

Decisions from Experience: From Monetary to Medical Gambles

TOMÁS LEJARRAGA*, THORSTEN PACHUR, RENATO FREY and RALPH HERTWIG

Max Planck Institute for Human Development, Berlin, Germany

ABSTRACT

The description–experience gap refers to the robust finding that learning about uncertain options via description or experience results in systematically different choices. This gap has previously been studied primarily with monetary gambles. Here, we examine search and choice processes in decisions from experience involving medical outcomes (side effects of medication). We compare these processes both to decisions from experience involving monetary gambles and to decisions from description involving the same medical outcomes. As in the monetary domain, we found a description–experience gap in medical choices. Yet we also found four striking differences between medical and monetary choices. First, medical choices were significantly less consistent with the maximization of expected value than were monetary choices from description or experience. Second, medical choices gave rise to more strongly inverse S-shaped probability weighting functions in decisions from description and experience, suggesting considerably lower probability sensitivity in the medical than the monetary domain. Third, participants gathered considerably less information in medical than in monetary decisions from experience. Finally, we found that mini-max—a simple decision rule that aims to minimize the maximum possible loss—predicted medical choices substantially better than monetary choices, in decisions from both description and experience. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS decisions from experience; information search; medical choice; monetary choice

INTRODUCTION

On 14 December 1799, George Washington, the first President of the USA, died in his bed at Mount Vernon at the age of 67. Suffering from fever, throat pain, and respiratory distress, Washington did not lack medical attention in his final hours. In fact, he underwent four episodes of bloodletting (Morens, 1999), a treatment he had previously experienced as effective. Though the exact cause of his death is still debated, the loss of 2.5 liters of his blood in a 12-hour period likely caused weakness and worsened some of his symptoms. In Washington's day, bloodletting was a popular and standard medical practice (referred to as “heroic” therapy; Morens, 1999), although some of its risks were known. Benjamin Rush, one of the most famous physicians of the time, championed it as a treatment for a host of illnesses. Yet, based on his experience, he acknowledged that it also had adverse side effects, including “fainting after bleeding,” of which he “saw several instances, in the yellow fever of 1794” (Rush, 1805). Both Washington, the patient, and Rush, the physician, thus drew on their personal experience of the effects of bloodletting to inform their choices.

In a more recent and well-publicized medical choice, philanthropist and Hollywood celebrity Angelina Jolie decided to have a preventive double mastectomy (both breasts removed) on the basis of statistical information provided by her doctors, who told her the surgery would reduce her chances of developing breast cancer “from 87 percent to under 5 percent” (Jolie, 2013). Unlike Washington, Jolie had access to a description of the risky medical prospects she faced, presumably including the risks of surgery, and so was able to make a *decision from description* (Hertwig, Barron, Weber, & Erev, 2004). Lacking such convenient descriptions of potential medical outcomes

and their likelihoods, Washington and physicians like Rush had no alternative but to make *decisions from experience* (Hertwig et al., 2004).¹

Nowadays, medical choices frequently take the form of decisions from description. For instance, the patient information leaflets provided with prescription medications (e.g., “approximately 1 in 500 patients experience a side effect”) communicate possible adverse effects as well as their likelihoods. Yet for numerous treatments—including experimental treatments, over-the-counter medications, homeopathic, and household remedies—there are no easily accessible descriptions of potential risks. Patients considering such treatments may thus find themselves making medical choices in the same way Washington and Rush did—by relying on their own experience and that of others. Moreover, information gleaned from the experience of others appears to have a strong impact on how people deal with risks (Pachur, Hertwig, & Steinmann, 2012), influencing decisions even when descriptive summary information is available (Haase & Betsch, 2012).

A recent line of research has found that decisions from description and decisions from experience can result in robust and predictable differences in choice that are collectively known as the *description–experience gap* (for a review, see Hertwig, in press). To date, however, the description–experience gap

¹Note that although Washington and Rush made decisions from experience, their experiences were of different kinds. Washington directly suffered from the side effects, as does any patient who undergoes medical treatment and experiences side effects. Doctors, in contrast, witness their patients' side effects but they do not experience them as directly as the patients do. This difference—the extent to which experience has immediate and costly consequences to the decision maker—is reflected in the distinction between the *feedback paradigm* and the *sampling paradigm* in research on decisions from experience (Hertwig & Erev, 2009). Admittedly, however, this distinction, like any dichotomy, is not as categorical in reality as in theory. As doctors carry some responsibility for their patients' well-being, they also carry some of the costs and benefits associated with the outcomes of a treatment.

*Correspondence to: Tomás Lejarraga, Max Planck Institute for Human Development, Berlin, Germany. E-mail: lejarraga@mpib-berlin.mpg.de

has been studied almost exclusively in the context of payoff distributions representing monetary outcomes. In this article, we explore whether it generalizes to medical outcomes such as side effects resulting from medication.

EXPERIENCE-BASED MONETARY CHOICES

Several experimental paradigms have been used to study decisions from experience (Hertwig & Erev, 2009); here we focus on the sampling paradigm. In this experimental approach, people sample as many outcomes as they wish from two (or more) options, or payoff distributions, before choosing one option for the final incentivized draw. Numerous studies comparing such experience-based choices with description-based choices have shown a systematic gap in choice: In decisions from experience, people behave as if rare events have less impact than they deserve according to their objective probabilities; in decisions from description, people behave as if rare events have more impact than they deserve.

Several factors have been proposed to contribute to the description–experience gap in the sampling paradigm (Camilleri & Newell, 2013; Hertwig, in press). One is reliance on relatively small samples, which do not reflect the objective probabilities of outcomes altogether faithfully, especially when those outcomes are rare. Another idea is that, in decisions from experience, rare events have less impact than they deserve according to their objective probabilities; this conclusion was originally derived from the systematic pattern of observable choices (e.g., Hertwig et al., 2004, p. 535; see also Weber, Shafir, & Blais, 2004). Subsequently, researchers have fitted cumulative prospect theory's probability weighting function to decisions from experience, with mixed results (e.g., Jarvstad, Hahn, Rushton, & Warren, 2013; Ungemach, Chater, & Stewart, 2009). Still another factor that has been demonstrated to contribute to the description–experience gap is search policy: People who switch frequently between options have been observed to deviate more strongly from the pattern of description-based choices than infrequent switchers (Hills & Hertwig, 2010). All these findings and conclusions have been drawn on the basis of behavior in monetary payoff distributions. Therefore, it remains unclear to what extent they carry over to the medical realm.

