Research Article

Psychological Science
2019, Vol. 30(12) 1767–1779

DOI: 10.1177/0956797619884324
www.psychologicalscience.org/PS
SAGE

How to Change the Weight of Rare Events in Decisions From Experience

TC

Jared M. Hotaling[1], Andreas Jarvstad[2,3], Chris Donkin[1], and Ben R. Newell[1]

[1]School of Psychology, University of New South Wales; [2]Department of Psychology, University of Oxford; and [3]Department of Psychology, City, University of London

Abstract

When people make risky choices, two kinds of information are crucial: outcome values and outcome probabilities. Here, we demonstrate that the juncture at which value and probability information is provided has a fundamental effect on choice. Across four experiments involving 489 participants, we compared two decision-making scenarios: one in which value information was revealed during sampling (*standard*) and one in which value information was revealed after sampling (*value ignorance*). On average, participants made riskier choices when value information was provided after sampling. Moreover, parameter estimates from a hierarchical Bayesian implementation of cumulative-prospect theory suggested that participants overweighted rare events when value information was absent during sampling but did not overweight such events in the standard condition. This suggests that the impact of rare events on choice relies crucially on the timing of probability and value integration. We provide paths toward mechanistic explanations of our results based on frameworks that assume different underlying cognitive architectures.

Keywords

decisions from experience, description–experience gap, decision making, sampling, cumulative-prospect theory, hierarchical Bayesian modeling, open data, open materials, preregistered

Received 10/31/18; Revision accepted 9/14/19

When people make risky choices, two kinds of information are crucial: outcome values and outcome probabilities (Bernoulli, 1954; von Neumann & Morgenstern, 1947). Although some decisions involve choosing between options with described values and probabilities (e.g., investing in stocks on the basis of earnings projections), others require learning about options through experience (e.g., finding a favorite restaurant in a new city). There is considerable evidence that these types of decisions differ systematically, a phenomenon known as the *description–experience gap* (see Wulff, Mergenthaler-Canseco, & Hertwig, 2018, for a recent review).

Decisions from description typically result in choices that imply overweighting of rare events (Kahneman & Tversky, 1979; Rieskamp, 2008), whereas decisions from experience do not (Camilleri & Newell, 2011; Hertwig, Barron, Weber, & Erev, 2004; Lejarraga & Gonzalez, 2011; Rakow & Newell, 2010; Yechiam & Busemeyer, 2006). Much has been made over the past decade of this key difference in behavior (Hertwig, Hogarth, & Lejarraga, 2018), yet its causes remain elusive (e.g., Glöckner, Hilbig, Henninger, & Fiedler, 2016; Wulff et al., 2018). Part of the problem may stem from focusing on the gap—which is not, in itself, behavior to be explained—instead of attempting to predict and explain the behaviors that exist along the decision continuum between description and experience (e.g., Camilleri & Newell, 2013; Rakow & Newell, 2010). Here, we focus on an aspect of information provision in risky choice that has received little attention in the literature.

We propose that an important aspect of the decision-making process lies in the juncture at which outcome values and probabilities are integrated (Jarvstad, Hahn, Rushton, & Warren, 2013, p. 16275). We explored this

Corresponding Author:
Jared M. Hotaling, School of Psychology, University of New South Wales, Sydney 2052, Australia
E-mail: jaredhotaling@gmail.com

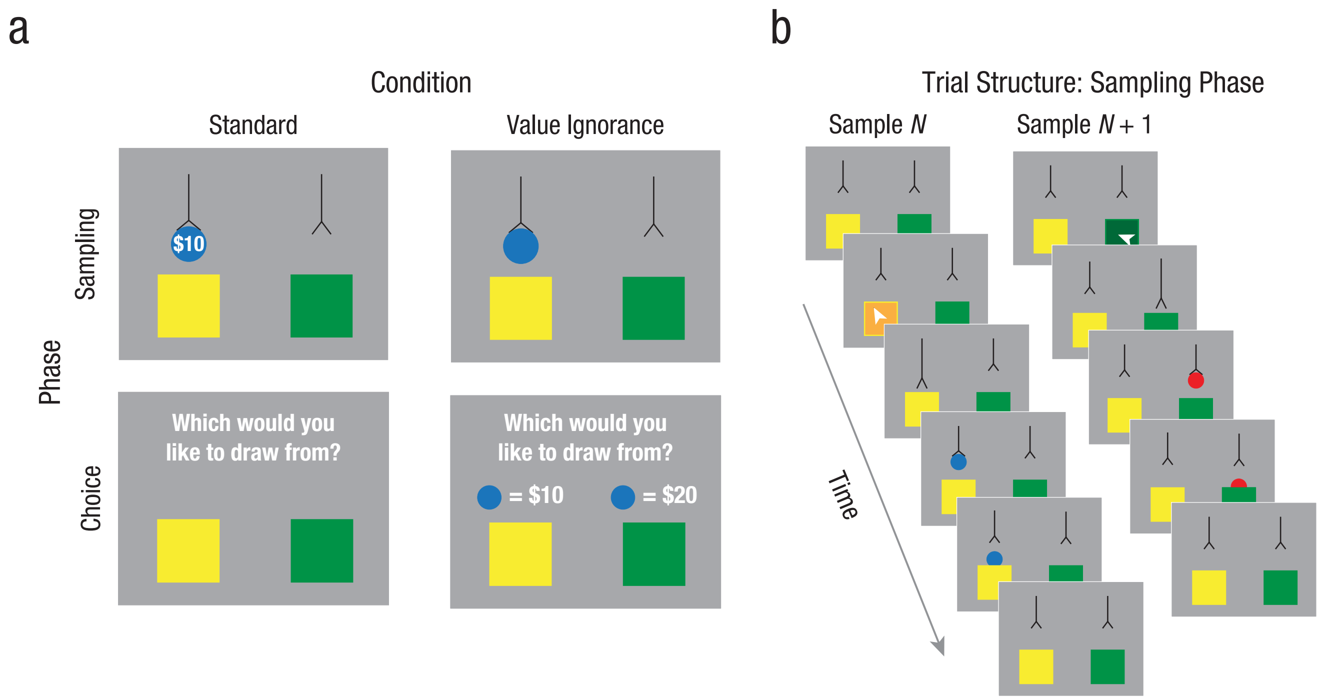

Fig. 1. Design and trial structure of the robot-sampling task. Each trial (a) was composed of a sampling phase and a choice phase. For a given sample, participants could observe either a blue ball or a red ball. Red balls were worth $0, and blue balls were worth some reward. In the *standard* condition, the value of the blue ball was revealed during sampling. In the *value-ignorance* condition, the value of the blue ball was not revealed until the choice stage. Thus, during sampling in the value-ignorance condition, the probability of drawing a blue ball could be learned, but its value could not. In the sampling sequence (b), once a box was clicked (i.e., selected for sampling), an animation showed the box shaking (to mix the balls). Then a robot arm reached down and grabbed a ball, lifted it up to reveal it, and then dropped it back down again (illustrating sampling with replacement). Participants were required to sample each box a set number of times but were free to sample the boxes in any order.

integration by comparing choices under two sampling-based, decision-from-experience conditions (Fig. 1). Participants directed a robot to sample balls from boxes. In the *standard* condition (Fig. 1a, left), the reward magnitudes were presented at the time of sampling, as in standard decision-from-experience tasks (e.g., Hertwig et al., 2004). In the *value-ignorance* condition (Fig. 1a, right), participants learned about outcome probabilities during the sampling phase but not about reward magnitude, which was revealed at the choice stage. The participants' goal was to learn which of two boxes they would prefer to draw from "for real" at the choice stage. Crucially, participants in both conditions learned about identical outcome distributions with deterministic reward magnitudes—the sole difference was when value information was available (cf. Hadar & Fox, 2009; see also the Discussion section).

