Online Supplemental Material for
Over-representation of extreme events in decision-making reflects rational use of cognitive resources

Falk Lieder
Helen Wills Neuroscience Institute, University of California, Berkeley

Thomas L. Griffiths
Department of Psychology, University of California, Berkeley

Ming Hsu
Haas School of Business, University of California, Berkeley

Deal or No Deal: Overweighting of extreme events in real-life high-stakes economic decisions

In the Main Text we found that people overweight extreme outcomes in judgment tasks and hypothetical and low-stakes decisions in the laboratory. Is this cognitive bias restricted to artificial laboratory tasks or does it also pervade the high-stakes economic decisions we make in real life? To answer this question, we analyze the high-stakes decisions of contestants in a popular TV game show called “Deal or No Deal” (Post, Van den Assem, Baltussen, & Thaler, 2008).

In this gameshow, the contestant is presented with up to 26 briefcases that contain prizes between \$0.01 and up to \$5,000,000. Knowing which prizes are available but not knowing which briefcase contains which prize, the contestant chooses one of the briefcases. In the first round, six of the remaining briefcases are opened and their contents are revealed. This narrows down the prize that might be in the contestant’s briefcase to the 20 remaining prizes. Next, the contestant receives a call from a banker offering to buy the contestant’s briefcase for a certain amount of money. If the contestant accepts the offer (“Deal”) the game is over and they receive the banker’s offer. If the participant rejects the offer (“No Deal”), then the second round begins. In the second round, five additional briefcases are opened and the participant receives a new offer that reflects the change in the expected value of their chosen briefcase brought about by the new information. Whenever the contestant rejects the offer the game proceeds to the next round and the process repeats. In the subsequent four rounds the number of briefcases opened is four, three, two, and one respectively. From there onward one briefcase will be opened on all subsequent rounds. The contestant’s chosen briefcase will be opened last, and when it is opened then the participant receives the prize contained therein and the game ends.

Post et al. (2008) extracted the round-by-round options and decisions of 151 contestants from the Netherlands, Germany, and the United States who were on the show between 2002 and 2007. Here, we reanalyze their data set to determine whether contestants overweighted extremely high prizes, such as $\$5,000,000$, and extremely low prizes, such as $\$0.01$, as predicted by utility-weighted sampling. To answer this question, we performed a formal model comparison between models that do versus models that do not overweight extreme outcomes. The results by Post et al. (2008) indicated that contestants evaluated prizes relative to a reference point that is adjusted gradually. They formalized this insight in a model called *dynamic prospect theory* (DPT). We therefore compared two dynamic reference point models with versus without utility-weighted sampling. In addition, we considered three models without dynamic reference points: a simple baseline model, a basic utility-weighted sampling model, and a basic representative sampling model. All of these models assume that contestants choose between the banker's offer o_r and their unknown prize $\tilde{X}$ based on which prizes $x^{(r)} = \{x_1^{(r)}, x_1^{(r)}, \cdots\}$ were still available in round r.

Models

The baseline model (m_{Random}) has one free parameter p_{accept}. It accepts the bank's offer o_r with probability p_{accept} and rejects it with probability $1 - p_{\text{accept}}$, that is

$$P(A = 1|o_r, x^{(r)}, m_{\text{Random}}, p_{\text{accept}}) = p_{\text{accept}}. \tag{1}$$

The static representative sampling model (m_{RS}) accepts the offer o_r with probability

$$P(A = 1|o_r, x^{(r)}, m_{\text{RS}}) = \Phi\left(\frac{\mathbb{E}\left[\Delta\hat{U}(o_r, x^{(r)})\right]}{\sigma_{\Delta\hat{U}(o_r, x^{(r)})}}\right), \tag{2}$$

where $\mathbb{E}\left[\Delta\hat{U}(o_r, x^{(r)})\right]$ is the expected value of the decision-maker's estimate of the difference between the utility of the offer and the utility of the unknown prize, that is

$$\mathbb{E}\left[\Delta\hat{U}(o_r, x^{(r)})\right] = \frac{1}{\#x^{(r)}} \cdot \sum_{x_i^{(r)} \in x^{(r)}} \left(u(o) - u(x_i^{(r)})\right), \tag{3}$$

where $\#x^{(r)}$ denotes the number of elements in the set $x^{(r)}$. $\sigma_{\Delta\hat{U}(o_r, x^{(r)})}$ is the standard deviation of this estimate, that is

$$\sigma_{\Delta\hat{U}(o_r, x^{(r)})} = \sqrt{\frac{1}{\#x^{(r)}} \cdot \sum_{x_i^{(r)} \in x^{(r)}} (u(o) - u(x_k))^2}. \tag{4}$$

As above, we assume that the utility-function normalizes each payoff by the range of possible outcomes according to efficient coding (Summerfield & Tsetsos, 2015):

$$u(o) = \frac{o}{\max\{x_1^{(r)}, \cdots, x_k^{(r)}\} - \min\{x_1^{(r)}, \cdots, x_k^{(r)}\}} + \varepsilon;\ \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon). \tag{5}$$

The static utility-weighted sampling model (m_{UWS}) is an analytic likelihood model that approximates the utility-weighted sampling model for decisions from description. It captures the central assumption that people approximate the expected utility difference in a stochastic fashion that over-weights extreme outcomes:

$$\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)}) = \sum_{x_i^{(r)} \in x^{(r)}} w_i \cdot \left(u(o) - u(x_i^{(r)})\right), \tag{6}$$

where the weight w_i of the i$^{\text{th}}$ potential value of the contestant's prize is defined by

$$w_i \propto P\left(\tilde{X} = x_i^{(r)}\right) \cdot \left|u(o_t) - u(x_i^{(r)})\right|^{\gamma}. \tag{7}$$

This formulation reflects that UWS over-simulates extreme outcomes and only partially corrects for it. How strongly UWS corrects for the bias of the sampling distribution depends on the number of samples and is captured by the parameter γ. As before, the utility function $u(o)$ normalizes outcomes by the range of the outcomes and adds normally distributed noise (Equation 16 in the Main Text). In order to obtain an analytic expression for the likelihood function we approximate the distribution of $\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})$ by a Gaussian with mean $\mathbb{E}\left[\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})\right]$ and variance

$$\sigma^2_{\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})} = \frac{1}{s} \cdot \mathbb{E}\left[\left(\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)}) - \mathbb{E}\left[\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})\right]\right)^2\right], \tag{8}$$

where s is a free parameter that approximately corresponds to the number of samples. Therefore, the likelihood function is given by

$$P(A = 1 \mid o_t, x^{(r)}, m_{\mathrm{UWS}}) = \Phi\left(\frac{\mathbb{E}\left[\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})\right]}{\sigma_{\Delta\hat{U}_{\mathrm{UWS}}(o_r, x^{(r)})}}\right). \tag{9}$$

The dynamic prospect theory model by Post et al. (2008) extends the utility-function of prospect theory, that is

$$u_{\mathrm{RP}}(o) = \begin{cases} (o - \mathrm{RP})^{\alpha}, & \text{if } o \geq \mathrm{RP} \\ -\lambda \cdot (\mathrm{RP} - o)^{\alpha}, & \text{else} \end{cases}, \tag{10}$$

by a dynamic model of its reference point (RP). According to this model, people gradually adjust their reference point RP to reflect the expected value of the possible payoffs in the current round, that is $\bar{x}^{(r)} = \frac{1}{\#x^{(r)}} \cdot \sum_{x_i^{(r)} \in x^{(r)}} x_i$. Because the adjustment is gradual, the reference point in round r is still influenced by the expected outcomes of earlier rounds:

$$\mathrm{RP} = B(x^{(r)}) \cdot (\theta_1 + \theta_2 \cdot d_t^{(t-2)} + \theta_3 \cdot d_t^{(0)}), \tag{11}$$

where $d_i^{(k)}$ is the relative difference between the average payoff in round i and the average payoff in round k, i.e. $d_i^{(k)} = \frac{\bar{x}^{(i)} - \bar{x}^{(k)}}{\bar{x}^{(k)}}$. Furthermore, the reference point is thresholded from below by the smallest possible payoff in the current round, and from above by the largest possible payoff in the current round. According to this model, the probability that the contestant will accept the deal is

$$P(A = 1 | o_r, x^{(r)}, x^{(r-1)}, x^{(0)}, m_{\mathrm{DPT}}) = \Phi\left(\frac{u_{\mathrm{RP}}(o_r) - \mathbb{E}\left[u_{\mathrm{RP}}(\tilde{X}) \mid \tilde{X} \in x^{(r)}\right]}{\sigma \cdot \sqrt{\mathrm{Var}\left[u_{\mathrm{RP}}(\tilde{X}) \mid \tilde{X} \in x^{(r)}\right]}}\right), \tag{12}$$

where σ is a free parameter that determines the choice variability.