EXPERIENCED-BASED MEDICAL CHOICES

In research on medical decision making, it is widely thought that patients' choices can and should be modeled in terms of the expected utility framework (Hunink et al., 2001; Russell & Schwartz, 2012). People are thus expected to behave as if they multiplied a subjective nonlinear value function of the outcomes in question (e.g., treatment success and adverse side effects) by a linear function of those outcomes' probabilities for each option and then maximized (i.e., chose the option with the higher expected utility). Yet ample evidence shows that patients' medical

choices systematically violate expected utility theory. One key violation is a relative disregard of probabilities. A number of investigations have found indications that, when making medical decisions from description, people appear to focus on the magnitude of an outcome at the expense of its actual probability (e.g., Pachur & Galesic, 2013; Pachur, Hertwig, & Wolkewitz, 2014; Russell & Schwartz, 2012; Waters, Weinstein, Colditz, & Emmons, 2007, 2009). To the extent that this focus on outcomes over probabilities, sometimes dubbed *probability neglect* (Sunstein, 2002), generalizes to decisions from experience, it is instructive to find out how it may influence search effort, search policy, and resulting choices.

Search effort

In decisions from experience, people striving to form an accurate representation of the worth of options may reasonably be expected to draw large samples from each; after all, the outcome probabilities in a larger sample are more likely to accurately reflect the underlying population parameters than are those in a smaller sample (law of large numbers; see Stigler, 1986). By contrast, people inclined to disregard probabilities may be satisfied with small(er) samples. If people making medical choices are indeed primarily interested in outcomes, their search can thus be expected to be less extensive than in the context of monetary choices. The opposite is also conceivable, however. Keen to find out the range of possible outcomes (but not attending to their respective probabilities), people might be inclined to search extensively to learn about all possible outcomes, including the worst-case scenario. Furthermore, pronounced affective responses to the prospect of adverse medical outcomes (Loewenstein, 2005; Pachur et al., 2014) may also prompt more rather than less exploration: Hypochondria, an extreme form of health anxiety, manifests itself as hypervigilance. Relatedly, cancer worry has been reported to result in more attention being paid to health information (Burke Beckjord, Finney Rutten, Arora, Moser, & Hesse, 2008) and to greater behavioral intentions to look for additional information about preventive medication (Dillard, Scherer, Ubel, Smith, & Zikmund-Fisher, 2013). In light of these findings, it is an open question whether medical choices will involve more or less search than monetary choices (see also Frey, Hertwig, & Rieskamp, 2014; Yechiam & Hochman, 2013).

Search policy

Another important property of information search in the sampling paradigm is how people search (Pachur & Scheibehenne, 2012). Hills and Hertwig (2010) found a link between how often people switched between payoff distributions in the process of exploring them and how they finally chose. Whereas infrequent switchers tended to make choices on the basis of summary comparisons (e.g., averaging), frequent switchers generally compared options in a round-wise manner, and tallied which option won most rounds. If people focus on outcomes at the expense of probabilities in a medical domain,

they may also exhibit a higher rate of switching between options, because switching permits them to compare outcomes directly.

The following investigation addressed two issues: First, does the description–experience gap generalize from monetary to medical outcomes? Second, to what extent do differences in search effort and search policy cause experienced-based medical choices to systematically differ from experienced-based monetary choices?

METHOD

Ninety students at the University of Basel participated in the experiment (24 male, $M_{\text{age}} = 23.7$ years, $SD_{\text{age}} = 6.56$), which consisted of an evaluation stage and a choice stage. In the evaluation stage, participants were asked to indicate the amount in Swiss francs (Sfr) they were willing to pay (i.e., their willingness-to-pay; WTP) to avoid each of 12 pharmaceutical side effects. The side effects were selected on the basis of a pilot study with 60 participants who expressed their WTP to avoid 40 side effects that the *Swiss Online Pharmaceutical Compendium* (<http://www.compendium.ch>) lists as associated with analgesics.

Participants read the following text:

You are receiving treatment for chronic pain. Several existing medications can be used to treat the pain. All are equally potent and differ only in their side effects. For each of these old medications, a new medication has recently been developed that is equally effective but does not have the side effect of the corresponding old medication. We are interested in the amount of money you would be willing to pay as a markup for the new medication without the side effect.

Side effects that we anticipated to be difficult to decipher were briefly explained (e.g., lalopathy: slow or unarticulated speech). Using a method developed and validated by Pachur and Galesic (2013) and Pachur et al. (2014), the WTPs obtained during the evaluation stage were subsequently used to construct monetarily equivalent medical and monetary choice problems for each participant. Participants were then presented with 18 choice problems, 9 from the medical domain and 9 from the monetary domain (domain order was counterbalanced across participants). In the medical domain, they chose between two equally effective treatments with different potential side effects, such as one treatment with a 70% probability of causing fatigue and an equally effective treatment with a 30% probability of causing dizziness. The monetary problems were identical to the medical problems, except that each side effect (e.g., fatigue) was replaced by the WTP the participant provided in the evaluation phase. The medical and monetary choice problems were thus monetarily equivalent, allowing within-participant comparison between domains. To illustrate, a participant who was willing to pay 10 Sfr to avoid fatigue and 30 Sfr to avoid dizziness was presented with an equivalent monetary choice between losing 10 Sfr with a probability of 70% and losing

30 Sfr with probability of 30%. The medical problems are listed in Table 1.

We sought to use side effects that participants would recognize (e.g., fatigue and diarrhea) and feel competent in evaluating. As recognizable side effects are typically not rare, 16 of the 18 problems did not include rare events (i.e., events with probabilities $\leq 20\%$). Thus, we were also able to examine the extent to which the description–experience gap arises in choice problems devoid of rare outcomes, an issue that has received scant research attention to date (but see Ludvig & Spetch, 2011, who examined this case in a repeated choice paradigm; for a description of the differences between the sampling and repeated choice paradigms, see Barron & Erev, 2003).

One group of participants ($n = 30$) made decisions from experience. Figure 1a is a screenshot of their task (translated from German). Another group ($n = 30$) made decisions from description, with outcome probabilities expressed as single-event percentages (Figure 1b). To examine whether the format of the described risk affected the results (it did not, as reported below), we presented a third group of participants ($n = 30$) with described risks expressed as absolute frequencies (the likelihood of medical side effects is often presented in terms of frequencies). This resulted in a 2×3 mixed design, with two domains (medical versus monetary) administered *within* participants, and three sources of information (experience versus description of probabilities or frequencies) administered *between* participants. Because we had no a priori estimate of the effect size, we arbitrarily decided to collect data from 90 participants and stopped when we reached that number.

All participants received a flat fee of 15 Sfr for their participation. Monetary and medical choices were thus hypothetical (in contrast to nearly all previous studies of decisions of experience with monetary outcomes; see, e.g., Hertwig et al., 2004). We implemented hypothetical choices because medical outcomes (e.g., dizziness) are difficult to actually implement in an experimental setting. Furthermore, hypothetical medical decisions made on the basis of vignettes have been shown to correlate highly with actual clinical practice (Peabody, Luck, Glassman, Dresselhaus, & Lee, 2000). By implementing all choices as hypothetical, we further avoided any confound between monetary and medical choices: Any differences between domains would not stem from incentivization or lack thereof. Finally, relying on hypothetical choices permitted us to examine to what extent the typical results found in previous investigations of the sampling paradigm hinged on incentivized choice.

