8

Decisions from Experience

Ralph Hertwig

Center for Adaptive Rationality (ARC), Max Planck Institute for Human Development, Germany

Introduction

Ward Edwards, the father of behavioral decision theory, demanded a great deal of his experimental participants. The title of one of his articles says it all: "Probability learning in 1000 trials" (Edwards, 1961). Individuals were asked to make 1,000 consecutive predictions about which of two events would occur, and were told after each prediction whether or not they had been correct. The high number of trials permitted Edwards to analyze behavior after learning was completed. In other investigations he proved equally rigorous. He and his colleagues were among the first to test experimentally whether human inference follows Bayes's theorem. Phillips and Edwards (1966, Experiment 1), for instance, asked each participant to make a total of 480 posterior probability estimates in light of ever-new evidence. Edwards's (1968) conclusion was that human inferences, although "conservative" (beliefs were revised less thoroughly than prescribed by Bayes's theorem), were usually proportional to the normatively correct values.

Edwards's protocol of experimentation can be contrasted with the approach that emerged in the early 1970s, with the ascent of the heuristics-and-biases research program (Tversky & Kahneman, 1974). Let us take, for illustration, investigations of Bayesian reasoning. In the engineer–lawyer problem, Kahneman and Tversky (1973) presented their participants with five written thumbnail descriptions of fictitious people, supposedly drawn at random from a population of 70 lawyers and 30 engineers. For each person described, participants estimated the probability that he or she was one of the 70 lawyers (or, in a condition with reversed base rates, one of the 30 lawyers). Learning was neither necessary nor possible. A slew of subsequent studies used this problem or variants of it, frequently providing participants with all relevant pieces of information (base rate, hit rate, and false alarm rate) that are entered into the standard probability format of Bayes's theorem.

The Wiley Blackwell Handbook of Judgment and Decision Making, First Edition.
Edited by Gideon Keren and George Wu.

Having made *inferences from experience* in Edwards's studies, people now made *inferences from description*. The former experienced a long sequence of new data (draws of a chip from a selected urn) and continuously revised their posterior estimates in light of that information. The latter responded once to the information displayed in front of them. Drastically different findings and conclusions resulted. For Edwards (1968), the human mind was Bayesian, albeit conservative Bayesian. For Kahneman and Tversky (1972), in contrast, Bayes's theorem failed entirely to describe the workings of the mind: People ignore base rates, and thus "in his evaluation of evidence, man is apparently not a conservative Bayesian: he is not Bayesian at all" (p. 450).

There may be more than one reason for Edwards and colleagues diagnosing conservatism where Kahneman and Tversky discern base-rate neglect. For instance, their experimental protocols differed on several dimensions. Vernon Smith (2001, p. 428) called Edwards one of the founders of experimental economics, and Edwards's protocol indeed foreshadowed economists' contemporary ideal of experimentation (Hertwig & Ortmann, 2001). There is, however, another potential explanation for the puzzling Gestalt switch from conservatism to base-rate neglect, namely that a finding referred to as the *description–experience gap* is not limited to choice (Hertwig & Erev, 2009) but may generalize, for instance, to probabilistic inference. Since the early 2000s, this gap has inspired numerous investigations, the key issue being to what extent choices and judgments change when people draw on online experience of the structure of the probabilistic world rather than simply being told about it, no learning required. In what follows, I will introduce this description–experience gap, its possible causes, and models used to capture experience-based choices.

The Description–Experience Gap in Risky Choice

If the Wason selection task is, as psychologists' folklore has it, the most studied "fruit fly" in cognitive psychology, then choice between monetary lotteries must be a close second. In theory, this *Drosophila melanogaster* can be studied from many angles. In practice, many researchers have grown accustomed to relying on a single approach (see Pleskac & Hertwig, 2014; Weber, Shafir, & Blais, 2004): lotteries in which the outcomes and their probabilities are explicitly stated (either numerically or visually in terms of, e.g., pie charts), and respondents thus make *decisions from description* (Hertwig, Barron, Weber, & Erev, 2004). This fruit fly – fully described lotteries – has populated both economists' and psychologists' laboratories. For instance, one of the most famous violations of EU theory, the Allais Paradox, involves choices between explicitly stated outcomes and probabilities (Allais, 1953, p. 514). Similarly, in his informal experiment designed to illustrate ambiguity aversion, another violation of EU theory, Ellsberg (1961, p. 650) relied on a setting with stated outcomes and probabilities (except in the ambiguous urn, in which probabilities were left unspecified). No learning necessary. The same held for Kahneman and Tversky's (1979) numerous demonstrations of violations of EU, including the reflection effect, the possibility effect, the common-consequence effect, and the common-ratio effect.