The hybrid model $m_{\text{UWS+DPT}}$ extends utility-weighted sampling by the utility function with a dynamic reference point postulated by dynamic prospect theory. This model is a conceptual analogue of the utility-weighted learning model for decisions from description: The utility-weighted learning model gradually adjusts its estimate of the reward expectancy $\bar{u}(t)$ (Equation 36 of the Main Text) which could be interpreted as the reference point of a utility function $\tilde{u}(o) = r(o) - \bar{u}(t)$. In utility-weighted learning, it is the absolute value of $\tilde{u}$ with its dynamic reference $\bar{u}(t)$ that determines the probability weighting according to $\tilde{q}(o) = p(o) \cdot |\tilde{u}(o)|$ just like in the hybrid model. The hybrid model's decision variable is

$$\Delta\hat{U}_{\text{UWS+DPT}}(o_r, x^{(r)}) = \sum_{x_i^{(r)} \in x^{(r)}} w_i \cdot \left(u_{\text{RP}}(o) - u_{\text{RP}}\left(x_i^{(r)}\right)\right), \tag{13}$$

where the weight w_i of the i^{th} possible prize is defined by

$$w_i \propto P(\tilde{X} = x_i^{(r)}) \cdot \left|u_{\text{RP}}(o_r) - u_{\text{RP}}(x_i^{(r)})\right|^{\gamma}. \tag{14}$$

The model's choice probability is thus given by

$$P(A = 1 \mid o_r, x^{(r)}, x^{(r-1)}, x^{(0)}, m_{\text{UWS+DPT}}) = \Phi\left(\frac{\mathbb{E}\left[\Delta\hat{U}_{\text{UWS+DPT}}(o_r, x^{(r)})\right]}{\sigma_{UWS+DPT}}\right), \tag{15}$$

where $\sigma_{UWS+DPT} = \frac{1}{s} \cdot \sqrt{\text{Var}\left[\Delta\hat{U}_{\text{UWS+DPT}}(o_r, x^{(r)})\right]}$.

Priors distributions on parameters of the models of the Deal No Deal dataset

For parameters that occurred in multiple models, the prior was always the same across all models.

For the random choice model the prior on the choice probability was the standard uniform distribution over the interval $[0, 1]$:

$$p(p_{\text{accept}}) = \begin{cases} 1, & \text{if } 0 \leq p_{\text{accept}} \leq 1, \\ 0, & \text{else} \end{cases} . \tag{16}$$

For the representative sampling model the prior on the noise parameter σ_ε of the stochastic utility function was a standard uniform distribution over the range $[0, 1]$ because 0 corresponds to no noise whereas 1 would entails that the magnitude of the noise is as high as the highest possible expected utility gain. The prior on the number of samples s was a uniform distribution over the range $[1, 1000]$ because the minimum number of samples is 1 and 1000 would be more than sufficient to estimate the expected utility gain accurate.

For the basic utility-weighted sampling model the prior on the utility-weighting parameter γ was a standard uniform distribution over the range $[0, 1]$ because 0 corresponds to no bias and 1 correspond to drawing only a single sample. The priors on the variability parameter σ and the number of samples s were the same as for the representative sampling model.

For the dynamic prospect theory models, the prior distribution on the exponent α of the utility was the uniform distribution over this parameters admissible range $[0, 1]$, because the utility function is concave if and only if $0 \leq \alpha < 1$ and linear for $\alpha = 1$. The prior distribution on the slope in the domain of losses λ was defined to be uniform distribution over the range $[0, 5]$ because it cannot be negative and the estimates obtained by ? (?) suggested that it is always smaller than five. The prior on the weights $\theta = (\theta_1, \theta_2, \theta_3)$ determining the updates of the reference point was a multivariate standard normal distribution:

$$p(\theta) = \mathcal{N}\left(\theta; \mu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right). \tag{17}$$

The prior on the noise parameter σ was an exponential distribution with mean 1 to express that the expected variability is the variance that is multiplied by σ and that less noise is more likely than more noise:

$$p(\sigma) = \exp(-\sigma). \tag{18}$$

For the combined model integrating UWS with dynamic prospect theory the priors on its parameters were the same as those reported above: The priors on the parameters of the utility function $(\alpha, \lambda, \theta)$ were the same as for the DPT model and the priors on the utility weighting parameter (γ) and the choice variability parameter σ were the same as those in the basic UWS model.

Results

We estimated the model parameters from all choices of the 151 contestants from the Netherlands, Germany, and the US using the maximum-a-posteriori method with the priors specified in Appendix D. To find these estimates we used a global optimization algorithm known as infinite-metric Gaussian process optimization (Kawaguchi, Kaelbling, & Lozano-Pérez, 2015). For all models this optimization algorithm was run for 1000 iterations. We then use the global maximum found by this derivative-free algorithm as the starting point for the gradient-based quasi-Newton algorithm (fminunc in Matlab 2015b) which was run until convergence. To find out which of these five models best explains people's choices in this high-stakes game show, we performed Bayesian model selection (Kass & Raftery, 1995) with a uniform prior over the five models. This method measures the goodness of each model by the marginal likelihood of the data given that model, which integrates over all possible settings of the model's parameters. The marginal likelihood thereby penalizes each model's fit by a complexity penalty that accurately reflects the model's flexibility and not just its number of parameters. Here, we estimate the marginal likelihood of each model using the Laplace approximation (Tierney & Kadane, 1986). Bayesian model selection then compares pairs of models by computing their Bayes factor (BF), which is the ratio of their posterior probabilities given the data.

Figure 1 shows the results of the model comparison. Consistent with the results of Post et al. (2008) we found that models with a dynamic reference point explained the contestants' decisions better than models with a fixed utility function. Most importantly, utility-weighted sampling performed better than unweighted decision mechanisms for either type of utility function: For models with the static, normalized stochastic utility function (Equation 16 of the Main Text), we found that our basic utility-weighted sampling model explained the contestants' choices substantially better than random choice ($\text{BF}_{\text{UWS, random}} = 3.7 \cdot 10^{50}$) or representative sampling ($\text{BF}_{\text{UWS, RS}} = 2.3 \cdot 10^{18}$). Among the models with dynamic utility functions, utility-weighted sampling with a dynamic reference point explained the contestants' choices substantially better than the unweighted decision mechanism of dynamic prospect theory ($\text{BF}_{\text{DPT+UWS, DPT}} = 1.1 \cdot 10^{7}$). In both cases, the data provided decisive evidence for utility-weighted sampling, because the Bayes factors are larger than 100 (Kass & Raftery, 1995). Furthermore, the models with the dynamic utility function captured the data significantly better than their counterparts with the static utility function ($\text{BF}_{\text{DPT+UWS, UWS}} = 1.4 \cdot 10^{16}$, $\text{BF}_{\text{DPT, RS}} = 5.6 \cdot 10^{30}$). Post et al. (2008) used a different task analysis according to which contestants choose between the current offer and anticipated next offer. To evaluate this alternative perspective, we adapted all models to their alternative task analysis and recomputed the model evidence scores. Quantitative model comparisons provided very strong evidence for our task analysis over the one by Post et al. (2008).