In the value-ignorance condition, participants could, by sampling each box, gradually learn the probability of drawing a blue ball (Hogarth & Einhorn, 1992; Sutton & Barto, 1998). At the moment of choice, they might recall that probability and integrate it with value information. Thus, participants might represent two pieces of information for each box: the probability of drawing a rewarded ball and its reward. An identical learning strategy could also be applied in the standard condition. However, unlike in the value-ignorance condition, a simpler approach was possible. Participants might represent only the average outcome of each box. This value-updating strategy did not require that participants separately learn the probability of drawing a blue ball, and similar mechanisms have been proposed to account for decisions from experience (Hertwig, Barron, Weber, & Erev, 2006).

If people used a single mechanism in both the standard and value-ignorance conditions, we would expect to observe similar choices across conditions. However, if manipulating the juncture at which participants received value information led them to use different mechanisms, we would expect to observe divergent choices across conditions. As is typical, we first examined the proportion of choices in favor of the riskier alternative in each condition. However, we wanted to understand the aspects of the decision mechanisms that produce different risk preferences, and so we used cumulative-prospect theory (CPT; Tversky & Kahneman, 1992) as a measurement model to estimate the effect of our manipulations on participants' treatment of probability and value information. This model has been used to measure risk preferences in a number of domains (e.g., Glöckner et al., 2016).

Across four experiments, we found that behavior at the choice stage differed systematically across conditions. Specifically, participants made a higher proportion of riskier choices in the value-ignorance condition than in the standard condition. We unpacked this result with a hierarchical Bayesian implementation of CPT, which showed that participants overweighted small probabilities in the value-ignorance condition but used neutral weighting in the standard condition. Moreover, the observed neutral weighting of outcomes in the standard condition was disrupted when rarer events were perceptually highlighted during sampling (Experiments 3 and 4), with this manipulation causing these events to be overweighted, as in the value-ignorance condition.[1] We argue that these results provide clear evidence for the operation of different decision mechanisms in the standard and value-ignorance conditions. In the Discussion section, we outline contrasting mechanistic explanations based on frameworks that assume different cognitive architectures.

Method

All data exclusions are reported and justified in this section. All independent variables and manipulations, whether successful or failed, and all dependent measures that were analyzed for this article's target research question are reported.

Participants

Ethical approval for all experiments was obtained through the institutional review boards of the School of Psychology at the University of New South Wales (UNSW). All participants were UNSW students and received course credit plus a monetary bonus ($0– $21; monetary amounts are in Australian dollars) based on a randomly selected trial. Eighty students[2] participated in Experiment 1, 83 students (45 female; age: range = 17–42 years, M = 19.44, SD = 2.95) participated in Experiment 2, 149 students (99 female; age: range = 18–53 years, M = 22.93, SD = 4.63) participated in Experiment 3, and 177 students (106 female; age: range = 18–58 years, M = 20.49, SD = 3.92) participated in Experiment 4. Within each experiment, participants were randomly distributed across conditions.

Procedure

After giving informed consent, participants were placed in a computer booth, where they read the following instructions:

> In this task you will draw balls from pairs of virtual boxes. In each box, there are 100 balls, some of which are blue and some of which are red. Blue balls are associated with reward and red balls are not (reward for a red ball = $0).

Participants first completed a practice trial to familiarize themselves with the task (Fig. 1) before beginning the experimental trials. Participants were told that each trial involved a new pair of boxes and that they would have to learn anew the values and proportions of balls within each box. To emphasize that boxes were different across trials, we gave each box a unique color. On each trial, participants were required to sample the entire set for both alternatives before making their choice. Participants were able to sample freely (e.g., alternating between boxes, sampling exhaustively from one and then the other), and sampling was disabled once the entire set from each box had been sampled. Participants were instructed that one of their choices would be used to draw a ball for a bonus payment at the end of the experiment.[3]

Importantly, the samples that participants observed matched the true underlying probabilities of each outcome, thus mitigating other factors that might have given rise to illusory gaps (e.g., biased sampling and reliance on small samples; Hau, Pleskac, Kiefer, & Hertwig, 2008; Hertwig & Pleskac, 2010; Rakow, Demes, & Newell, 2008).

Materials and design

Experiments 1 and 2. The values of boxes (monetary gambles) were determined as follows. Each choice alternative was defined by a reward value, ν (range = $1–$21), and a probability of reward, π (range = .083–1). With these values, we created a sample set for each alternative representing the proportion of red and blue balls. The size of the sample set ranged from 10 to 12, and the frequency of rewards was determined by π.

Red balls were always worth $0. The value of blue balls was fixed within each box but varied across boxes and trials from $1 to $21. For example, the value of a blue ball might be $16 in the left box and $2 in the right box. In the standard condition (Fig. 1a), each sampled ball was labeled with the outcome value. In the value-ignorance condition, sampled balls were not labeled with values, though the instructions indicated that red balls were worth $0 and that the blue balls were worth some reward. Participants could therefore learn the relative proportions of balls in each box but not their values; the value was revealed in the choice phase (see Fig. 1a).

Choice pairs were constructed with the goal of exposing participants to a range of problems. For example, problems could involve zero, one, or two risky options (i.e., $\pi < .5$), and equal or unequal expected values. To better understand the task, consider an example trial involving a riskier option on the left, offering a 10% chance of winning \$16, and a safer option on the right, offering an 80% chance of winning \$2. When sampling from the riskier box, participants would observe one blue ball and nine red balls; from the safer box, they would sample eight blue balls and two red balls.

In Experiment 1, each participant received the same underlying decision problems, but in Experiment 2, five of the decision problems were the same across participants, and six of the problems were randomly generated. Experiment 2 also involved minor graphical changes, such as a larger font size. In all other aspects, Experiments 1 and 2 were identical (see the Supplemental Material available online for decision-problem details).

Experiments 3 and 4. In Experiment 3, we introduced a new manipulation motivated by our finding of the differential influence of rare events in Experiments 1 and 2 (see the Results section). Specifically, we sought to examine whether the lower weighting placed on rare events in the standard condition, compared with the value-ignorance condition, could be increased by emphasizing the rare outcome to participants. To do this, we introduced a *salience* condition in which, during sampling, some balls were highlighted. When a highlighted ball was drawn, a tone played, and the ball flashed for approximately 700 ms before being returned to the box as usual.