There are, no doubt, real-world analogs of such convenient summary descriptions of options. Newspaper weather forecasts now commonly present outcomes and quantitative probabilities (e.g., of precipitation; Gigerenzer, Hertwig, Van den Broek, Fasolo, & Katsikopoulos, 2005); similarly, actuarial tables, mutual-fund brochures, and package-inserts offer descriptions of possible outcomes and probabilities. Yet there is also a world devoid of descriptions. Many of our behaviors – falling in love, job interviews, marital arguments, crossing the street – come without a package-insert detailing possible outcomes and their probabilities. In his famous distinction, Knight (1921) delineated the world of *risk* from the world of *uncertainty*. In the former, our actions lead to sets of possible outcomes, each occurring with a *known* probability; in the latter, probabilities are unknown, and we have no choice but to navigate the "twilight ... of probability," as John Locke (1690/1959) put it. Although the modern world of risk has made notable inroads into this descriptionless territory – for instance, by recording and tabulating our social behavior in greater detail than Quetelet (1842/1969) could ever have imagined – many things still remain hidden in the gloom of uncertainty.

We do, however, have torchlights. Many uncertain environments permit us to benefit from the vicarious experience of others or to bring to bear relevant experience and knowledge of similar situations stored in memory (Fox & Tversky, 1998; Tversky & Fox, 1995). Sometimes they allow us to gain hands-on experience of the initially unknown payoff distributions. We have referred to people's drawing on experience of payoff distributions and making choices on this basis as *decisions from experience* (Hertwig et al., 2004; Weber et al., 2004). Studying such decisions from experience is, of course, nothing new. As described above, Edwards and other godfathers of modern decision science did exactly that (see Luce & Suppes, 1965, for a review). Although, starting in the 1970s, behavioral-decision researchers began to turn away from the transients of learning (with some notable exceptions; e.g., Busemeyer, 1985), concerns for learning and experience-based decision making remained alive in areas such as economics (see the literature on reinforcement learning in experimental games; e.g., Erev & Roth, 1998) and operation research (see the literature on multiarmed bandit problems; Sutton & Barto, 1998). What is novel, however, is that since the early 2000s various researchers have systematically pitted decisions from experience against decisions from description, commonly using monetary lotteries. Their investigations have revealed a systematic and robust difference between the two kinds of decisions, and this description–experience gap has rekindled interest in decisions from experience (for an early review, see Rakow & Newell, 2010).

Before turning to this gap, let me introduce the simple tool employed to capture decisions from experience in the domain of monetary lotteries. The tool is a "computerized money machine." Participants see two buttons on a computer screen, each representing an initially unknown payoff distribution. Clicking a button results in a random draw from the specified distribution. Three variations of this experimental tool have commonly been employed (but hybrids, such as combinations involving both descriptions and experience can also be constructed; see, e.g., Abdellaoui, L'Haridon, & Paraschiv, 2011; Erev, Glozman, & Hertwig, 2008; Jessup, Bishara, & Busemeyer, 2008; Ludvig & Spetch, 2011). In the *sampling paradigm* (e.g., Hertwig et al., 2004; Weber et al., 2004) participants first sample as many outcomes as they like and only then decide from which distribution to make a single draw for real. In the *full-feedback paradigm* (e.g., Yechiam & Busemeyer, 2006) each draw contributes to participants' earnings, and they receive draw-by-draw feedback on their obtained and forgone payoffs (i.e., the payoff they would have received had the other option been selected). The *partial-feedback paradigm* (e.g., Barron & Erev, 2003; Erev & Barron, 2005) is identical to the full-feedback paradigm, except that participants receive draw-by-draw feedback only on the obtained payoffs. All three paradigms, along with an instance of the "common fruit fly" (description-based choice), are depicted in Figure 8.1.