RESULTS

Figure 2 shows participants' hypothetical WTPs to avoid each of the 12 side effects. For example, participants were willing to pay a median of 15 Sfr to avoid fatigue, but they were willing to pay more than three times as much, 50 Sfr, to avoid memory loss.

Table 1. Choice problems, proportion of choices of the option with the higher expected value (EV), and size of the description–experience (D–E) gap in the decision from description (D) and decision from experience (E) conditions

						% of EV choices				Absolute size of the D–E gap	
Medication A		Medication B		% times A is the higher EV option	Monetary		Medical				
Problem	Side effect	Probability	Side effect		Probability		D	E	D	E	Monetary
1	Flatulence	1.00	Hallucinations	0.25	0.38	72	71	63	68	0	5
2	Fatigue	0.90	Memory loss	0.25	0.54	73	70	61	43	3	18
3	Itching	0.70	Depression	0.25	0.30	75	83	57	57	8	0
4	Fever	0.40	Hallucinations	0.20	0.48	83	56	71	67	27	4
5	Insomnia	0.55	Depression	0.30	0.29	88	60	56	47	28	9
6	Lalopathy	0.30	Memory loss	0.20	0.27	73	63	58	67	10	8
7	Fatigue	0.70	Dizziness	0.30	0.38	80	67	58	47	13	11
8	Itching	0.60	Trembling	0.50	0.34	87	73	73	67	13	7
9	Flatulence	0.70	Diarrhea	0.40	0.30	77	90	75	38	13	37
					<i>M</i>	78.6	70.3	63.5	55.5	8.2	8.1

Note. Each medical problem involves a choice between two medications. For example, problem 1 involves a choice between medication A, which has the side effect of flatulence (with a probability of 1.0) and medication B, which causes hallucinations with a probability of 25% (and no side effects otherwise). Medications A and B were described as equally effective, differing only in the type and likelihood of the side effect. The monetary problems, which match the probability structure of medical problems, were constructed on the basis of the individuals' willingness-to-pay amounts (see Method). Therefore, the option with the higher EV can differ across participants. The sixth column indicates the percentage of participants (across conditions) for whom option A was the higher EV option.

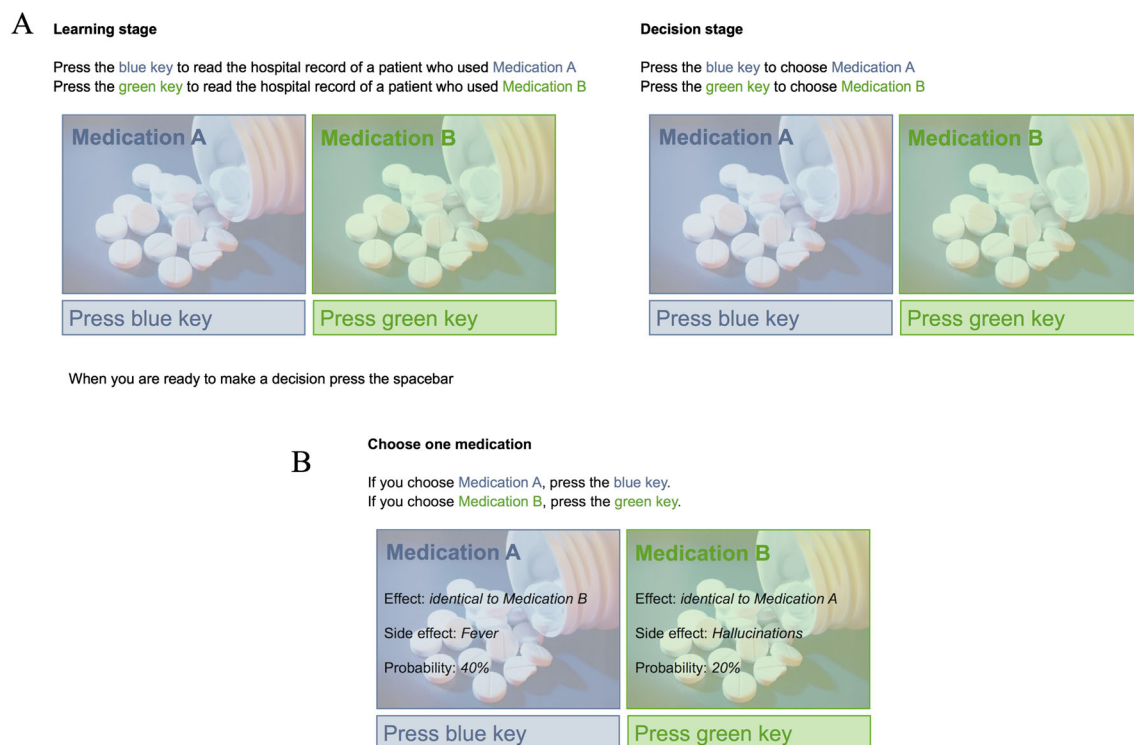


Figure 1. Screenshots with examples of a task requiring medical decisions from experience (1A; sampling and choice phase; see Hertwig et al., 2004) and medical decisions from description (1B)

Choices

Do medical and monetary choices maximize expected value to a similar degree? To answer this question, we used a mixed-effects logistic regression to predict whether participants chose the option with the higher expected value (hereafter, EV choices). Source of information and domain of choice were treated as fixed effects, and choice problem and participants as random effects. The expected value of an option was defined on the basis of its objective probability

(Table 1) and outcome value (i.e., person-specific WTPs). Table 1 shows the problems and choice proportions separately for each condition. As a consequence of variability in the monetary evaluation of the side effects, the option with the higher expected value in each problem can differ across individuals. As the format of described risk (probability versus frequency) did not influence the results, $\chi^2(1)=.03$, $p=.86$, we collapsed across both description-based conditions in the following analyses. We used a step-wise procedure in the

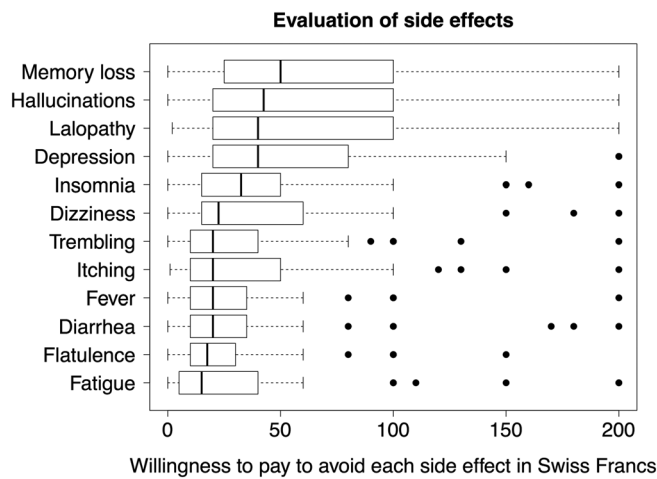


Figure 2. Boxplots showing the amounts that people were willing to pay to avoid each of the 12 side effects. The right and left limits of the boxes denote the first and third quartile. The whiskers denote the lowest datum within 1.5 interquartile range (IQR) of the 1st quartile, and the highest datum within 1.5 IQR of the 3rd quartile. Amounts larger than 200 Sfr are not shown in the figure (but are included in the boxplots)

logistic regression analysis, first testing a full model (with all relevant variables and the interaction between domain and source of information) and then examining the effect of excluding one relevant variable on the model fit (using a chi-squared test).