Our analyses support the hybrid model ($m_{\text{UWS+DPT}}$) that combines utility-weighted sampling with a utility function with a gradually adjusting reference point. For this model, the estimated utility-weighting coefficient $\hat{\gamma}$ was significantly larger than zero ($\hat{\gamma} = 0.5721$, 95% CI: $[0.5668; 0.5774]$). This is consistent with the hypothesis that contestants performed utility-weighted sampling with an intermediate number of samples. Furthermore, fixing the probability-weighting parameter to 0, which yields the DPT model, led to a significantly worse fit that is not offset by the corresponding gain in parsimony. The estimated value of the number of samples was $s \approx \frac{1}{\sigma^2} = 11.88$ suggesting that contestants simulated their potential prize about 12 times on average. Note that in UWS some of these imagined outcomes would have been identical so that the number of considered prizes can be smaller. Note also, that s only approximately corresponds to the number of samples, because the number of samples is also reflected by the value of γ. The maximum-a-posteriori estimates for the remaining parameters were $\hat{\alpha} = 0.6721$, $\hat{\lambda} = 1.1346$, and $\hat{\theta} = (1.0049, -0.0070, -0.0313)$. For these parameter values, the hybrid model correctly predicts 87.1% of the contestants choices, meaning that 87.1% of the time the predicted probability of the contestant's choice was greater than 0.5.

Discussion

In conclusion, we found that people overweight extreme potential outcomes not only in hypothetical and low-stakes laboratory tasks but also in high-stakes real-life decisions whose outcomes do count. This finding is consistent with utility-weighted sampling. In fact utility-weighted sampling predicts that the overweighting of extreme outcomes is larger for high-stakes decisions than for low-stakes decisions, because their highest possible outcomes are more extreme. However, we cannot conclude that the contestants' choices were resource-rational because the normative status of the dynamic reference point of the winning model's utility function is unclear. On the one hand, the reference point can be seen as an estimate of the expected utility gain $\mathbb{E}[\Delta U]$. Therefore, the difference between what would otherwise be the outcome's utility and the reference point can be interpreted as an approximation to the term $u(o) - \mathbb{E}_p[\Delta U]$ used in optimal importance sampling (Equation 10 in the Main Text). Hence, the model's use of the absolute value of the reference-point-dependent

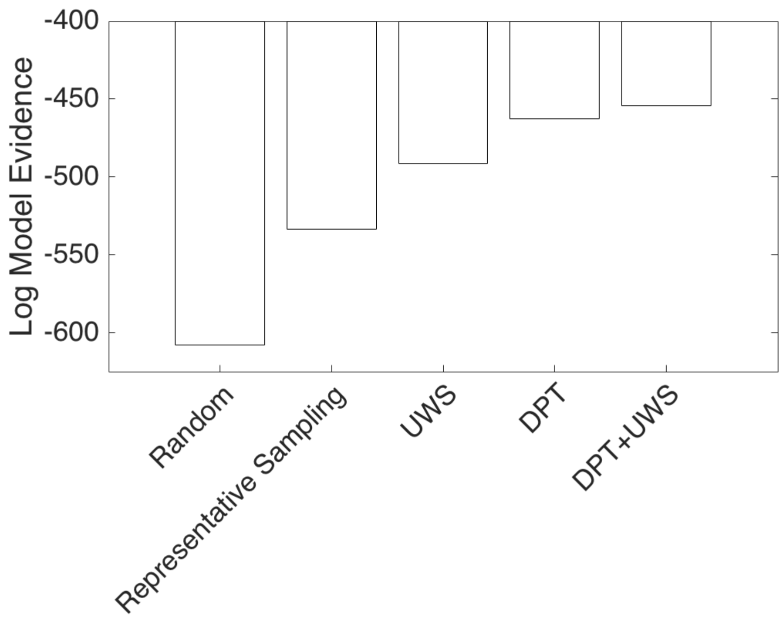

Figure 1. Model comparison for “Deal No Deal” data set. Better models have a *higher* log-model evidence.

utility to weight the probabilities of the corresponding outcomes can be interpreted as an approximation to optimal importance sampling (Equation 10 in the Main Text). Since the reference point is an estimate of the expected utility gain, it is rational to update it when additional outcomes are observed. The update equation for the dynamic reference point emphasizes recent outcomes. This is consistent with estimating the expected utility gain of decisions in dynamic environments like Deal No Deal where the expected utility gain of future decisions changes every time an outcome is observed. Therefore, the winning model could be a rational extension of UWS to dynamic decision environments. However, this rational interpretation has to be taken with a grain of salt, because the update rule for the dynamic reference point was not derived from first principles, and the normative status of other aspects of the utility function is also unclear. Deriving a fully-principled form of UWS for dynamic environments and testing it against the models examined here is a possible direction for future work.

Interestingly, the estimated number of samples (s) was substantially higher for the high-stakes decisions in *Deal or No Deal* than for the low-stakes decisions in the Technion choice prediction competition. This finding is consistent with the hypothesis that people make rational use of their finite time and limited computational resources: Raising the stakes increases the expected gain in reward for performing an additional simulation but its time cost remains the same. Once the expected gain in reward exceeds the time cost, it becomes resource-rational to perform an additional simulation.

Payoff-variability effects in decisions with very many possible outcomes

The decisions from experience simulated above were very simple in that each option had only two possible outcomes, but in the real world a choice can have very many outcomes. To investigate whether utility-weighted sampling can capture these more complex decisions from experience, we simulated Experiment 1 by Barron and Erev (2003) where outcomes were sampled from normal distributions with different means and variances. Participants were instructed to maximize their earnings by repeatedly choosing between two buttons but received no further information about the task other than that the experiment would last for about 30 minutes. After each decision an outcome was sampled from the chosen option's payoff distribution and shown to the participant. There were three groups who made 200 choices each: In the first condition the outcome of the first option was sampled from a normal distribution with mean 25 and standard deviation 17.7, and the outcomes of the second option were sampled from a normal distribution with mean 100 and standard deviation 354. The second condition was like the first, except that both means were shifted upwards by 1000. The third condition was like the second one except that the standard deviation of the second option was reduced to 17.7. Barron and Erev (2003) found that the high variability of the payoffs in the first and second condition interfered with people's ability to discover that the first option was better than the second option. This is known as the *payoff-variability effect*.

We simulated the experiment with the parameters estimated from the experiment by Madan, Ludvig, and Spetch (2014). The largest and the smallest possible outcome ($o^c_{\max}$ and $o^c_{\max}$ in Equation 16 of the Main Text) were initialized by ± 10 and continuously updated to always equal the largest and smallest outcome observed so far respectively. We conducted 1000 simulations for each of the 3 conditions. Figure 2 shows the average frequency with which our model choose the option with the higher expected value as a function of time in the experiment. To evaluate the effect of learning we compared the average choice frequencies between the first 5 trials and the last 100 trials. We found that the model captures the outcome variability effect (see Figure 2): When the payoff variability of the better option was large compared to the expected values and their difference (Condition 1), then participants came to avoid the better option as their choice frequency dropped from 51.3% to 43.3% ($\chi^2(1) = 123.17, p < 10^{-15}$). When the means were increased to be substantially higher than the payoff variability (Condition 2), then the frequency of the maximizing choice increased slightly to 49.37% ($\chi^2(1) = 740.4, p < 10^{-15}$) but remained below chance level ($p < .0001$), and their choice frequency did not change significantly over time ($\chi^2(1) = 3.07, p = 0.08$). But when the payoff variability was reduced (Condition 3), then people learned to choose the better option: the predicted frequency of choosing the better option rose significantly from 51.5% in the first five trials to 60.8% in the last 100 trials ($\chi^2(1) = 174.5, p < 10^{-15}$) and surpassed the chance level ($p < 10^{-15}$). Thus, our utility-weighted learning model correctly predicted the detrimental effects of payoff variability on decisions from experience.