The highlighting occurred whenever participants sampled the rare event for the riskier alternative. This resulted in two types of problems. For Type 1 problems (*best outcome salient*), salience highlighted a rare reward and was expected to increase the likelihood of choosing the risky option. For Type 2 problems (*worst outcome salient*), salience highlighted an outcome of \$0 and could be expected to decrease the likelihood of choosing the risky option. The *no-salience* conditions replicated the conditions in Experiments 1 and 2, and no highlighting occurred. Each participant completed the same set of fourteen Type 1 trials and six Type 2 trials in a random order. Experiment 4 was a preregistered replication of Experiment 3 (details can be found on the Open Science Framework at https://osf.io/zsbwt/). See the Supplemental Material for the specific gambles used in Experiments 3 and 4.

Statistical analyses

Because each experiment used nearly identical methods, procedures, and designs; because we expected small effect sizes given our well-controlled design (see Camilleri & Newell, 2011; Glöckner et al., 2016; Wulff et al., 2018); and because we wanted to make use of the large amount of data collected across experiments—both in terms of number of participants and different choice options—we report and analyze them together. In designing each experiment, we selected sample sizes such that pooling across experiments with hierarchical Bayesian analyses would yield a total sample size sufficient to demonstrate reliable evidence for the effects we observed and one considerably larger than in previous decision-from-experience studies (see Wulff et al., 2018). We pooled in two different ways: (a) across Experiments 1 through 4 for standard versus value-ignorance contrasts and (b) across Experiments 3 and 4 to examine the effect of the salience manipulation (absent in Experiments 1 and 2).

Type 1 gambles (best outcome salient) and Type 2 gambles (worst outcome salient). In pooling these data, we treated Type 2 gambles differently. When considering the effect of value information, we used all data from Experiments 1 through 4 because we expected the effect of value information to be the same for both Type 1 and Type 2 gambles. However, our salience manipulation carried the risk of introducing a demand characteristic whereby participants were encouraged to choose the riskier option, regardless of which outcome was highlighted. Type 2 problems therefore served as a manipulation check because salience highlighted an outcome of \$0 rather than a rare reward. Because we expected the effect of salience to be different for Type 1 and Type 2 gambles, we analyzed only Type 1 problems when considering the effect of salience. In the Supplemental Material, we show that the salience manipulation had no effect on choices for Type 2 problems.

Proportion of risky choices. We modeled the proportion of risky choices made by participant i in condition j as coming from a binomial process with rate parameter θ_{ij}. Our data across the four experiments can be thought of as contributing different amounts of information to four conditions in a 2 (sampling) × 2 (salience) design. Because there was no salience manipulation in Experiments 1 and 2, these data contributed only to the no-salience conditions. Experiments 3 and 4, on the other hand, contributed data to all four conditions. Finally, we assumed that the rate parameters for each participant came from a group-level normal distribution, $\theta_{ij} \sim N(\mu^{\theta_j}, \sigma^{\theta})$, truncated to be between 0 and 1, with a group-level mean, μ^{θ_j}, for each condition and a variability parameter, σ^{θ}, that was shared across conditions for parsimony. We used uniform (0, 1) priors for each group-level mean, μ^{θ_j}, and the prior for group-level precision was set so that $1/\sigma^2_{\theta_j} \sim \gamma(.001, .001)$.[4]

We also examined the difference between the (population-level) proportion of risky choices in the standard and value-ignorance conditions. We calculated the posterior distribution of the differences between μ^{θ} in the two sampling conditions, $\Delta\mu^{\theta}$. To calculate a Bayesian equivalent of a frequentist *p* value, we evaluated the empirical cumulative density function of $\Delta\mu^{\theta}$ values at 0. This Bayesian *p* value tells us how likely it is that the difference between μ^{θ} in the two sampling conditions is below zero. Extremely small or large *p* values were thus associated with a high probability that participants behaved differently in the standard and value-ignorance conditions.

Model-based analysis of risk preferences. We used a hierarchical Bayesian implementation of prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) to estimate participants' risk preferences. This analysis allowed us to interpret our results in terms of psychologically interesting latent parameters while naturally accounting for the variability observed across gambles, which is ignored in many traditional analyses. One can think of this as extending our pooled data analysis to include a psychological model of the θ parameters used earlier, accounting for participants' risk preferences with four latent parameters.

Prospect theory was parameterized following Stott's (2006) methodology: A power utility function, $v(x) = x^{\alpha}$, captured preferences for money, and a one-parameter probability weighting function, $w(p) = e^{(-(-\ln p)^{\gamma})}$ (Prelec, 1998), captured preferences for probabilities. Choice options were modeled in terms of whether they appeared on the left or right of the screen, with the differences in prospect values between the left and right options given by $\Delta_{\text{prospect}} = v(x)_{\text{left}} w(p)_{\text{left}} - v(x)_{\text{right}} w(p)_{\text{right}}$. We mapped this difference onto the probability of choosing the left option using a two-parameter logistic function, $p_{\text{left}} = \dfrac{1}{1 + e^{-1\,\kappa\,[\beta + \Delta_{\text{prospect}}]}}$. Higher values of the sensitivity parameter, κ, indicate better discrimination between prospects (i.e., more deterministic choice). The bias parameter, β, captured the extent to which one side of the screen was favored irrespective of prospect value.

In a first analysis, we fitted the model to the standard and value-ignorance conditions, pooling data across Experiments 1 through 4. Participants' risk preferences, described by the four parameters in the model, were modeled as coming from four separate population-level distributions: $\alpha_i \sim e^{N(\mu^{\alpha},\, \sigma^{\alpha})}$, $\gamma_i \sim e^{N(\mu^{\gamma},\, \sigma^{\gamma})}$, $\kappa_i \sim e^{N(\mu^{\kappa},\, \sigma^{\kappa})}$, $\beta_i \sim N(\mu^{\beta}, \sigma^{\beta})$. For α, γ, and κ, participant-level samples were drawn from exponentiated normal distributions. Priors for the population-level means were $\mu^{\alpha} \sim U(-10,10)$, $\mu^{\gamma} \sim U(-10,10)$, $\mu^{\kappa} \sim U(-2,10)$, $\mu^{\beta} \sim U(-10,10)$, which are sufficiently broad to capture extreme distortions of money and probability and a wide range of logistic functions. Population-level parameters in log space accounted for the fact that α and γ have qualitatively different meanings between 0 and 1 and between 1 and $+\infty$. The priors for α and γ were symmetric around 0 to assign equal prior probability to under- and overweighting of small probabilities and increasing or diminishing marginal utility, respectively. The prior for k was not symmetric, as discriminability increases monotonically with k. The priors for all population-level standard deviations were identical across parameters—for example, $\sigma^{\alpha} \sim U(1 \times 10^{-5}, 6)$. Sampling was truncated by the respective prior range. For plotting purposes, we report best-fitting population-level parameters in standard parameter space.

To evaluate the effect of the salience manipulation on risk preferences, we pooled data from Experiments 3 and 4 and estimated population level α and γ for each of the four conditions (standard and salience, standard and no salience, value ignorance and salience, value ignorance and no salience). On the basis of the standard versus value-ignorance analysis (for which we observed no condition-specific effects on κ and β; see the Results section), we modeled these parameters as coming from the same population-level distributions. In all other aspects, this model was identical to the model above.