Monetary versus medical choices

Domain proved to be a significant predictor of EV choices, $\chi^2(1)=43.41$, $p<.01$.² Table 1 shows the proportion of EV choices in each domain. In decisions from experience, the proportion of EV responses was 70% in monetary choices, relative to 56% in medical choices. The same pattern emerged in decisions from description, with 79% EV responses in monetary choices and 64% in medical choices.

Description versus experience

Consistent with previous findings (e.g., Hertwig et al., 2004), the proportion of EV choices differed between the experience and the description conditions, $\chi^2(1)=7.01$, $p<.01$. The size of this description–experience gap was similar in the medical and monetary domains, as the interaction of domain and source of information was nonsignificant, $\chi^2(1)=.10$, $p=.74$. Table 1 shows the proportion of EV choices for each of the problems. These findings demonstrate that the description–experience gap extends beyond monetary choices; equally importantly, it extends beyond options with rare events and also occurs in hypothetical choices.

²The chi-squared statistic and corresponding p value were derived from comparing a model excluding domain of choice with a full model including domain of choice. Accordingly, this comparison has one degree of freedom. Including domain of choice significantly improved the model fit.

Probability weighting

Pachur et al. (2014) contrasted monetary and medical choices in terms of cumulative prospect theory's (CPT; Tversky & Kahneman, 1992) probability weighting function. This function expresses how objective probabilities are translated into decision weights, with more strongly curved functions reflecting lower sensitivity to differences in probabilities. In decisions from description, Pachur et al. found that choices in the medical domain gave rise to a considerably more strongly curved weighting function than did choices in the monetary domain, indicating reduced probability sensitivity in medical choices. To examine whether a similar difference in probability weighting emerges in decisions from experience, we modeled each participant's choices using a hierarchical Bayesian approach (see Appendix for details). We estimated the parameters in the experience condition by using the objective probabilities of the problems (see Table 1; the estimates for the other CPT parameters as well as a probability weighting analysis based on the probabilities derived from the frequencies of outcomes that participants actually experienced are reported in the Appendix).³

Figure 3 shows the resulting individual-specific probability weighting functions, separately for each of the four conditions, with the monetary and medical domains in the left and right panels, respectively, and the description and experience conditions in the upper and lower panels, respectively. There is considerably lower probability sensitivity in the medical domain than in the monetary domain, indicated by strongly inverse S-shaped probability weighting functions for medical choices. This result generalizes the findings of Pachur et al. (2014) to decisions from experience: medical choices appear to be characterized by reduced probability sensitivity. Further, note that the probability weighting functions in the monetary domain suggest linear probability weighting in decisions from description, but a small degree of overweighting in decisions from experience. This suggests that underweighting of rare events caused by their underrepresentation in experienced samples due to limited search is not the major factor prompting the description–experience gap in the present experiment (in contrast to the findings of Hertwig et al., 2004). One possible reason is that only around 10% of the decision problems in our study included a rare event (see Method).

Search effort

Are people inclined to search for more or less information in the medical than in the monetary domain? Figure 4 plots the average number of samples drawn by each of the 30 participants. Within participants, sample sizes in the medical and the monetary domains correlated substantially (Spearman rank correlation), $r_s=.68$, $S=1438$, $p<.001$. At the same time, the amount of sampling differed notably between domains. Participants drew considerably fewer samples in

³We prioritize the analysis using objective probabilities—rather than “experienced” probabilities—because we want to characterize the patterns of choices irrespective of the effects of differences in search effort.

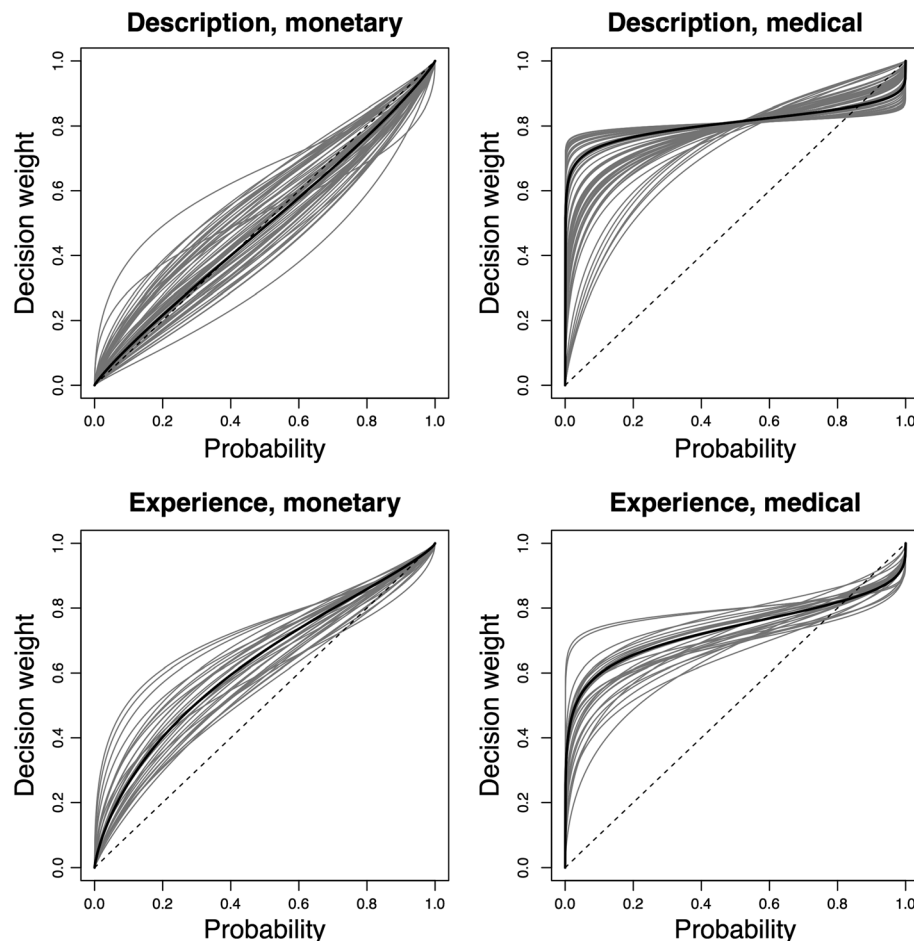


Figure 3. Probability weighting functions estimated for each participant's choices in the monetary domain (left) and the medical domain (right), in the description (above) and experience (below) conditions. The gray lines indicate the curves for each individual (60 curves in the description condition; 30 curves in the experience condition); the black lines represent the curves based on the group-level mean parameters