This illustrates that utility-weighted sampling can capture people's ability to make decisions with (infinitely) many possible outcomes as well as people's biases in the face of high payoff variability. According to utility-weighted sampling, people's apparently irrational aversion to choices with superior expected value but higher payoff variability in

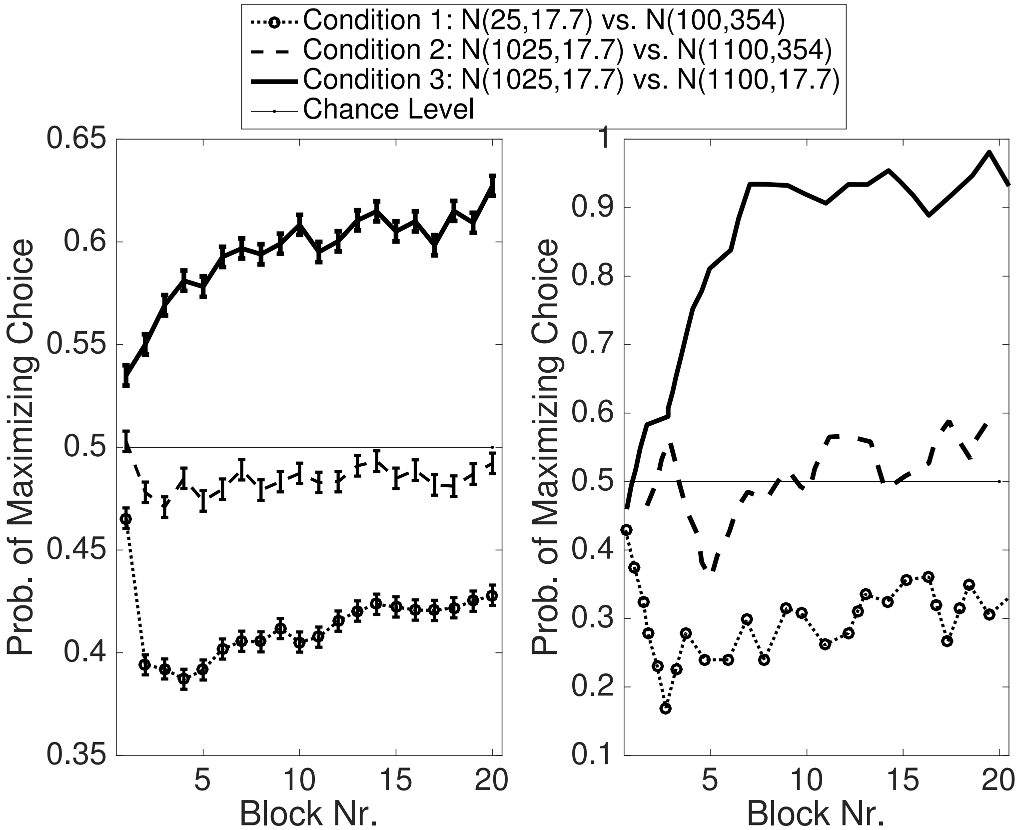

Figure 2. Simulation of Experiment 1 by Barron and Erev (2003) according to the utility-weighted learning model. Each line represents the frequency of choosing the first option in each of the 20 blocks averaged across 1000 simulations. The error bars indicated standard errors of the mean.

decisions from experience arises because people overweight the salient memories of large losses.

Comparison of the risk preferences of UWL to people's risk preferences in the Technion choice prediction tournament

We found that the average risky-choice frequency of the UWL model was $41.6 \pm 2.0\%$ whereas the average risky-choice frequency of people was $38.1 \pm 2.2\%$. This shared overall preference for the safe option suggests that utility-weighted learning captures that people underweight rare gains in classic decisions-from-experience paradigms. However, according to a paired t-test, the predictions of UWL were significantly less risk averse than people ($-3.5 \pm 1.4\%$,$t(59) = 2.59$, $p = 0.01$). This apparent bias towards risk seeking does, at least in part, result from a regression towards the "mean" frequency of 50%. Consistent with this interpretation, UWS was less risk averse than people primarily when they chose the risky option less than half of the time (36.47% vs. 31.05%; $t(46) = 3.68, p < .0006$), but when they chose it more than 50% of the time, then UWL was less risk-seeking than people

(60.14% vs. 63.58%[1]). For this particular data set, regression to the mean increased the overall frequency of choosing the risky option because people were risk averse in 47 of the 60 problems, and chose the risky option only 38% of the time on average.

To understand why the UWL model's risk preferences were less extreme than human risk preferences, we inspected the decision problems on which UWL was much more risk-seeking than people. We found that the two problems where the bias was largest were the only problems in which the risky option was dominated by the safe option. In these problems the outcome of the safe option was slightly higher than the best possible outcome of the risky option that occurred with a frequency of 97%. Here, people chose the dominated risky option only about 15% of the time, whereas UWL chose it 40% of the time. The choice frequency of UWL was closer to 50% because the difference between the safe outcome and the high outcome of the risky option was small relative to the noise of its utility function.

Examining these results, it seems that people can exploit obvious dominance better than UWL. For instance, when people recognize dominance they can switch to a different decision strategy (Lieder & Griffiths, 2015, under review). People's advantage on problems with obvious dominance contributed to the apparent bias of UWL, because the safe option dominated the risky option twice as often as vice versa. When the three problems with dominance were excluded, the bias decreased to 2.9% but remained statistically significant ($t(56) = -2.29, p = 0.0257$). We therefore also inspected the problem where UWL had the third largest bias towards risk seeking. In this problem the probability of the high payoff was very low ($p_{\text{high}} = 0.06$), and the low payoff (o_{low}) differed from the sure payoff (o_{sure}) by less than 2% of its value (-20.5 vs. -20.3). For this problem many participants may thus never have sampled the high payoff. This would again create the dominance scenario in which the noisy utility function of UWL induces more random choices, and hence more risk-seeking, than the heuristic that people appear to use for problems with dominance. When we additionally removed the six problems where the safe option was very likely to slightly dominate the risky option according to the sampled outcomes ($p_{\text{high}} < 0.1$ and $0 < \frac{o_{\text{sure}} - o_{\text{low}}}{\max\{|o_{\text{sure}}|, |o_{\text{high}}|\}} < 0.025$), then the average difference in the frequency of risk-seeking dropped to 1.5% and was no longer statistically significant ($t(49) = -1.0, p = 0.32$).

Taking these results into account, it appears that the risk-seeking bias we observed in the predictions of the UWL model may arise from situations where the safe option dominates the risky one according to their observed outcomes. On those trials UWL does not capture people's choices frequencies. One possible reason for this is that the model's assumptions about the normalized, stochastic utility function are invalid when the difference between the observed outcomes is very small relative to the range of possible outcomes. Another possible reason is that people switch to a specialized heuristic when they encounter dominance. Investigating these possibilities is an interesting direction for future research.

UWS captures that people's performance approaches optimality as the options become more different

Resource-rationality predicts that as the stakes increase people should become increasingly more accurate. Consistent with this prediction, Jarvstad, Hahn, Rushton, and

[1]This difference was not statistically significant ($t(12) = -1.37, p = 0.20$), but the test was highly underpowered because people were risk-seeking for only 13 of the 60 problems.

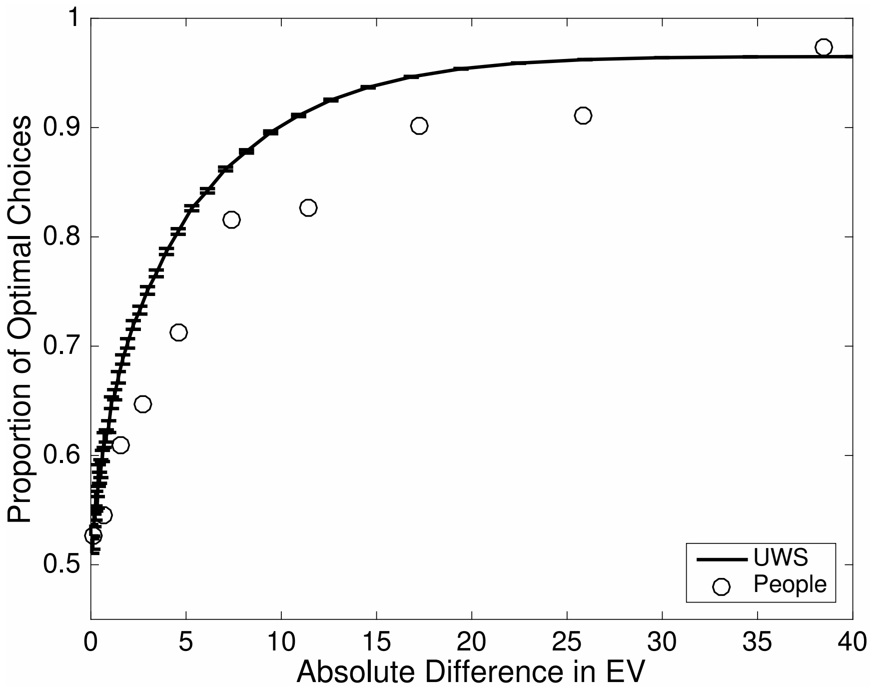

Figure 3. UWS captures that people reach (near) optimal performance as the difference between the options' expected values increases. The human data was taken from Jarvstad, et al. (2013). Error bars enclose 95% confidence intervals.