Results

The left panel of Figure 2 shows our analysis of choice-proportion data pooled across Experiments 1 to 4. As the figure shows, participants who did not know the values associated with each outcome during sampling (value-ignorance condition) made more risky choices than participants for whom the value information was present during sampling. The right panel of Figure 2 plots the posterior distribution of the difference in the proportion of risky choices between the standard and value-ignorance conditions. Here, we see that the effect of value information on risky choice was reliable, with the posterior distribution of the difference in the proportion of risky choices between the standard and value-ignorance conditions sitting almost entirely above zero ($p = .009$).

Figure 3 pools data from Experiments 3 and 4 to show the effects of value information and salience. In the top panel, we see that in both the no-salience and salience conditions, participants made more risky choices if they did not know the values associated with outcomes during sampling. The bottom panel of Figure 3 shows that in the no-salience condition, the effect of value information on risky choice was relatively large and reliable when there was no salience manipulation: The posterior distribution of the difference in the proportion of risky choices between the standard and value-ignorance conditions again was almost entirely

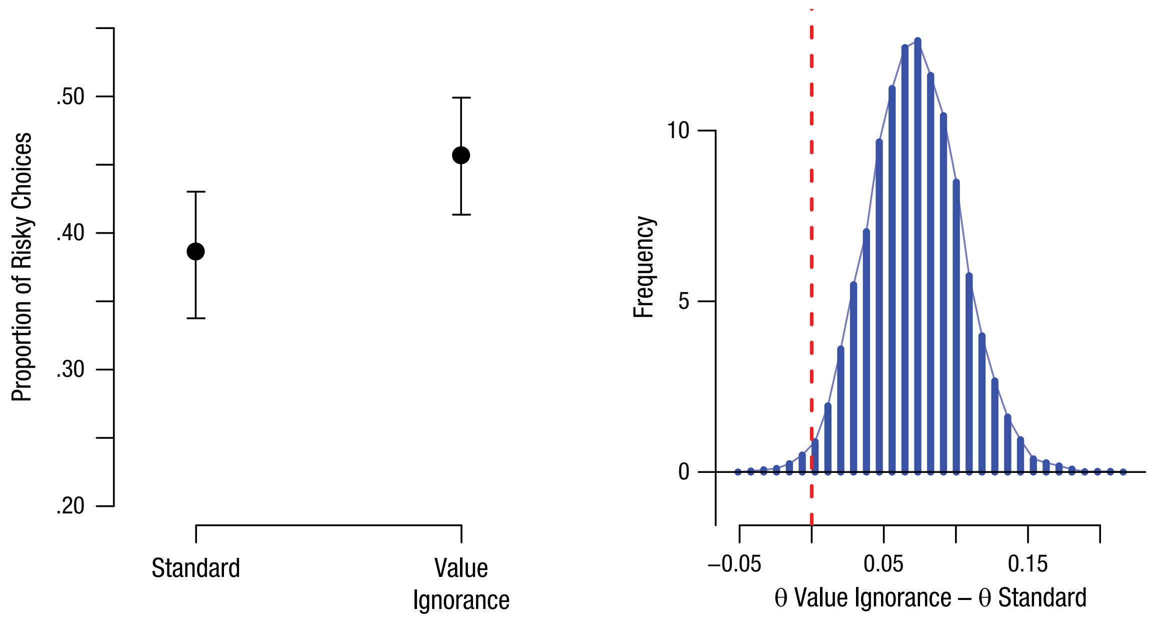

Fig. 2. Choice-proportion data pooled across Experiments 1 to 4. The left panel shows the median and central 95% of the posterior distributions for the population-level mean of the proportion of risky choices, θ, in the standard and value-ignorance conditions. The right panel shows the posterior distribution of the difference between population-level θ parameters in the standard and value-ignorance conditions. Positive values on the *x*-axis reflect more risky choices made when value information was absent during sampling (value-ignorance condition). The dashed vertical line indicates no difference between conditions.

above zero (p = .004). When the salience manipulation was present, the difference between the standard and value-ignorance conditions was less robust because the salience manipulation increased risky choices in the standard conditions. There was, however, some evidence of its reliability: In the posterior distribution of the difference in the proportion of risky choices between the standard and value-ignorance conditions, most of the density was above zero (p = .08).

The previous analysis shows that participants made riskier choices when they had to wait until the choice phase to integrate probability and value, compared with when they could integrate this information at the sampling phase. To see whether this pattern held at the level of individual gambles, we explored gamble-wise choices from Experiments 3 and 4 using a wide range of reward probabilities and substantial data for each gamble. Results are shown in Figure 4, which plots the proportion of risky choices as a function of the probability of drawing a blue ball from the riskier box (i.e., the box with the lower probability of reward). As can be seen, when the riskier option was unlikely (p < .5), participants in the value-ignorance condition made riskier choices than participants in the standard condition. This result matches the overall effect on risky choice observed previously. However, for gambles in which the riskier option was likely (p > .5), we found either no appreciable difference between conditions or a trend toward less risky choices in the value-ignorance condition. This is the pattern one would expect if the value-ignorance condition induced prospect-theory-like weighting of probability at the time of choice (overweighting of small probabilities, underweighting of large probabilities).

The previous analyses showed that having to integrate probability and value at the time of choice induced riskier choices relative to the standard condition and that this average effect was mainly due to a strong preference for choosing the risky option when its probability was low (Fig. 4). What these previous analyses did not show is which of the two critical pieces of information was affected—value or probability. To quantify this, we used prospect theory as a measurement model and estimated its parameters using hierarchical Bayesian methods.

Figure 5 shows the result of fitting prospect theory to the data, with the mean posterior utility and probability-weighting functions (Figs. 5a and 5b, respectively) for the standard and value-ignorance conditions. Participants in both conditions (standard: μ^{α} = .71, value ignorance: μ^{α} = .59) showed decreasing sensitivity to increasing monetary values (diminishing marginal utility), as expected (Fig. 5a). Across the two conditions, there was a small difference, with participants in the standard condition showing more linear preferences than those in the value-ignorance condition (p = .032; Fig. 5e).