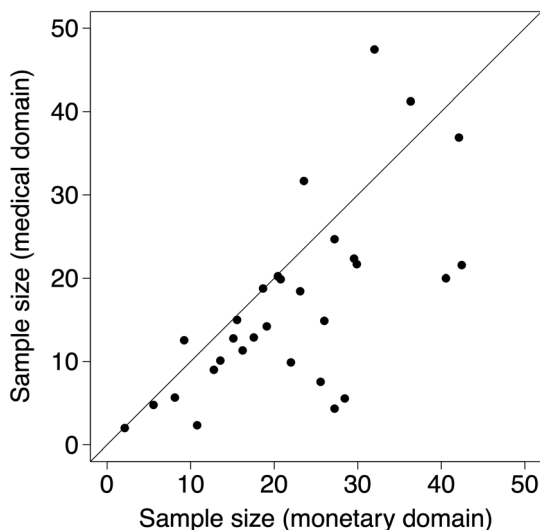


Figure 4. Search effort (in terms of sample size) in the medical and the monetary domain, respectively. The data points represent the mean sample size for each individual across the nine problems in each domain. Data points below the diagonal indicate larger samples in the monetary domain

the medical than in the monetary domain, $M=16.7$, $SD=11.2$ vs. $M=22.1$, $SD=10.5$, $t(29)=-3.3$, $p<.01$ (paired t -test), and most data points (25 out of 30) in Figure 4 fall below

the diagonal. It is important to note that the average sample size in the monetary domain was larger than the median of 16 found in a meta-analysis of previous studies (Hertwig, in press). This finding is especially noteworthy given that monetary choices were not incentivized. Two properties of our choice problems are likely to have led to the larger sample size: All options involved (a) losses and (b) relatively common events (thus causing people to experience variability in outcomes even on the basis of few samples). Both the danger of losses and exposure to variability of outcomes have been found to prompt substantially more search (Lejarraga, Hertwig, & Gonzalez, 2012).

Search policy

To examine differences in search policy, we calculated the ratio between a participant's number of switches and the maximum possible number of switches (i.e., the total sample size minus 1; Hills & Hertwig, 2010) for each problem. We then determined individuals' average switching rate separately in each domain. In both the monetary and the medical domain, the distribution of participants by switching rate was bimodal, with a large cluster of infrequent switchers and a smaller cluster of frequent switchers (Figure 5). These results are consistent with Hills and Hertwig's observation that

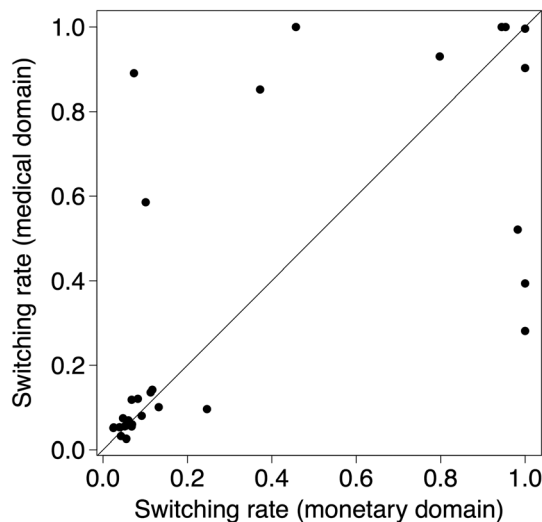


Figure 5. Individual switching rates in the medical and the monetary domain, respectively. Each data point represents one participant from the experience condition. Data points below the diagonal indicate more switching in the monetary domain

people seem to adopt one of two distinct search policies. Most participants appeared to carry their policy across domains, as suggested by the high correlation between domains, $r_s = .84$, $p < .01$ (Spearman's rank correlation). Moreover, on average, policies did not appear to differ across domains: monetary ($M = .34$, $SD = .4$) versus medical ($M = .36$, $SD = .39$), $t(29) = .40$, $p = .60$.

Choice policy

Still another way to examine the hypothesis that people tend to focus on adverse medical outcomes at the expense of their likelihoods is to determine the proportion of decisions in which participants chose the option promising the less bad outcome. Savage (1951) called this choice policy *minimax*. In those problems in which the maximum loss differed between options (i.e., the problems that permit this analysis), minimax predicted 68% of medical choices (65% in experienced-based and 70% in description-based decisions), relative to 52% of monetary choices (55% in experienced-based and 51% in description-based decisions), $\chi^2(1) = 36.5$, $p < .01$ (Figure 6). When the analysis was run on the distributions of outcomes that participants actually experienced from the two options in each problem, minimax predicted 73% of medical and 59% of monetary choices, $\chi^2(1) = 10.6$, $p < .01$.

DISCUSSION

Our investigation of experienced-based and description-based choices involving monetary and medical outcomes produced several notable findings. First, we found a description–experience gap of similar size in both domains (Table 1). This finding supports the observation of Ludvig and Spetch (2011) that the presence of rare events is an enabling, but not a necessary, condition for the description–experience

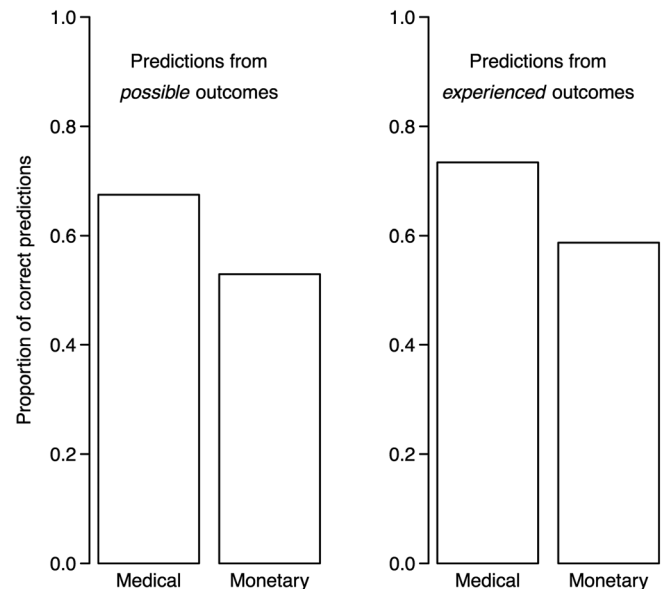


Figure 6. Proportion of correct predictions of minimax derived from the lowest possible loss (irrespective of whether an individual experienced it) and from the outcomes that were experienced (i.e., the lowest experienced loss)