Warren (2013) found that as the difference between two gambles' expected values increases people's decision quality increases gradually. To test whether UWS can capture this effect, we simulated the decisions from description experiment from Jarvstad et al. (2013) according to our binary choice model with the parameters estimated from the Technion tournament for decisions from description. As shown in Figure 3, our model captures that people err primarily when the options' expected values are very close but come to choose the optimal action almost 100% of the time as the difference in expected value increases (Jarvstad et al., 2013). These findings suggest that the biases and suboptimalities that classic laboratory experiments have demonstrated for choices between options with very similar expected values are not representative of decision-making in the real world where the (relative and absolute) difference in the options' expected values tends to be larger. Instead, the fact that the difference in expected value has to be small to elicit biases and sub-optimalities in people is consistent with the rational use of limited cognitive resources. Indeed, it is resource-rational to save time and mental effort when the return for investing additional cognitive resources is less than their cost.

Comparison to previous theories of memory, judgment, and decision-making

Comparison to previous theories of memory and frequency judgment. Anderson's rational analysis of memory demonstrated that the availability of a memory rationally reflects how likely it is going to be needed according to its frequency and recency of occurrence (Anderson, 1990; Anderson & Schooler, 1991). Here, we have demonstrated another rational aspect of availability: eventualities that are more important for making a decision are more available in memory than their equally probable counterparts. We have shown that the rational availability of extreme events can account for the memory biases observed by Madan et al. (2014). Our model of frequency judgments is consistent with the availability-by-recall model (Hertwig, Pachur, & Kurzenhäuser, 2005; Pachur, Hertwig, & Steinmann, 2012) of the availability heuristic (Tversky & Kahneman, 1973), but it goes one step further by predicting how many instances of each event people will recall from memory and how this number depends on the event's frequency and extremity. This allowed our model to correctly predict that people overestimate the frequency of extreme events regardless of whether they are rare as in Experiment 1 or frequent as in the Experiment by Madan et al. (2014). Our theory thereby reconciles seemingly irrational availability biases with Anderson's rational analysis of memory, and our results resolved the open question whether biases in frequency estimation are due to availability or regression to the mean (Hertwig et al., 2005) in favor of a rational version of availability.

Comparison to previous theories of decisions from experience. Which events are retrieved from memory is critical to people's decisions from experience. Several models of experience based choice assume that memory recall rationally reflects past experience (Lejarraga, Dutt, & Gonzalez, 2010; N. Stewart, Chater, & Brown, 2006) and this is also true of the exploratory sampler with recency that won the Technion choice prediction competition (Erev et al., 2010). Concretely, instance-based learning theory (C. Gonzalez, Lerch, & Lebiere, 2003; C. Gonzalez & Dutt, 2011; Lejarraga et al., 2010) assumes that previous instances of similar past decisions are recalled with a probability that reflects their frequency and recency according to Anderson's rational analysis of memory (Anderson, 1990; Anderson & Schooler, 1991). Our analysis suggests that these models' assumption of rational memory recall implies that events with extreme utilities should be recalled more frequently than would be warranted by how often they have been encountered in the past, whereas equally frequent events with unremarkable utilities should be recalled less often. Other models of decisions from experience emphasize that people's memory is fallible (Hawkins, Camilleri, Heathcote, Newell, & Brown, 2014; Marchiori, Di Guida, & Erev, 2015). The Technion choice prediction tournament also included reinforcement learning models and an ACT-R model of instance-based learning, and Plonsky, Teodorescu, and Erev (2015) have proposed a new model according to which people's decision mechanisms are tuned to dynamic environments. Yet, as far as we know, no previous model of decisions from experience captures the over-weighting of events with extreme utilities. We now compare our utility-weighted learning model to each of these theories in turn

Decision-by-sampling assumes that outcomes are sampled from memory in a manner that reflects the structure of the environment but is also subject to availability biases (N. Stewart et al., 2006). This view is consistent with our model but decision-by-sampling does not explain why some past experiences are more available than others. The explorative sampler with recency (Erev et al., 2010; pp. 29-31) stochastically chooses to explore or to exploit. When it explores, it chooses at random. When it exploits, it estimates each option's value and chooses the option with the highest value estimate. To estimate an alternative's value the sampler retrieves a randomly generated number of past experiences with that alternative from memory. The retrieved experiences always include the most recent outcome and all earlier experiences are retrieved with equal probability. The retrieved outcomes are regressed towards the mean outcome and passed through a concave utility function. In the Technion choice prediction competition, the performance of the exploratory sampler with recency was not significantly higher, and its lower mean-squared deviation might reflect that it captures that people face an exploration-exploitation dilemma and assume that the environment is changing. Incorporating this idea into the UWL model might lead to even better predictions.

The exemplar-confusion model by Hawkins et al. (2014) assumes that people store a new memory trace every time they experience an outcome. Every time a new memory trace is added to the store of the chosen lottery, every stored memory trace has a small probability that its outcome will be confused. It this happens, then that memory's outcome will be replaced by a value that is sampled uniformly at random from the set of values that have been observed for that lottery so far regardless of how often each value has been observed. This model predicts both choices and probability judgments by assuming that people average over all of their memory traces. When evaluated on the Technion choice prediction tournament for repeated decisions from experience, the exemplar-confusion model's risk preferences agreed with people's risk preferences slightly less often than the risk preferences of our utility-weighted learning model (83.3% vs. 90% agreement). While the exemplar confusion model focusses on errors during encoding, the noisy retrieval models focus on errors during retrieval (Marchiori et al., 2015). Concretely, noisy retrieval models assume that people retrieve only a very small number of experienced outcomes and erroneously recall outcomes of unrelated decisions and use them as if they pertained to the current choice. These models reconcile the under-weighting of rare events in repeated decisions from experience with their being over-weighted in one-shot decisions under risk, and the overestimation of their probabilities. However, none of these models captures the effect of extremity on memory recall, frequency judgments, and choice. Furthermore, in contrast to these theories, UWS is based on a rational model of memory.

The basic reinforcement learning model from the Technion choice prediction tournament (Erev et al., 2010) probabilistically chooses the option with the higher recency-weighted average payoff, and the normalized reinforcement learning model normalizes the options' weighted average values by the variability of their payoffs. The value assessment model by Barron and Erev (2003) and the model by Shteingart, Neiman, and Loewenstein (2013) are similar to the basic reinforcement learning model. The main difference in the model by Shteingart et al. (2013) is that it gives special weight to the first outcome of each action, and the model by Barron and Erev (2003) includes a separate exploration mechanism and a utility function that captures loss aversion. These models differ from UWL in that they do not simulate potential outcomes and do not overweight extreme outcomes relative to moderate outcomes. As reported above, our model achieved a significantly lower mean-squared deviation than the basic reinforcement learning model. While the normalized reinforcement learning model was about as accurate as our UWL model by assuming that payoff variability has a deterring effect, our model provides a mechanistic explanation for why this was the case for many problems in the choice prediction competition. As reported above, the instance-based learning model by Lejarraga et al. (2010) predicted decisions in the Technion choice prediction tournament significantly better than our UWL model. This might be because it incorporates additional psychological insights such as people's optimism in the face of uncertainty and the implicit assumption that the environment is changing.