In terms of probability weighting (Fig. 5b), the standard condition resulted in near-linear weighting of small

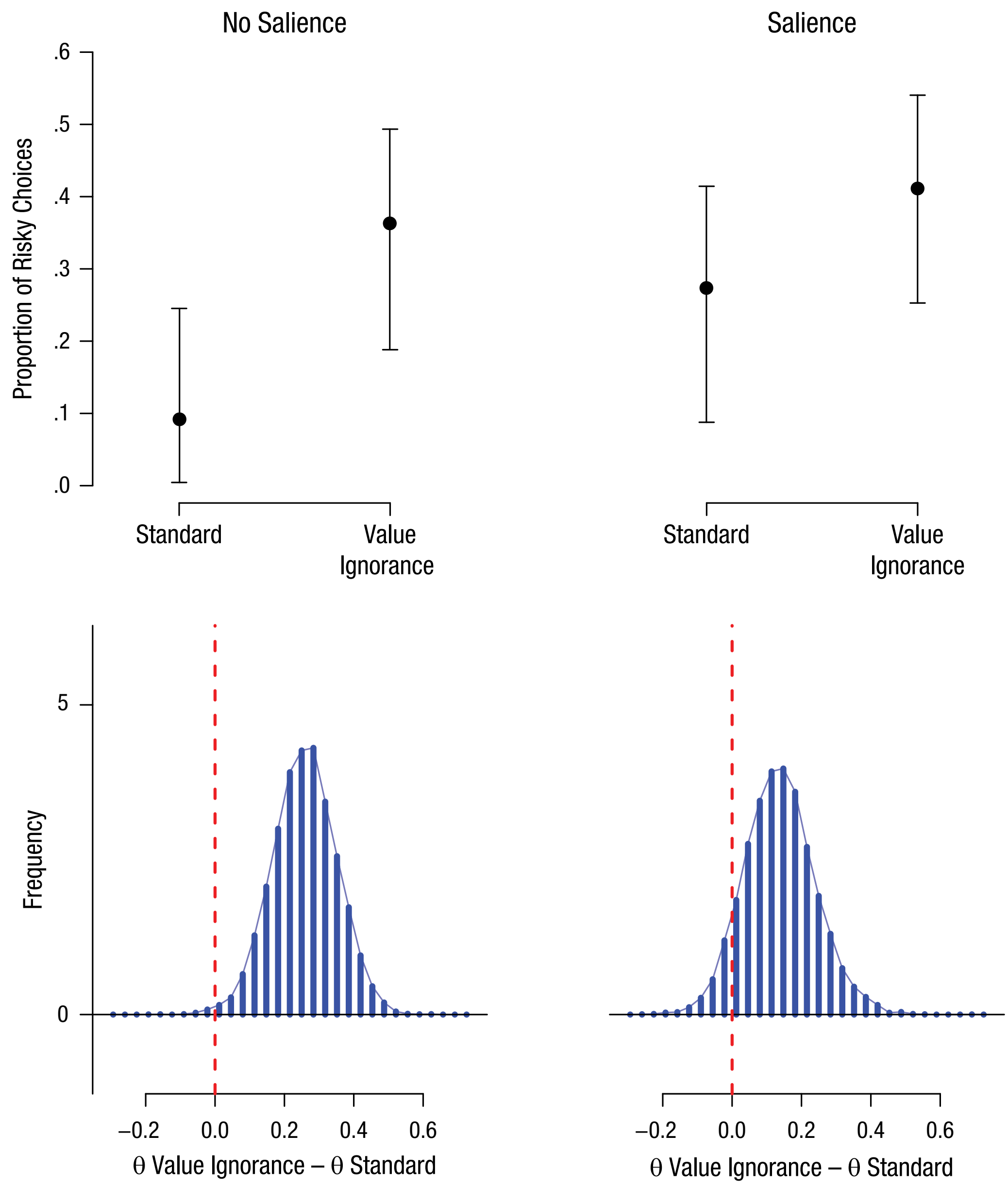

Fig. 3. Effects of value information and salience in Experiments 3 and 4. The top row shows the median and central 95% of the posterior distributions for the population-level mean of the proportion of risky choices, θ, in the standard and value-ignorance conditions, separately for the no-salience and salience conditions. The bottom row shows the posterior distribution of the difference between population-level θ parameters in the standard and value-ignorance conditions for the no-salience and salience conditions. Positive values on the *x*-axis reflect more risky choices when value information was absent during sampling (value-ignorance condition). The dashed vertical line indicates no difference between conditions. *Salience* refers to the perceptual highlighting of a drawn ball during sampling.

probabilities ($\mu^{\gamma} = 1.02$), and the value-ignorance condition resulted in overweighting of small probabilities ($\mu^{\gamma} = 0.69$). This between-conditions difference in probability weighting was highly reliable ($p = 1 \times 10^{-4}$; Fig. 5f). This pattern corresponds well to the standard description–experience gap in that we observed near-neutral weighting for the standard experience-based condition (i.e., experience-like weighting; e.g., Camilleri & Newell, 2011) and overweighting of small probabilities when value information was absent at the time of sampling (i.e., description-like weighting; e.g., Kahneman & Tversky, 1979). There was very little evidence for differences between conditions in the parameters of the logistic choice function (Figs. 5c, 5d, 5g, and 5h).

In Experiments 3 and 4, we also manipulated the salience of the riskier option. We found that perceptually

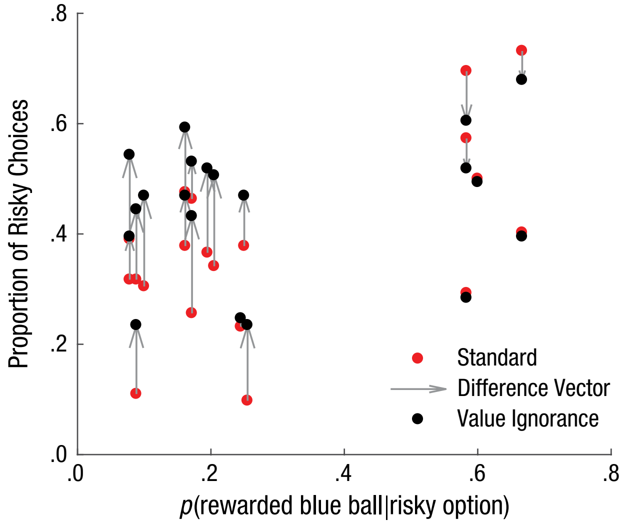

Fig. 4. Gamble-wise proportion of risky choices as a function of the probability of drawing a blue ball from the riskier option in Experiments 3 and 4. Each data point comes from a unique gamble in one of the two conditions. In both experiments, there was a wide range of reward probabilities and approximately 80 observations per gamble. The arrows indicate the direction of change in the value-ignorance condition relative to the standard condition. Values on the *x*-axis have been jittered to improve readability.

highlighting rare events during sampling increased the proportion of risky choices in the standard condition but had little effect on sampling in the value-ignorance condition (Fig. 3). The corresponding prospect-theory analyses are shown in Figure 6.[5]

Figures 6a and 6c show the mean posterior functions for the standard condition, and Figures 6b and 6d show the mean posterior functions for the value-ignorance condition. The best-fitting utility functions (top row) show that there were trends toward a lower sensitivity to value (more nonlinear utility functions) in the salience conditions (standard: $\mu^{\alpha} = .58$, value ignorance: $\mu^{\alpha} = .61$) compared with the no-salience conditions (standard: $\mu^{\alpha} = .75$, value ignorance: $\mu^{\alpha} = .73$; standard: $p = .057$, value ignorance: $p = .196$, respectively).