gap (Hertwig & Erev, 2009). Moreover, our finding extends their observation from the repeated choice paradigm to the sampling paradigm. Second, participants were less likely to choose the option with the higher EV in the medical “gamble” than in the monetary gambles (Table 1). Third, people’s decisions from experience in the medical domain were accompanied by a probability weighting pattern suggesting reduced probability sensitivity. Fourth, these regularities held in decisions from both experience and description (Table 1). Fifth, taking advantage of the sampling paradigm’s ability to lay bare the processes of information search, we found that people sampled substantially fewer outcomes in the medical than in the monetary domain (Figure 4). This finding conforms with previous reports that people rendering decisions from description search less extensively for probability information in the medical than in the monetary context (as also observed in the process-tracing study by Pachur et al., 2014). Finally, participants were more likely in the medical than in the monetary domain to choose as if their goal was to avoid the worst possible outcome (Figure 6).

In the medical domain, we found two behavioral regularities to coexist: reliance on relatively small samples (at least, smaller than in the monetary domain) and a tendency to choose with the goal of minimizing the worst possible loss. Of course, reliance on small samples may well be inconsistent with the objective of minimizing the worst possible loss, especially if that loss happens to occur rarely. For instance, bloodletting may have been popular because deaths like Washington’s—likely to have been caused by the co-occurrence of grave illness and general weakness caused by the “heroic therapy”—were still rare enough not to receive sufficient attention (or weight, in modern terminology) in physicians’ limited sample of experience. Reliance on small samples and the aspiration to avoid the worst loss have

consequences for the weighting of rare events. Individuals drawing small samples may not encounter a rare side effect and, consequently, act as if they were underweighting its probability. In contrast, participants who encounter rare events and choose with the goal of avoiding them may choose as if they were underweighting them to a lesser extent—or even overweighting them. Our analysis, based on side effects with mostly intermediate probabilities, suggests the latter (Figure 6). Future work should extend this analysis to a wider range of probabilities.

Limitations

In order to make medical and monetary choices monetarily equivalent, we elicited WTPs, which some have criticized as unreliable (Ariely, Loewenstein, & Prelec, 2003). Could the WTP technique have influenced our results? We doubt it, for two reasons. First, our analyses of search effort and policy are independent of the WTPs. Second, whereas unreliable WTPs would be expected to make choices more noisy and thus less predictable, our analysis suggested a relatively consistent pattern: Minimax correctly predicted 70% of the description-based and 73% of the experienced-based medical choices. In the medical domain, no matter whether information came in the form of description or in the form of experience, many participants appear to have aspired to avoid the worst loss. Furthermore, the WTP elicitation procedure we used has been shown to yield reliable WTPs over time (Pachur & Galesic, 2013) and not to cause more random guessing in the medical than in the monetary domain (Pachur et al., 2014).

Another limitation is the extent to which the medical problems used in our study represent real medical choices. One way in which our problems differ from many real problems is that the alternative medical treatments were presented as “equally effective,” meaning that participants could not trade off benefits against harms, as they would be likely to do in actual decisions between different medications. Would probability neglect also emerge in problems involving treatments that differed in effectiveness, permitting people to trade off benefits against harms? Waters et al. (2009) examined medical decisions in such situations. For example, one of their cases stated that “... [t]he doctor says that a new drug would decrease your risk of developing colon cancer from 23% to 4%. But the drug has an independent side effect that would increase your risk of stomach cancer from 2% to 6% ...” (p. 208). Even in this potential cost–benefit tradeoff scenario, participants seemed to ignore probability information.

Yet the participants of Waters et al. (2009) made hypothetical rather than real medical choices, as did our respondents. Peabody et al. (2000) examined to what extent experimental medical choices made on the basis of vignettes depicting hypothetical scenarios are appropriate proxies for real medical choices. The authors used an ingenious method: Trained actors visited real clinical practices and reported specific symptoms. Several weeks later, corresponding vignettes were sent to the same physicians. It was, thus, possible to compare physicians’ responses to the vignettes with

their actual medical responses to the “simulated” symptoms. Both responses were highly correlated, leading the authors to conclude that vignettes appear to be a valid method for probing the process of care provided in actual medical practice. Of course, our participants were acting as patients rather than physicians, and we subjected them to a sampling process rather than vignettes. Overall, however the various lines of research suggest that it is not unreasonable to examine hypothetical choices and to suspect that the aspiration to avoid the worst possible medical loss (at the expense of disregarding probabilities) found in the laboratory may generalize to real choices.

CONCLUSIONS

Because a complete description of the world’s potential dangers is unattainable, people’s perceptions of harms and risks and their subsequent choices will also always be shaped by personal experience. And even if people enjoy access to stated probabilities and outcomes, they often also bring personal experience to bear on their decision. That was indeed the case for Angelina Jolie: Her mother died at the age of 56, after nearly a decade fighting cancer (Jolie, 2013). Our study contributes to a better understanding of the search and choice policy in “pure” decisions from experience in the medical context. Future studies can now begin to analyze how the simultaneous presence of descriptive and experiential information interacts in people’s medical choices (Lejarraga & Gonzalez, 2011; Newell & Rakow, 2007) and explore how treatment benefits (gains)—not only side effects (losses)—affect medical decisions from experience.

APPENDIX

Cumulative Prospect Theory

According to cumulative prospect theory (CPT), in the two-outcome lotteries with only one nonzero outcome used in our study the subjective valuation V of option A is determined as

$$V(A) = \sum_{i=1}^n v(x_i)w(p_i), \quad (A1)$$

where $v(x_i)$ is the subjective value of outcome x_i , defined according to the following value function:

$$v(x_i) = \begin{cases} x_i^\alpha, & \text{if } x_i \geq 0 \\ -(-x_i)^\alpha, & \text{if } x_i < 0 \end{cases} \quad (A2)$$

The parameter α reflects the sensitivity to differences in outcomes and is assumed to lie in the range $[0, 1]$. This yields a concave value function for gains and a convex one for losses. Note that because the choice problems we investigated did not involve options with both gains and losses, we did not fit a loss-aversion parameter. In Equation A3, $w(p_i)$ is the probability weighting function that translates objective probabilities p_i into subjective decision weights:

$$w(p_i) = \frac{\delta p_i^\gamma}{\delta p_i^\gamma + (1 - p_i)^\gamma}. \quad (\text{A3})$$

The parameter γ reflects the sensitivity to differences in probabilities, with values <1 yielding a more inverse S-shaped curvature (indicating overweighting of small probabilities) and values >1 yielding a more S-shaped curvature (indicating underweighting of small probabilities). The parameter δ governs the elevation of the weighting function and can be interpreted as a measure of risk aversion (with $\delta > 0$; Gonzalez & Wu, 1999). CPT predicts that the option with the more attractive V is preferred.