Incorporating these assumptions into the UWL model or incorporating the heightened availability of extreme events into the instance-based learning model might lead to even better predictions. The ACT-R model of instance-based learning (T. C. Stewart, West, & Lebiere, 2009) was very similar to the model by Lejarraga et al. (2010). The main difference was that the ACT-R model learned separately about the contexts established by the history of previous choices and outcomes. Concretely, the model by T. C. Stewart et al. (2009) recalls only those previous outcomes that followed the sequence of choices and outcomes observed in the preceding k trials. Like, the ACT-R model of instance-based learning, the contingent average and trend (CAT) model by Plonsky et al. (2015) postulates that people assume that the same choice will lead to different outcomes depending on the outcomes that preceded it. Concretely, this model assumes that people learn a separate reward expectation for every possible sequence of the k preceding outcomes. In addition, the model probabilistically responds to trends: If the last three outcomes suggest an increase or decrease in an action's average payoff, then the CAT model estimates the expected value of that action by its most recent payoff with some probability. This model captures people's sensitivity to patterns, the underweighting of rare events, and the non-monotonic effect of recency on the weight of previous outcomes (Plonsky et al., 2015). The CAT model is complementary to UWS in the sense that it describes how people learn when they assume that the environment is dynamic whereas UWL describes how people learn when they assume that the environment is static. The two theories could be combined into an integrated model of utility-weighted learning in dynamic environments. Overall, comparing UWL to models of decisions from experience highlights that extending our model to dynamic environments is an important direction for future work.

Comparison to previous theories of decisions from description. Most descriptive theories of decisions from description modify expected utility theory in order to account for some of the ways in which people deviate from its predictions (Starmer, 2000). Some of these theories postulate that people's choices optimize not only the expected utility of their payoffs but also additional experiential qualities like regret (Loomes & Sugden, 1982) or disappointment (Bell, 1985; Loomes & Sugden, 1984, 1986), or assume that people have additional preferences about the variance (Allais, 1979) and skewness (Hagen, 1979) of a prospect's outcome distribution. Other theories maintain that people maximize their subjective expected utility with respected to weighted probabilities (Edwards, 1962; Quiggin, 1982; R. Gonzalez & Wu, 1999). By contrast to all of these theories, utility-weighted sampling is derived from the assumption that people are striving to maximize the expected utility of their outcomes but are constrained by their finite time and limited cognitive resources. Hence, unlike these earlier theories, UWS does not propose that people behave as if they were optimizing a certain preference function. Instead, UWS is a procedural theory. Like prospect theory (Kahneman & Tversky, 1979), cumulative prospect theory (Tversky & Kahneman, 1992), dynamic prospect theory (Post et al., 2008), rational inattention theory (Sims, 2003), and salience theory (Bordalo, Gennaioli, & Shleifer, 2012), it is informed by people's limited cognitive resources, but it goes beyond these theories by providing a decision strategy that is optimal given the constraints imposed by those limited resources under certain assumptions. While rational inattention (Sims, 2003) prescribes how much time and attention a decision-maker should allocate to each of their choices, but it does not specify how that decision should be made. Utility-weighted sampling complements rational inattention by specifying a decision strategy that makes the best possible use of the limited amount of attention that has been allocated to a choice. Conversely, rational inattention complements UWS by specifying how many samples it should generate.

What sets UWS apart from all of theories mentioned above, is that it provides a process model. Process models of decisions from description that are similar to UWS include the *priority heuristic* (Brandstätter, Gigerenzer, & Hertwig, 2006), *decision-by-sampling* (N. Stewart et al., 2006), the *exemplar-confusion* model (Lin, Donkin, & Newell, 2015), query theory (Johnson, Häubl, & Keinan, 2007; Weber et al., 2007), *selective integration* (Tsetsos et al., 2016), *drift-diffusion models* of value-based choice (Shadlen & Shohamy, 2016; Krajbich, Armel, & Rangel, 2010; Krajbich & Rangel, 2011), and the *associative accumulation model* (Bhatia, 2013). We now discuss the similarities and differences between these models and UWS.

The priority heuristic (Brandstätter et al., 2006) is a fast-and-frugal heuristic for binary decisions from description. It sequentially compares the two alternatives on a list of criteria and stops after comparing the options on the first criterion on which they are sufficiently different. In the domain of gains, the priority heuristic first considers the minimum gain. If that does not lead to a decision, then it considers the probability of the minimum gain, and the remaining criteria are the maximum gain and its probability. UWS is similar to the priority heuristic in that it prioritizes important information. On the other hand, the two theories are very different in that UWS uses the probabilities to simulate outcomes whereas the priority heuristic treats the probabilities as just another attribute. Furthermore, the prioritization of UWS is stochastic allowing it to predict choice probabilities. While both the priority heuristic and UWS qualitatively capture the violations of expected utility theory in decisions from description, we found that UWS outperformed the priority heuristic on the Technion choice prediction tournament. Unlike the priority heuristic, UWS was derived from first principles and is more widely applicable.

The heuristic we derived by applying utility-weighted sampling to binary choices from description is similar to decision-by-sampling (N. Stewart et al., 2006) in that both mechanisms rely on drawing samples, comparing them, and counting how often the comparison favored the option to be evaluated. Although the decision-by-sampling model was originally proposed as a model of magnitude judgments, it has since been extended to predict choice probabilities (N. Stewart & Simpson, 2008; N. Stewart, 2009; N. Stewart, Reimers, & Harris, 2015; Noguchi & Stewart, 2016). Decision by sampling is consistent with the over-weighting of extreme events predicted by UWS because it assumes that the samples are drawn from memory and thus are subject to the availability bias in memory retrieval (N. Stewart et al., 2006; Tversky & Kahneman, 1973), but unlike UWS it does not specify why extreme events are more available than mundane events and how strong their availability should be.

Like UWS, the exemplar confusion model of decisions from description (Lin et al., 2015) assumes that people mentally simulate the outcomes of choosing either option. But unlike UWS, it simulates outcomes representatively according to their true stated probabilities and for each simulated outcome there is a small chance that its value will be confused. When a confusion occurs a value is chosen uniformly at random from the set of possible values and the sampled value replaces the value of the simulated outcome. This model captures that small probabilities tend to be overweighted whereas large probabilities tend to be underweighted in decisions from description. However, unlike UWS, the exemplar confusion model does not capture that overweighting depends on extremity.

UWS is similar to query theory (Johnson et al., 2007; Weber et al., 2007) in that both assume that preferences are constructed by the sequential consideration of a limited number of aspects or possible outcomes. Both accounts agree that cognitive constraints lead decision-makers to give more weight to the desiderata that are processed first. The main advance of UWS is to provide a rational process model of the order and frequency with which potential outcomes are queried and how the considered outcomes are translated into a decision. The similarity between UWS and query theory suggests that the process tracing methods that provided support for query theory could also be used to test UWS.

The strengths of UWS are complementary to early drift-diffusion models of value-based choice (Krajbich et al., 2010; Krajbich & Rangel, 2011). Both models sequentially accumulate evidence. But while the focus of UWS is on which outcomes should be generated to generate the most informative evidence, the focus of drift-diffusion models is on when the process of evidence generation should be terminated. Furthermore, while most applications of the drift-diffusion model focus on evidence that is generated by the environment, UWS focuses on evidence that is internally generated by memory recall or mental simulation. Recent work has applied to the drift-diffusion model to decisions from memory (Shadlen & Shohamy, 2016) to capture the relationship between response times and choice frequency. Our model is complementary in that it offers a quantitative account of which potential outcomes are sampled from memory. Combining UWS with the drift-diffusion model is one of the directions for future research we will discuss below.

Utility-weighted sampling is similar to selective integration (Tsetsos et al., 2016) in that both provide a rational explanation for violations of expected utility theory. However, the mechanism of utility-weighted sampling is different: In binary choice, utility-weighted sampling overweights attributes on which the alternatives differ by a large amount relative to attributes on which their values are similar. By contrast, selective integration always underweights the weaker attribute value by the same factor regardless of how much larger the stronger attribute value is. Furthermore, while the normative explanation of selective integration emphasizes noise in the decision stage, the normative justification of utility-weighted sampling is that most real-life decisions have to be made from a small subset of the available information because time is valuable. Our article complements the normative explanation of intransitivity by Tsetsos et al. (2016) by explaining a different set of cognitive biases that might result from a different mechanism.