For probability weighting, the salience manipulation affected the standard condition differently from the value-ignorance condition. For the latter condition, the manipulation had essentially no effect. In other words, when outcome values were absent during sampling, highlighting a rare event did not induce a change in how learned probabilities were treated at the time of choice (no salience: $\mu^{\gamma} = 0.84$, salience $\mu^{\gamma} = 0.78$, $p = .333$; Fig. 6d). For the standard condition, however, highlighting the rewards for the riskier option had a dramatic impact on how the gambles were evaluated, resulting in an overweighting of low-probability outcomes (no salience: $\mu^{\gamma} = 1.06$, salience $\mu^{\gamma} = 0.75$, $p = .005$; Fig. 6c).

The population-level estimates of κ (sensitivity to differences between prospects) were comparable with those from the model used to contrast the standard and the value-ignorance conditions ($\mu^{\kappa} = 2.2$, $\sigma^{\kappa} = 1.08$), and there was a similar trend toward a left-option bias ($\mu^{\beta} = 0.015$, $\sigma^{\beta} = 0.013$).

General Discussion

Our findings show that subtle changes to an experience-based decision task can substantially affect both choice behavior and risk preferences. By manipulating whether outcome values were present while participants learned the probabilities of receiving a reward, we produced effects reminiscent of the much-publicized description–experience gap. Across four experiments, we observed riskier choices, on average, when outcome values were absent during sampling, compared with when they were present. This average increased preference for risk was driven by situations in which there was a small probability of the best outcome (see Fig. 4). These patterns suggest that having to integrate value and probability information at the time of choice leads people to treat rare events differently.

The logic behind this conclusion is based on the observation that the two versions of our task can be solved using different methods. In the value-ignorance condition, participants must somehow integrate probability information—which they learned during sampling—with value information presented on the choice screen. On the other hand, participants in the standard condition can avoid this direct integration by learning the expected reward produced by each option during sampling. The behavioral differences we observed across conditions indicate that our manipulation did affect participants' decision processes, though this occurred in subtle ways, depending on the exact combinations of probabilities and rewards (see Fig. 4). To better examine this interpretation, we used a hierarchical Bayesian CPT model to analyze the risk preferences underlying participants' choices. Here, too, we found compelling evidence for a difference between conditions, most notably in probability weighting.[6] Participants in the value-ignorance condition overweighted rare events, while those in the standard condition did not. These results emphasize the importance of mapping and measuring behavior on the continuum between decisions from description and decisions from experience (Camilleri & Newell, 2013; de Palma et al., 2014; Rakow & Newell, 2010).

Our results raise important questions regarding psychological mechanisms. Why does the weighting of rare events increase when outcome values are presented only at the time when a choice has to be made? How

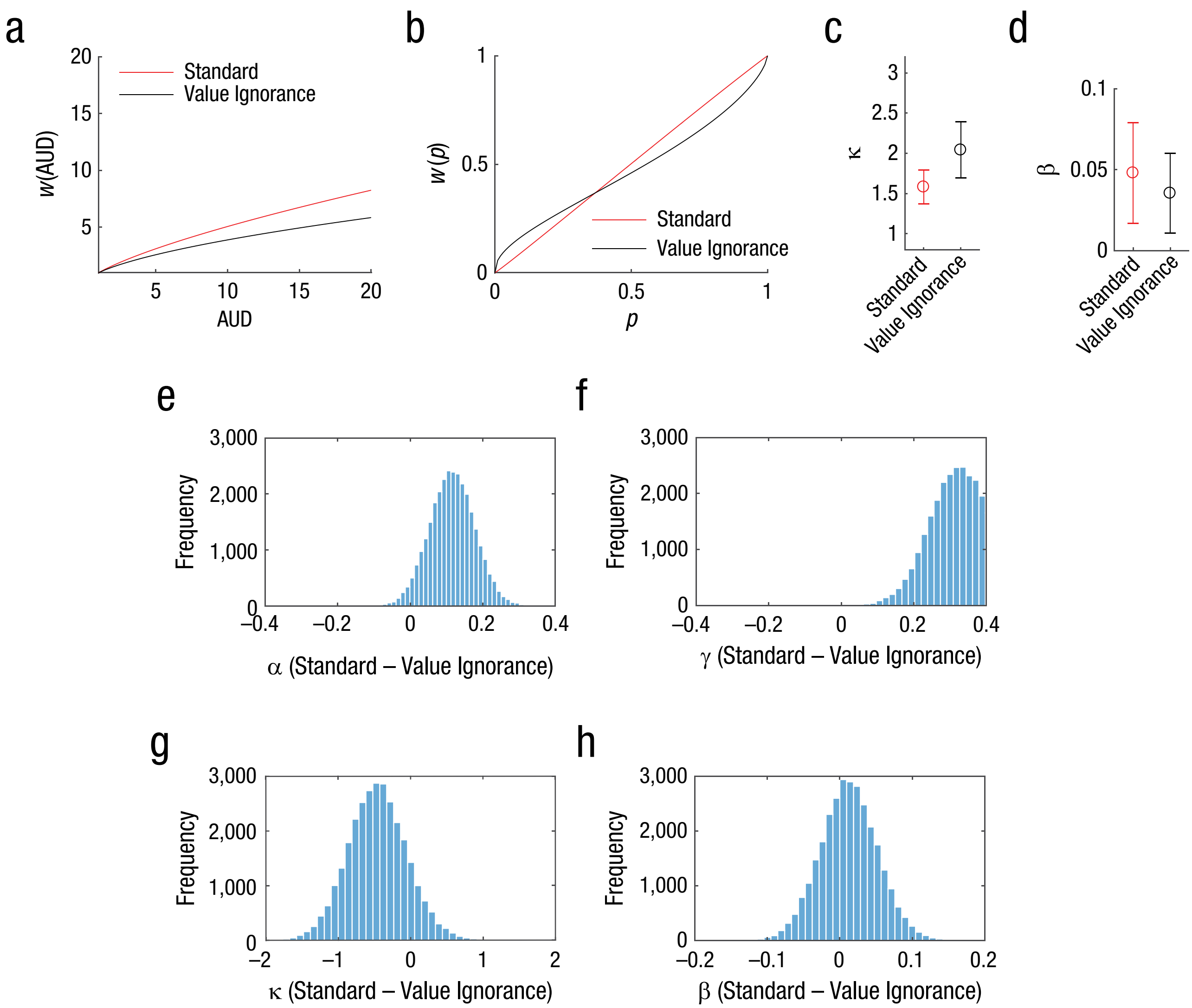

Fig. 5. Risk preferences for the standard and value-ignorance conditions in Experiments 1 through 4. The leftmost column shows (a) mean posterior utility functions, mapping money onto subjective utility, and (b) mean posterior probability weighting functions, mapping probability onto subjective decision weights. The next column illustrates (c) the posterior mean of the sensitivity parameter of the logistic function and (d) the posterior mean of the side bias (positive = left-side bias). Error bars in (c) and (d) show standard deviations. The histograms in the right-hand columns show posterior distributions of differences in population-level parameters between the standard and the value-ignorance conditions for the (e) α, (f) γ, (g) κ, and (h) β parameters. Values on the *x*-axis were scaled to facilitate comparison of effect sizes; for *x*-axis scaling to fit the distributions, see Figure S1 in the Supplemental Material. AUD = Australian dollars.

do manipulations of outcome salience influence choice and risk preference? Several possibilities exist, but we focus on two potential frameworks of explanation.

