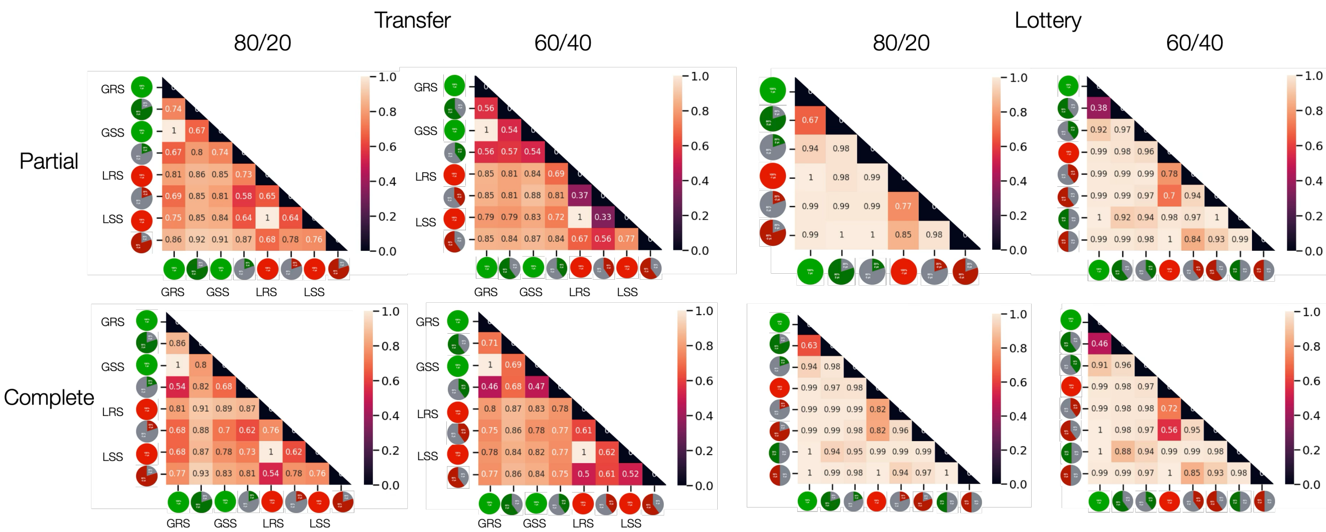

Fig. S5. Pairwise EV-maximization in transfer and lottery phases. Lower-triangular heat maps show the mean probability of choosing the expected-value (EV)–maximizing option for every ordered pair of stimuli (rows vs. columns). Panels are arranged by phase—Transfer (two left columns) and Lottery/explicit (two right columns)—and by feedback regime—partial (top row) and complete (bottom row). Within each phase, columns separate the probability structures (80/20 and 60/40). Axes list the options; in Transfer the three-letter code uses the first two letters for the context (GR, GS, LR, LS) and the final letter for Risky or Safe (e.g., LRS = safe option from the LR context). In Lottery panels, the options correspond to the lotteries themselves (matching the transfer stimuli), with two additional 50/50 lotteries (gain: 2 or 0; loss: -2 or 0) present in all experiments except the 80/20 partial condition. Color indicates EV-maximization rate from 0 to 1 (scale bars at right), with the numerical value printed in each cell; darker/lighter shades reflect lower/higher EV adherence. The diagonal is masked. Each pair was presented twice per participant (left–right order counterbalanced), and the cell value is the average EV-maximization rate across participants. Sample size N=100 per experiment; error bars are not shown in matrices, as cells represent across-participant means. Values near 1 indicate near-optimal EV choice, whereas values near 0 indicate systematic selection of the lower-EV option.