To derive the probability that option A is preferred over option B, we used an exponential version of Luce's (1959) choice rule:

$$p(A, B) = \frac{e^{\phi \cdot V(A)}}{e^{\phi \cdot V(A)} + e^{\phi \cdot V(B)}}, \quad (\text{A4})$$

where ϕ is a choice-sensitivity parameter governing how sensitively the choice probability reacts to differences in the valuation of options A and B. A higher ϕ indicates more deterministic behavior; with $\phi=0$, choices are random.

In this implementation, CPT has four adjustable parameters (Equations A1–A4): outcome sensitivity (α), probability sensitivity (γ), elevation (δ), and choice sensitivity (ϕ). These parameters were estimated for each participant using a

yielding a uniform distribution from 0 to 1 on the individual level. In order to extend the range of these distributions from -2 to 2 for δ and from 0 to 5 for β , γ , and ϕ , we interposed an additional linear linkage function. All hierarchical group-level means were assumed to be normally distributed, with a mean of 0 and a variance of 1 (yielding uniform distributions on the individual level). The priors on the group-level standard deviations were uniformly distributed, ranging from 0 to 10 (thus avoiding extreme bimodal distributions on the individual level).

The joint posterior parameter distributions were estimated using Markov chain Monte Carlo methods implemented in JAGS, a sampler that utilizes a version of the BUGS programming language (version 3.3.0) called from MATLAB. We ran three chains, each with 30,000 iterations, which were drawn from the posterior distributions after a burn-in period of 1000 samples. To reduce autocorrelations during the sampling process, we recorded only every 10th sample. The sampling procedures were efficient, as indicated by low autocorrelations of the sample chains, Gelman–Rubin statistics, and visual inspections of the chain plots.

The group-level means for each parameter are reported in Table A1. Based on the individually estimated parameters, CPT accommodated, on average, 86.3% ($SD=17.2$) and 74.3% ($SD=17.1$) of participants' monetary and medical choices, respectively, in the description condition. In the experience condition, CPT accommodated 75.2% ($SD=21.6$) and 72.6% ($SD=22.9$) of participants' monetary and medical choices, respectively.

Table A1. Group-level means of the CPT parameters. For decisions from experience (E), the results are shown separately for estimates based on objective (obj.) and experienced (exp.) probabilities, respectively

Domain		CPT parameter			
		Outcome sensitivity (α)	Probability sensitivity (γ)	Elevation (δ)	Choice sensitivity (ϕ)
D	Monetary	0.630	0.894	0.955	3.844
	Medical	0.389	0.202	4.331	2.244
E (obj.)	Monetary	0.376	0.794	2.018	3.294
	Medical	0.289	0.310	2.921	1.481
E (exp.)	Monetary	0.497	0.814	0.868	2.484
	Medical	0.481	0.534	3.816	1.570

hierarchical Bayesian approach (Nilsson, Rieskamp, & Wagenmakers, 2011; Scheibehenne & Pachur, 2015), separately for the medical and monetary conditions and also separately for decisions from experience and decisions from description. The advantage of a hierarchical approach is that individual parameters are partially pooled through group-level distributions, which promises to yield more reliable estimates than the traditional, nonhierarchical approach. The priors for the parameters on the individual level were set to uniform probability distributions spanning a reasonable range that excluded theoretically impossible values, but included parameter values found in previous research. Specifically, the ranges were 0 to 5 for ϕ , δ ; 0 to 1 for α ; and 0 to 2 for γ . The group-level parameters were linked with the individual level through probit transformations (Scheibehenne & Pachur, 2015),

Estimating CPT Parameters Using Experienced Probabilities

In addition to estimating the probability weighting functions based on the problems' objective probabilities, we also estimated CPT parameters using the experienced probabilities (i.e., the individual-specific experienced frequency for each outcome). The results, reported in Table A1 and Figure A1, provide additional evidence of considerably lower probability sensitivity in the medical domain, whereas probability weighting in the monetary domain is essentially linear. Based on the parameters estimated for each participant, CPT accommodated, on average, 72.6% ($SD=22.9$) and 70.4% ($SD=17.8$) of participants' monetary and medical choices, respectively.

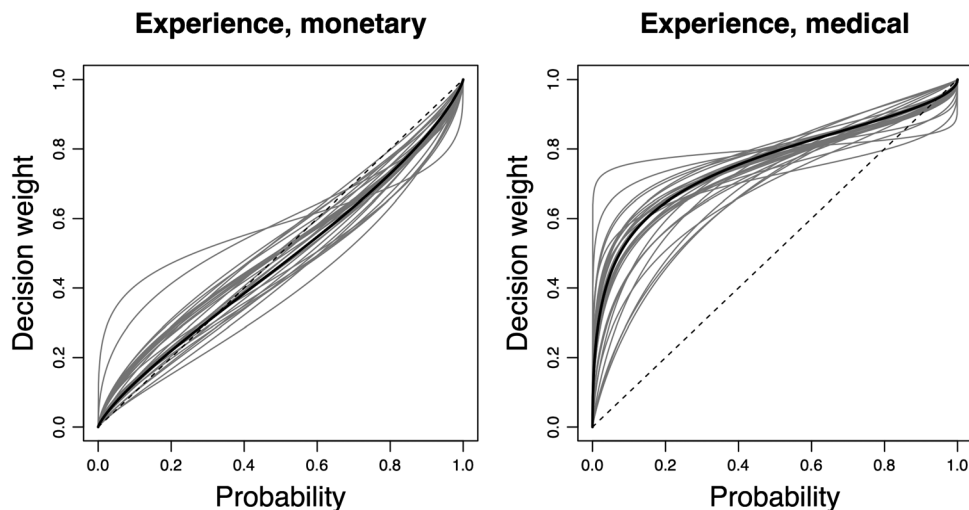


Figure A1. Probability weighting functions estimated for each participant's choices in the experience condition in the monetary domain (left) and the medical domain (right). The gray lines indicate the curves for each individual; the black lines represent the curves based on the group-level mean parameters. The estimates are based on the experienced probabilities for each option.