Reinforcement learning

Our results could be cast in the form of reinforcement learning. Imagine a reinforcement-learning agent making choices in the standard condition. After sampling a ball from Box A, the value for Box A is updated according to the formula $v_A^{t+1} = v_A^t + \delta\,(o_A^t - v_A^t)$, where v_A^t is the estimated value at sample t, o_A^t is the value of the sampled ball, and δ is a learning-rate parameter between 0 and 1. After completing sampling, the agent compares the estimated value of each option and chooses accordingly. Contrast this mechanism with that of another agent in the value-ignorance condition. This agent learns the probability of receiving a reward from Box A according to $p_A^{t+1} = p_A^t + \delta\,(c_A^t - p_A^t)$, where p_A^t is the estimated probability of a reward at sample t, and c_A^t is the outcome of the sample (1 = reward, 0 = no reward). On the choice screen, the value of a reward from Box A is revealed, and the reinforcement-learning agent integrates this value with the learned probability and would thus—unlike the box-averaging agent—explicitly represent both value and probability. These different

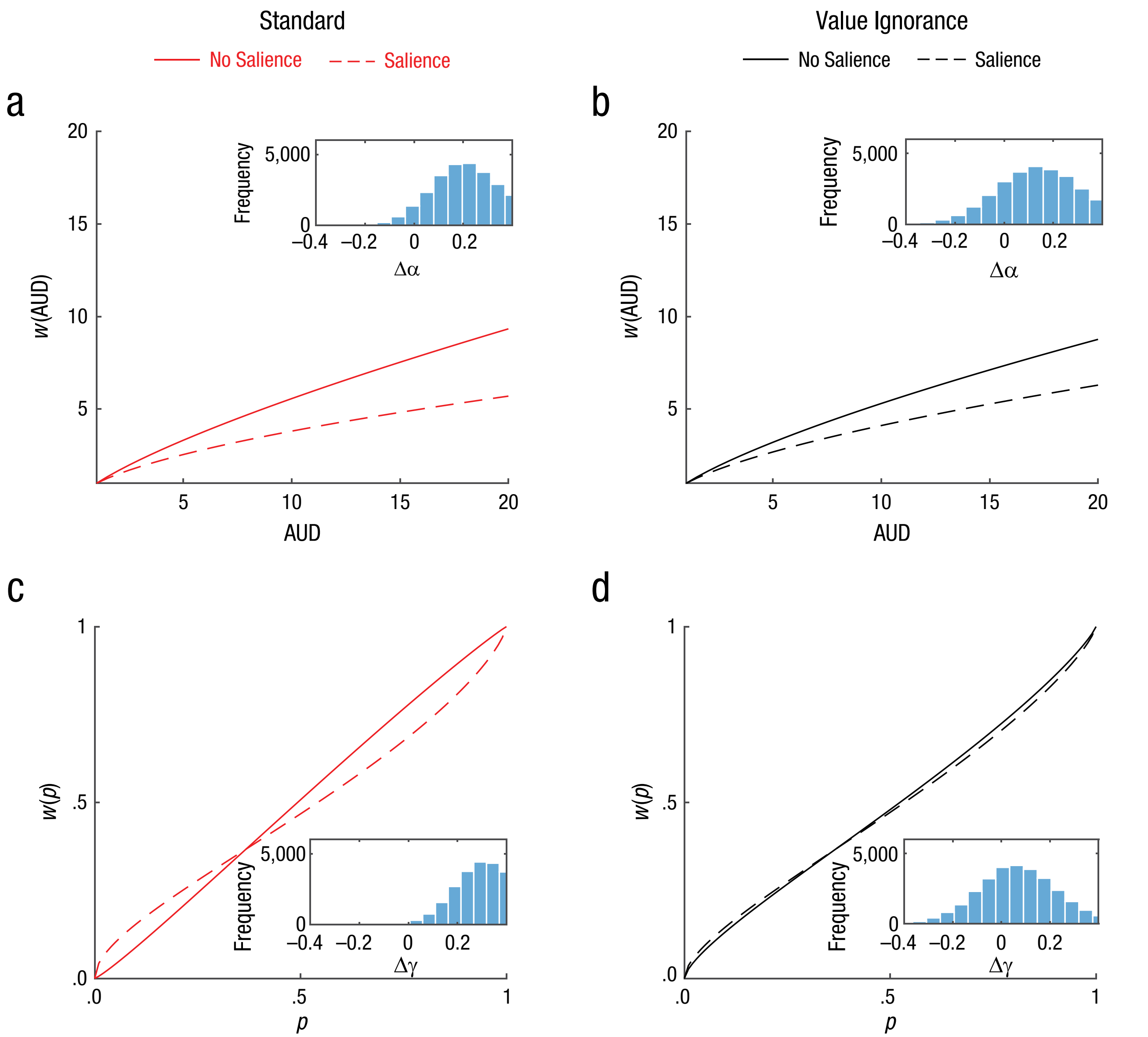

Fig. 6. Value and probability-weighting functions for the salience manipulation in the standard and value-ignorance conditions of Experiments 3 and 4. Each slope shows the best-fitting population-level utility function (a, b) or probability-weighting function (c, d) for the no-salience or salience condition. Histograms show the posterior distribution of population-level differences in parameters (i.e., no salience − salience). AUD = Australian dollars.

methods of combining value and probability—value updating in the standard condition and direct integration in the value-ignorance condition—may give rise to different behaviors. For instance, if the explicit representation of value and probability triggers cognitive mechanisms responsible for CPT-like preferences (as seen in decisions from description), the models can produce the behavior we see in our data.

The salience effect could also be explained within a reinforcement-learning framework: modeled as a multiplicative gain on the impact of highlighted outcomes, where highlighting boosts value signals leading to an overestimation of the riskier option. This could also explain the absence of a salience effect for Type 2 problems (worst outcome salient; see Section 2 in the Supplemental Material) because a multiplicative gain would have no impact on an outcome of \$0. Future studies could test this idea by manipulating the value of the lower outcome.

Memory encoding and retrieval

The allocation of attentional resources provides another mechanistic account. In the value-ignorance condition, participants must retrieve information about the frequency of outcomes from memory in order to make a choice. Our results suggest that when participants retrieve this information, they place additional attention (and higher decision weight) on the rare outcome. This possibility is also supported by the observation that prompting decision makers to mentally "repack" experienced events (sampled in the absence of outcome knowledge) leads to choices more consistent with decisions from description (de Palma et al., 2014).

A related explanation suggests that the mere presentation of outcome values at the choice phase in the value-ignorance condition might lead to more equal emphasis being placed on those outcomes than is warranted. This argument, similar to that proposed by Erev, Glozman, and Hertwig (2008), invokes the notion of a *propositional* (symbolic) *representation* (e.g., blue ball = \$2 vs. blue ball = \$16), leading to greater attentional allocation to the rare outcome in memory. In contrast, when outcomes and frequencies are learned simultaneously, Erev et al. (2008) suggest that an *analogical representation* is formed (e.g., 0, 0, 0, 0, 16, 0, 0, 16, 0), from which the "frequency of the option's events can be read off directly" (Hertwig, 2016, p. 258). This "read-off" of frequencies from memory would lead to the near-linear weighting we observed in the standard condition.