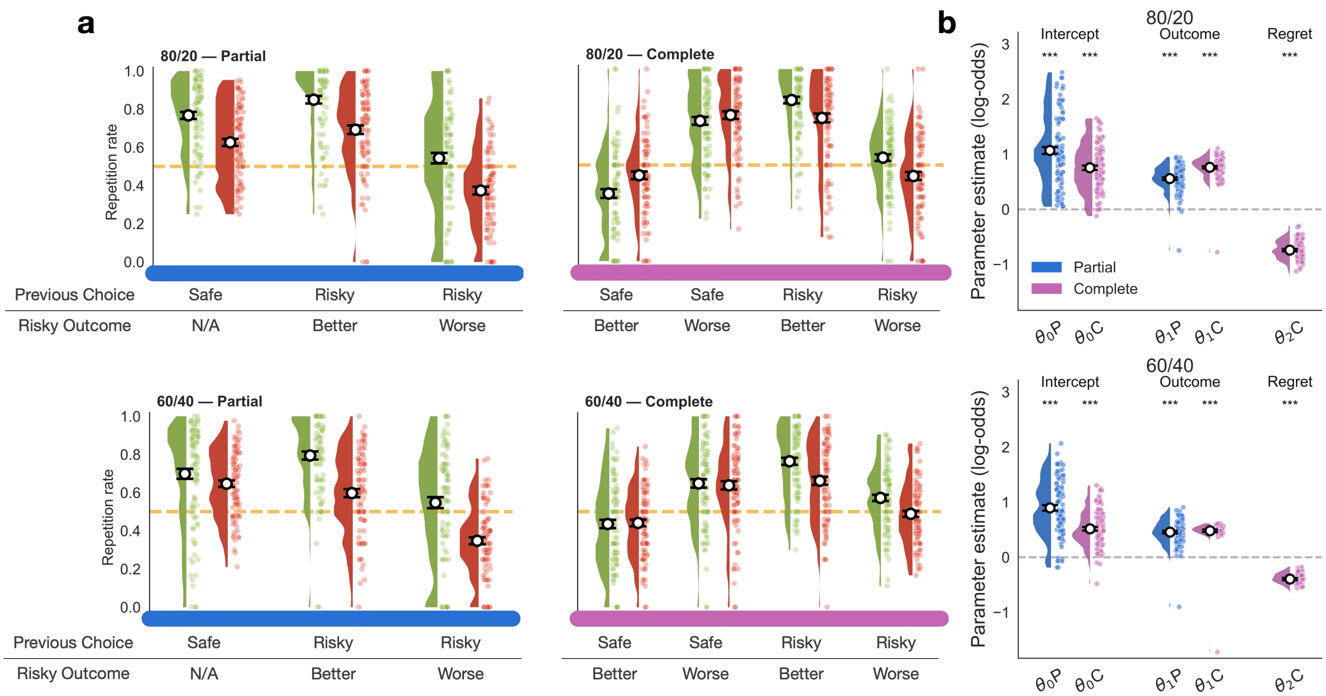

Fig. S6. Choice repetition rates across all experiments. **a**, The repetition ratio during the learning phase for each option type in experiments with partial feedback (left), complete feedback (right), 80/20 (top) and 60/40 probability conditions (bottom). The repetition ratio represents the probability of selecting the same option on consecutive trials within each choice context, regardless of whether the choice maximizes EV. **b**, Regression coefficients predicting the probability of stay in the next trial in 80/20 (top) and 60/40 (bottom) probability conditions. θ_0: Intercept, θ_1: chosen outcome, θ_2: counterfactual (regret) outcome. ***$P<0.001$, one-sample Wilcoxon against zero (N=100 per condition); Error bars represent s.e.m.

References

1. D Prelec, The probability weighting function. *Econometrica* pp. 497–527 (1998).
2. D Kahneman, A Tversky, Prospect theory: An analysis of decision under risk in *Handbook of the fundamentals of financial decision making: Part I.* (World Scientific), pp. 99–127 (2013).
3. NC Barberis, Thirty years of prospect theory in economics: A review and assessment. *J. economic perspectives* **27**, 173–196 (2013).
4. R Hertwig, I Erev, The description–experience gap in risky choice. *Trends cognitive sciences* **13**, 517–523 (2009).
5. S Palminteri, M Lebreton, The computational roots of positivity and confirmation biases in reinforcement learning. *Trends cognitive sciences* **26**, 607–621 (2022).
6. DU Wulff, M Mergenthaler-Canseco, R Hertwig, A meta-analytic review of two modes of learning and the description-experience gap. *Psychol. bulletin* **144**, 140 (2018).
7. R Hertwig, TJ Pleskac, T Pachur, *Taming uncertainty.* (Mit Press), (2019).