Utility-weighted sampling is also related to the associative accumulation model by Bhatia (2013) according to which an attribute of a choice alternative will be sampled more frequently if its value is high. This is similar to our model except that in our model the sampling frequency would increases with the extremity of the attribute value's utility instead of its value per se. Utility-weighted sampling provides a strong rational explanation for the importance of extremity whereas the alternative assumption of the associative accumulation model appears to be less principled.

A critical feature of UWS is that it overweights the probability of extreme events. While prospect theory (Kahneman & Tversky, 1979) assumed that the overweighting of outcomes depends only on their probability, utility-weighted sampling predicted that over-weighting is driven by the outcome's utility, and the results reported here support this assumption very strongly. Rank-dependent expected utility theories (Quiggin, 1982) like cumulative prospect theory (Tversky & Kahneman, 1992) accommodate the effect of utility on probability weighting by applying the weighting function to the cumulative outcome distribution ($P(O \leq o)$). This captures that the probabilities of the worst and the best outcomes tend to be overweighted. Utility-weighted sampling adds to cumulative prospect theory by identifying cognitive mechanisms that might give rise to this effect. Furthermore, while Kahneman and Tversky (1979) assumed that the overweighting of outcome probabilities in decision-making was independent of the overestimation of event frequencies they attributed to the availability heuristic, we have argued that both originate from the same utility-weighted sampling mechanism.

In addition, utility weighted sampling predicts the distribution of people's choices whereas cumulative prospect theory was created to predict their modal response. This difference allowed utility-weighted sampling to capture people's choice frequencies in the Technion choice prediction competition more accurately than cumulative prospect theory. Our model performed on par with a stochastic extension of cumulative prospect theory that predicts choice distributions (Erev et al., 2010) and other probabilistic extensions of cumulative prospect theory (Rieskamp, 2008; Stott, 2006) might perform similarly. Furthermore, in cumulative prospect theory overweighting only depends on the rank of the outcome's utility. Thus, if the largest outcome is very close to all other outcomes, then it should be overweighted just as much as when it is orders of magnitudes larger than all other outcomes. By contrast, according to utility-weighted sampling, the largest outcome should be overweighted more heavily in the latter case than in the former.

Previous descriptive theories of choice, including disappointment theory (Bell, 1985; Loomes & Sugden, 1984, 1986), regret theory (Loomes & Sugden, 1982), and salience theory (Bordalo et al., 2012) also assert that people overweight extreme events. Our resource-rational analysis provides a rational justification for this assumption. Despite this commonality, UWS is qualitatively different from all of these previous theories. While all three previous theories are descriptive theories that predict what people will choose, utility-weighted sampling and utility-weighted learning are process models that specify the mechanism of how people decide and how this mechanism changes with learning. Unlike any of the previous theories, these mechanisms predict that people's memory recall and frequency estimates should be biased to overrepresent extreme events and both predictions were confirmed in the experiments reported above. We now discuss the similarities and differences between UWS and each of these three theories in turn.

Our UWS models of frequency estimation and decisions from experience bear a surprising similarity to disappointment theory (Bell, 1985; Loomes & Sugden, 1984, 1986) in that the optimal sampling distribution (Equation 10 in the Main Text) is proportional to the absolute value of the disappointment or elation that the decision maker would experience about the outcome, and according to the UWL model, the absolute value of the disappointment or elation that the decision-maker experiences determines how much the association between an action and its outcome is strengthened. Likewise, our model of binary choice from description is similar to regret theory and salience theory in that it amplifies the impact of large utility differences. Like regret theory and salience theory, this model assumes that decision-makers reason about the difference between the outcomes of the two actions instead of evaluating each action separately. Due to this commonality, our model of binary decisions from description shares some of the strengths and weaknesses of regret theory. On the positive side, this assumption allows all three theories to explain the Allais paradox, the fourfold pattern of risk preferences, and preference reversals. Furthermore, this shared property also predicts violations of weak stochastic transitivity (Tversky, 1969) for some triplets of gambles. For instance, with the parameters estimated from the Technion choice prediction tournament our UWS model of binary choice prefers a 50% chance of \$38 to a 35% chance of \$58 ($p_{2\succ 1} = 51.1\%$), prefers a 70% chance of \$30 over the 50% chance of \$38 ($p_{3\succ 2} = 51.7\%$), and yet prefers the 35% chance of \$58 over the 70% chance of \$30 ($p_{1\succ 3} = 52.1\%$). On the negative side, this commonality entails that, unlike disappointment theory, neither UWS nor regret theory can capture the common-ratio effects in problems that control for regret (Starmer & Sugden, 1989). Nor can UWS capture the specific intransitivity of people's preferences demonstrated by Tversky (1969).

Despite the commonality, the mechanism by which UWS overweights extreme events deviates from regret theory. Specifically, while regret theory and disappointment theory amplify the subjective utility of extreme events, UWS postulates that extremity increases the decision-maker's propensity to consider an outcome and thereby increases its subjective probability without affecting its utility. This entails a non-linear, non-monotonic interaction between probability and extremity: For unlikely outcomes the effect of extremity increases with their probability but for likely outcomes the effect of extremity decreases with their probability because subjective probabilities cannot be larger than 1. Furthermore, in UWS the overweighting of a large difference also depends on the magnitude of the utility differences between other pairs of outcomes. Depending on the magnitude of the differences of those other pairs, UWS can underweight the same pair of outcomes in one context and overweight it in a different context. By contrast, in regret theory, the same difference is always amplified to the same extent and its weight increases linearly with the event's probability. These differences did manifest in our simulations of the common-ratio effects reported by Starmer and Sugden (1989). Concretely, we found that UWS with the parameters estimated from the Technion choice prediction tournament for decisions from description failed to predict the common ratio effects that regret theory did capture (Starmer & Sugden, 1989). Furthermore, disappointment theory and regret theory make the very intuitive prediction that expectations and counterfactual outcomes modulate the satisfaction that people experience when they attain a certain outcome. This too, is not captured by UWS. On the other hand, UWS correctly predicted that extreme events come to mind first and that people overestimate their frequency. Taken together, these findings suggest that extremity affects both subjective utilities and subjective probabilities. UWS thus appears to be complementary to disappointment theory and regret theory because it captures different effects and explains them at a different level of analysis.

Although salience theory and UWS both assume that the subjective probability of extreme events is inflated, our account offers three advances over salience theory. First, we do not only describe the effect of utility on probability-weighting, but we also model the cognitive strategy that generates it. Second, our theory reconciles this seemingly irrational effect with rational information processing. Concretely, the resource-rational basis of the salience of a utility difference $\Delta U = u(O_1) - u(O_2)$ is the relative frequency with which it should be simulated, i.e. the importance distribution $\tilde{q}(\Delta u) \propto p(\Delta u) \cdot |\Delta u|$.[2] This provides a resource-rational justification for salience and a mechanistic account of its effect on decision-making. Third, since our explanation instantiates a more general theoretical framework – resource-rationality – it can also capture many additional phenomena such as decisions from experience, memory biases, and biases in frequency estimation.

Counterintuitive Model Prediction: Inconsistency increases with mental effort

Concretely, the gap between the UWS heuristic's risk seeking in choices between a small chance of winning a lottery versus the lottery's expected value (e.g., a 1% chance of winning \$100 vs. \$1 for sure) and its risk aversion for choices between a small chance of losing a gamble versus losing its expected value for sure (e.g., a 1% chance of losing \$100 vs. losing \$1 for sure) widens with the number of simulated outcomes due to the compounding of random errors in the evaluation of the utility of the simulated outcomes (see Figure 4). This entails that, in this very particular situation, manipulations that reduce mental effort, such as time pressure, should make people appear more rational in these decisions, whereas manipulations that increase mental effort should make them appear less rational.[3] This counter-intuitive relationship could also be used to test whether people allocate their cognitive resources rationally: While incentives for high performance should increase measures of mental effort (Mulder, 1986) on most tasks, people should always exert the minimal amount of cognitive effort on decisions problems where effort fails to improve performance. Regardless of how much mental effort a person exerts on these tasks they should always be biased at least as much as a person who simulates the outcome only once.

[2] This definition satisfies two of Bordalo et al.'s (2012) three axioms of salience.