Allocation of attentional resources could also explain the impact of outcome salience on choice and risk preference. In the standard condition, highlighting rare outcomes caused overweighting of rare events and brought risk preferences more in line with those in the value-ignorance condition. This could be explained—in keeping with the mere presentation proposal of Erev et al. (2008)—by an analogical representation in which rare outcomes (e.g., 16) are promoted in the mental sample (e.g., 0, 0, 0, 0, **16**, 0, 0, **16**, 0) during encoding, leading to greater attention and higher decision weighting at the time of choice. The absence of a salience effect under the value-ignorance condition suggests that events must be tied to outcome values to receive this promotion. That the effect would also disappear for Type 2 problems (see Section 2 in the Supplemental Material) likewise indicates the importance of outcome values. Perhaps participants view receiving \$0 as a nonevent, rendering it immune to attentional distortions.

Consistency with the existing literature and a way forward

Although we observed consistent effects across all experiments (see the Supplemental Material), our results highlight the importance of statistical power, experimental control, and good measurement models when studying choice behavior. We analyzed data from 489 participants who chose between many different gambles, and we used a hierarchical Bayesian CPT model to understand how behavior varied across trials, conditions, and participants (see also Glöckner et al., 2016). Our findings show that the parameters of CPT support stronger, more psychologically grounded inferences, compared with the parameters from our statistical models.

Methodological considerations may also shed light on discrepancies between our results and those reported by Hadar and Fox (2009). As part of a larger study focused on the undersampling of rare events, they contrasted conditions in which reward information was revealed at the time of choice with more standard experience-based conditions, but they found no differences. Perhaps one reason we observed differences was because, unlike Hadar and Fox, we ensured that the sample of outcomes that participants observed were representative of the properties of the gambles, thus eliminating noise arising from sampling error. Another important factor is statistical power. With our much larger sample size, we observed consistent but modest effects that may have gone undetected by Hadar and Fox (2009).

Our results provide new insights into the relationships among task demands, experimental procedures, risk preferences, and decision processes. Using subtle manipulations—adjusting when value information was presented—we found notable differences in the way that value information and probability (frequency) information were combined. These results offer novel empirical and modeling pathways for investigating how risk preferences develop in other domains (e.g., losses) across the entire description–experience continuum. At present, frameworks assuming very different cognitive architectures provide plausible accounts of our results, auguring fertile avenues of future investigation for distinguishing between them.

Action Editor

Timothy J. Pleskac served as action editor for this article.

Author Contributions

J. M. Hotaling and A. Jarvstad contributed equally to this research. A. Jarvstad and B. R. Newell conceived and designed Experiments 1 and 2. J. M. Hotaling, B. R. Newell, C. Donkin, and A. Jarvstad developed Experiments 3 and 4. A. Jarvstad wrote the software, and J. M. Hotaling modified it for Experiments 3 and 4. J. M. Hotaling and B. R. Newell supervised data collection. A. Jarvstad, J. M. Hotaling, C. Donkin, and B. R. Newell analyzed and interpreted the data. J. M. Hotaling wrote the first draft of the manuscript, and A. Jarvstad, B. R. Newell, and C. Donkin provided critical revisions. All authors approved the final version of the manuscript for submission.

ORCID iDs

Jared M. Hotaling https://orcid.org/0000-0001-7265-1310
Andreas Jarvstad https://orcid.org/0000-0002-3175-8733

Acknowledgments

We thank Garston Liang and Dominic Tran for help with data collection and Casimir Ludwig for valuable discussions.

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

J. M. Hotaling, C. Donkin, and B. R. Newell were supported by the Australian Research Council (Grant No. DP160101186). A. Jarvstad was supported by a British Academy Postdoctoral Fellowship (No. PF150005).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619884324

Open Practices

TC

All data and materials have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/ry75j. The methods, hypotheses, and analyses for Experiment 4 were preregistered on the Open Science Framework (https://osf.io/zsbwt/). Specific plans for the model-based analysis of risk preferences and the statistical analysis of risky choices were developed after preregistration, and these are detailed in the text. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797619884324. 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 lowest probability outcomes ranged from approximately .1 to .4 across problems (see the Method section), but for simplicity we define rare events as those that occur with a probability of less than .5.
2. Demographic information was not collected in Experiment 1.
3. In Experiment 1, we collected responses to described versions of the gambles from a separate group of participants. In Experiment 2, all participants completed a description phase following the sampling phase in which they responded to the same set of gambles. We do not consider those data further in this article. No description data were collected in Experiments 3 and 4. See the Supplemental Material for further information.
4. The population-mean parameter, μ^{θ}, in this hierarchical model was not equivalent to the average proportion typically used to summarize behavior. We used the hierarchical Bayesian approach here because it outperforms the standard approach in most contexts (Gelman et al., 2013), but we present the standard measures in Tables S1 through S3 in the Supplemental Material.
5. These analyses were performed on Type 1 gambles only. Analyzing both Type 1 and Type 2 gambles produced near-identical results (see Figure S2 in the Supplemental Material).
6. We found some evidence that utility functions were affected by our manipulations, but we refrain from discussing these until more robust evidence is found.

References

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*, 23–26.

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review*, *18*, 377–384.

Camilleri, A. R., & Newell, B. R. (2013). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In V. S. Chandrasekhar Pammi & N. Srinivasan (Eds.), *Decision Making: Neural and Behavioural Approaches* (Vol. 202, pp. 55–71). Amsterdam, The Netherlands: Elsevier.

de Palma, A., Abdellaoui, M., Attanasi, G., Ben-Akiva, M., Erev, I., Fehr-Duda, H., . . . Picard, N. (2014). Beware of black swans: Taking stock of the description–experience gap in decision under uncertainty. *Marketing Letters*, *25*, 269–280.

Erev, I., Glozman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty*, *36*, 153–177.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Boca Raton, FL: CRC Press.

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.

Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*, 317–325.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description–experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*, 493–518.

Hertwig, R. (2016). Decisions from experience. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making* (Vol. 1, pp. 239–267). Oxford, England: Wiley-Blackwell.

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

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). The role of information sampling in risky choice. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 72–91). New York, NY: Cambridge University Press.

Hertwig, R., Hogarth, R. M., & Lejarraga, T. (2018). Experience and description: Exploring two paths to knowledge. *Current Directions in Psychological Science*, *27*, 123–128.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*, 225–237.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, *24*, 1–55.

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, USA*, *110*, 16271–16276.

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

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, *116*, 286–295.

Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes*, *106*, 168–179.

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.

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1446–1465.

Stott, H. P. (2006). Cumulative prospect theory's functional menagerie. *Journal of Risk and Uncertainty*, *32*, 101–130.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1). Cambridge, MA: MIT Press.

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

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

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.

Yechiam, E., & Busemeyer, J. R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making*, *19*, 1–16.