ACKNOWLEDGEMENTS

We thank Patrick Riepl for collecting the data and Susannah Goss for editing the manuscript. This research was supported by Swiss National Science Foundation grant CRSII1_136227. Correspondence concerning this article should be addressed to Tomás Lejarraga, Max Planck Institute for Human Development, Center for Adaptive Rationality (ARC), Lentzeallee 94, 14195 Berlin, Germany

REFERENCES

- Ariely, D., Loewenstein, G., & Prelec, D. (2003). Coherent arbitrariness: Stable demand curves without stable preferences. *Quarterly Journal of Economics*, 118, 73–105.
- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, 16, 215–233.
- Burke Beckjord, E., Finney Rutten, L. J., Arora, N. K., Moser, R. P., & Hesse, B. W. (2008). Information processing and negative affect: Evidence from the 2003 Health Information National Trends Survey. *Health Psychology*, 27, 249–257.
- Camilleri, A. R., & Newell, B. R. (2013). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In N. Srinivasan, & P. Chandrasekhar (Eds.), *Progress in brain research: Vol. 202. Decision making: Neural and behavioural approaches* (pp. 55–72). Oxford, UK: Elsevier.
- Dillard, A. J., Scherer, L., Ubel, P. A., Smith, D. M., Zikmund-Fisher, B. J., McClure, ... Fagerlin, A. (2013). Breast cancer anxiety's associations with responses to a chemoprevention decision aid. *Social Science and Medicine*, 77, 13–19.
- Frey, R., Hertwig, R., & Rieskamp, J. (2014). Fear shapes information acquisition in decisions from experience. *Cognition*, 132, 90–99.
- Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology*, 38, 129–166. DOI:10.1006/cogp.1998.0710.
- Haase, N., & Betsch, C. (2012). Parents trust other parents: Lay vaccination narratives on the web may create doubt about vaccination safety. *Medical Decision Making*, 32, 645.
- Hertwig, R. (in press). Decisions from experience. In Keren, G., & Wu, G. (Eds.), *Blackwell handbook of decision making*. Oxford, UK: Blackwell. <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-1118468392.html>
- 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., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13, 517–523.
- Hills, T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science*, 21, 1787–1792.
- Hunink, M., Glasziou, P., Siegel, J., Weeks, J., Pliskin, J., Elstein, A., et al. (2001). *Decision making in health and medicine: Integrating evidence and values*. Cambridge, UK: Cambridge University Press.
- 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, 16271–16276.
- Jolie, A. (2013). My medical decision. *The New York Times*, p. A25. Retrieved from <http://www.nytimes.com/2013/05/14/opinion/my-medical-choice.html>
- Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, 117, 286–295.
- Lejarraga, T., Hertwig, R., & Gonzalez, C. (2012). How choice ecology influences search in decisions from experience. *Cognition*, 124, 334–342.
- Loewenstein, G. (2005). Hot–cold empathy gaps and medical decision-making. *Health Psychology*, 24, 49–56.
- Luce, R. D. (1959). *Individual choice behavior*. New York, NY: Wiley.
- Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description–experience gap in risky choice limited to rare events? *PLoS ONE*, 6, e20262.
- Morens, D. M. (1999). Death of a president. *The New England Journal of Medicine*, 341, 1845–1849.
- Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin and Review*, 14, 1133–1139.
- Nilsson, H., Rieskamp, J., & Wagenmakers, E.-J. (2011). Hierarchical Bayesian parameter estimation for cumulative prospect theory. *Journal of Mathematical Psychology*, 55, 84–93.
- Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making*, 26, 260–271.
- Pachur, T., Hertwig, R., & Steinmann, F. (2012). How do people judge risks: Availability heuristic, affect heuristic, or both? *Journal of Experimental Psychology: Applied*, 18, 314–330.

- Pachur, T., Hertwig, R., & Wolkewitz, R. (2014). The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision, 1*, 64–78.
- Pachur, T., & Scheibehenne, B. (2012). Constructing preference from experience: The endowment effect reflected in external information search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1108–1116.
- Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., & Lee, M. (2000). Comparison of vignettes, standardized patients, and chart abstraction. A prospective validation study of 3 methods for measuring quality. *The Journal of the American Medical Association, 283*, 1715–1722.
- Rush, B. (1805). A defence of blood-letting, as a remedy for certain diseases. In *Medical inquiries and observations*, Vol. 4 (pp. 276–361). Philadelphia, PA: J. Conrad & Co. Retrieved from <http://resource.nlm.nih.gov/2569003RX4>
- Russell, L. B., & Schwartz, A. (2012). Looking at patients' choices through the lens of expected utility: A critique and research agenda. *Medical Decision Making, 32*, 527–531.
- Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association, 46*, 55–67.
- Scheibehenne, B., & Pachur, T. (2015). Using Bayesian hierarchical parameter estimation to assess the generalizability of cognitive models of choice. *Psychonomic Bulletin and Review, 22*, 391–407.
- Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Cambridge, MA: Harvard University Press.
- Sunstein, C. R. (2002). Probability neglect: Emotions, worst cases, and law. *Yale Law Journal, 112*, 61–107.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.
- Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science, 20*, 473–479.
- Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. M. (2007). Reducing aversion to side effects in preventive medical treatment decisions. *Journal of Experimental Psychology: Applied, 13*, 11–21.
- Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. (2009). Explanations for side-effect aversion in preventive medical treatment decisions. *Health Psychology, 28*, 201–209.
- Weber, E. U., Shafir, S., & Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 111*, 430–445.
- Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin, 139*, 497–518.

Authors' biographies:

Tomás Lejarraga is a researcher at the Center for Adaptive Rationality at the Max Planck Institute for Human Development in Berlin. He studies how individuals and groups make judgments and decisions from experience.

Thorsten Pachur is a senior researcher at the Center for Adaptive Rationality at the Max Planck Institute for Human Development in Berlin. He studies learning and memory processes in decision-making, strategy selection, and the psychology of risky choice, using computational modelling and process analyses.

Renato Frey is a researcher at the Center for Cognitive and Decision Sciences at the University of Basel and an adjunct researcher at the Max Planck Institute for Human Development in Berlin. He studies risk taking and in particular experience-based decision-making using computational models of cognition.

Ralph Hertwig is the director of the Center for Adaptive Rationality at the Max Planck Institute for Human Development in Berlin. He held positions at the University of Chicago and Columbia University, and served as the chair for Cognitive and Decision Science and Dean of Research at the University of Basel.

Authors' address:

Tomás Lejarraga, Center for Adaptive Rationality, Max Planck Institute for Human Development. Lentzeallee 94, 14195, Berlin, Germany.

Thorsten Pachur, Center for Adaptive Rationality, Max Planck Institute for Human Development. Lentzeallee 94, 14195, Berlin, Germany.

Renato Frey, Fakultät für Psychologie, Missionsstrasse 64A, CH-4055, Basel, Switzerland.

Ralph Hertwig, Center for Adaptive Rationality, Max Planck Institute for Human Development. Lentzeallee 94, 14195, Berlin, Germany.