[3] This prediction is very specific to the particular decisions described here, the normalized, stochastic utility function, and the estimated noise level but not representative of UWS in general.

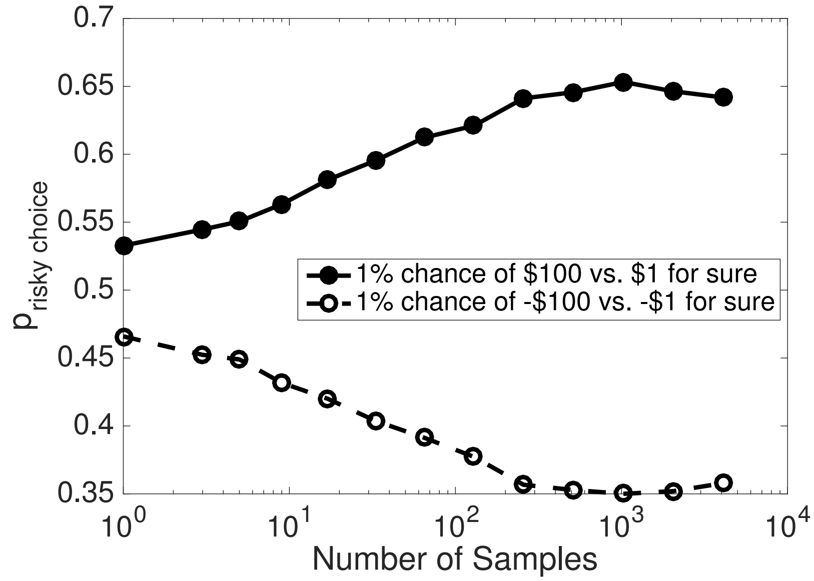

Figure 4. Counterintuitive prediction of UWS: Investing more mental effort can increase the inconsistency of people's risk preferences in choices between gambles and their expected values. Each line shows the frequency with which the UWS heuristic for binary decisions from description chose the risky option, averaged across 50000 simulations.

References

Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the american school (1952). In M. Allais & O. Hagen (Eds.), *Expected utility hypotheses and the allais paradox: Contemporary discussions of the decisions under uncertainty with allais' rejoinder* (pp. 27–145). Dordrecht: Springer Netherlands.

Anderson, J. R. (1990). *The adaptive character of thought.* Psychology Press.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological science*, *2*(6), 396–408.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations research*, *33*(1), 1–27.

Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological Review*, *120*(3), 522.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, *127*(3), 1243–1285.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, *113*(2), 409–432.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological review*, *69*(2), 109.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., ... Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making*, *23*(1), 15–47.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological review*, *118*(4), 523.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591–635.
Gonzalez, R., & Wu, G. (1999). On the Shape of the Probability Weighting Function. *Cognitive Psychology*, *38*(1), 129–166.
Hagen, O. (1979). Towards a positive theory of preferences under risk. In M. Allais & O. Hagen (Eds.), *Expected utility hypotheses and the allais paradox: Contemporary discussions of the decisions under uncertainty with allais' rejoinder* (pp. 271–302). Dordrecht: Springer Netherlands.
Hawkins, G. E., Camilleri, A. R., Heathcote, A., Newell, B. R., & Brown, S. D. (2014). Modeling probability knowledge and choice in decisions from experience. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 595–600). Austin, TX: Cognitive Science Society.
Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523.
Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 621.
Jarvstad, A., Hahn, U., Rushton, S. K., & Warren, P. A. (2013). Perceptuo-motor, cognitive, and description-based decision-making seem equally good. *Proceedings of the National Academy of Sciences*, *110*(40), 16271–16276.
Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: a query theory of value construction. *Journal of experimental psychology: Learning, memory, and cognition*, *33*(3), 461.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.
Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.
Kawaguchi, K., Kaelbling, L. P., & Lozano-Pérez, T. (2015). Bayesian optimization with exponential convergence. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 28* (pp. 2809–2817).
Kellen, D., Pachur, T., & Hertwig, R. (2016). How (in) variant are subjective representations of described and experienced risk and rewards? *Cognition*, *157*, 126–138.
Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, *13*(10), 1292–1298.
Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.
Lejarraga, T., Dutt, V., & Gonzalez, C. (2010). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making*, *25*(2), 143–153.
Lieder, F., & Griffiths, T. L. (2015). When to use which heuristic: A rational solution to the strategy selection problem. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th annual conference of the cognitive science society.* Austin, TX: Cognitive Science Society.
Lieder, F., & Griffiths, T. L. (under review). Strategy selection as rational metareasoning.
Lin, D., Donkin, C., & Newell, B. R. (2015). The exemplar confusion model: An account of biased probability estimates in decisions from description. In D. C. Noelle et al. (Eds.), *Proceedings of the 37th Annual Meeting of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.
Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*(368), 805–824.
Loomes, G., & Sugden, R. (1984). The importance of what might have been. In *Progress in utility and risk theory* (pp. 219–235). Springer.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282.
Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times: Memories for extreme outcomes bias risky decisions. *Psychonomic Bulletin & Review*, *21*(3), 629–636.
Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2016). The role of memory in distinguishing risky decisions from experience and description. *The Quarterly Journal of Experimental Psychology*, 1–12.
Marchiori, D., Di Guida, S., & Erev, I. (2015). Noisy retrieval models of over-and undersensitivity to rare events. *Decision*, *2*(2), 82.
Mulder, G. (1986). The concept and measurement of mental effort. In G. Hockey, A. W. Gaillard, & M. Coles (Eds.), *Energetics and human information processing* (pp. 175–198). Heidelberg: Springer.
Noguchi, T., & Stewart, N. (2016). Multialternative decision by sampling. (Manuscript submitted for publication)
Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied*, *18*(3), 314.
Plonsky, O., Teodorescu, K., & Erev, I. (2015). Reliance on small samples, the wavy recency effect, and similarity-based learning. *Psychological review*, *122*(4), 621.
Post, T., Van den Assem, M. J., Baltussen, G., & Thaler, R. H. (2008). Deal or no deal? decision making under risk in a large-payoff game show. *The American Economic Review*, *98*, 38–71.
Quiggin, J. (1982). A theory of anticipated utility. *J. Econ. Behav. Organ.*, *3*(4), 323–343.
Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(6), 1446.
Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927–939.
Shteingart, H., Neiman, T., & Loewenstein, Y. (2013). The role of first impression in operant learning. *Journal of Experimental Psychology: General*, *142*(2), 476.
Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665–690.
Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of economic literature*, *38*(2), 332–382.
Starmer, C., & Sugden, R. (1989). Probability and juxtaposition effects: An experimental investigation of the common ratio effect. *Journal of Risk and Uncertainty*, *2*(2), 159–178.
Stewart, N. (2009). Decision by sampling: the role of the decision environment in risky choice. *Quarterly Journal of Experimental Psychology*, *62*(6), 1041–1062.
Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.
Stewart, N., Reimers, S., & Harris, A. J. (2015). On the origin of utility, weighting, and discounting functions: How they get their shapes and how to change their shapes. *Management Science*, *61*, 687–705.
Stewart, N., & Simpson, K. (2008). A decision-by-sampling account of decision under risk. In N. Chater & M. Oaksford (Eds.), (pp. 261–276). Oxford University Press, Oxford, UK.
Stewart, T. C., West, R., & Lebiere, C. (2009). Applying cognitive architectures to decision making: How cognitive theory and the equivalence measure triumphed in the technion prediction tournament. In N. Taatgen, H. van Rijn, L. Schomaker, & J. Nerbonne (Eds.), *Proceedings of the 31st annual meeting of the cognitive science society* (pp. 561–566). Houston, TX: Cognitive Science Society.
Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and uncertainty*, *32*(2), 101–130.

Summerfield, C., & Tsetsos, K. (2015). Do humans make good decisions? *Trends in Cognitive Sciences*, *19*(1), 27–34.

Tierney, L., & Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, *81*(393), 82–86.

Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences*, *113*(11), 3102-3107.

Tversky, A. (1969). Intransitivity of preferences. *Psychological review*, *76*(1), 31.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323.

Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice a query-theory account. *Psychological science*, *18*(6), 516–523.