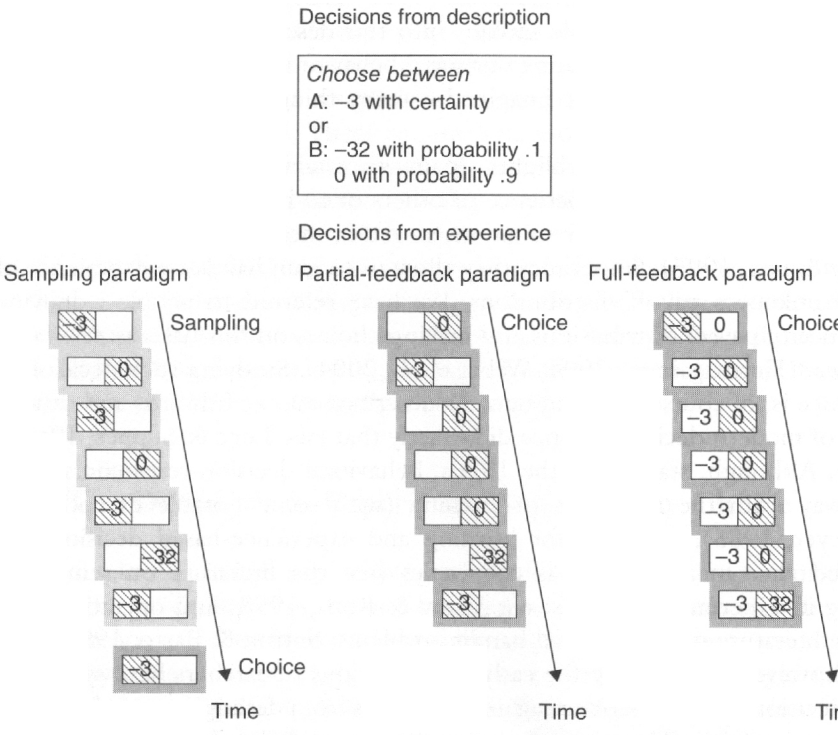
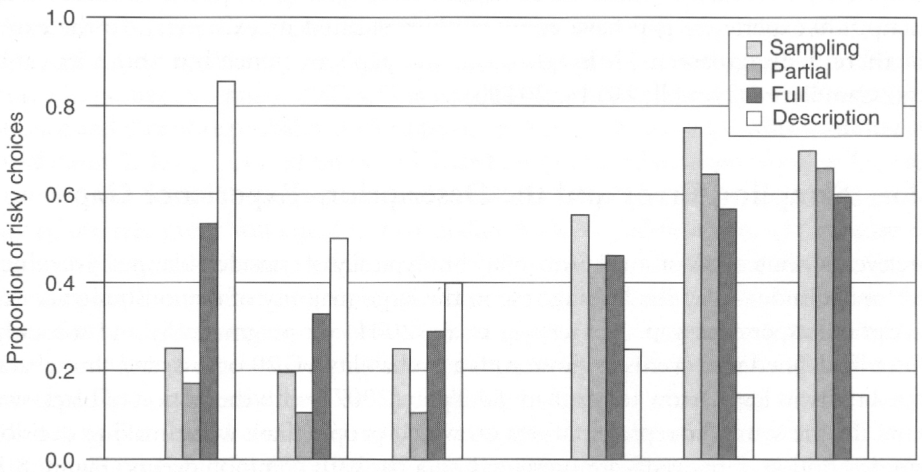


Figure 8.1 How to study decisions from description and experience? The choice task in decisions from description (upper panel) often consists of two lotteries with explicitly stated outcomes and probabilities. In research on decisions from experience (lower panel), three paradigms (and hybrids thereof) have been employed: The *sampling paradigm* includes an initial sampling stage (represented by seven fictitious draws) during which the participant explores two payoff distributions by clicking on one of two buttons on a computer screen (light gray screen). After terminating sampling, the participant sees a choice screen (here shown in dark gray) and is asked to draw once for real. The buttons chosen during sampling (exploration) and choice (exploitation) are hatched diagonally. The *partial-feedback paradigm* merges sampling and choice, and each draw simultaneously represents exploration and exploitation. The participant receives feedback on the obtained payoff after each draw (hatched box). The *full-feedback paradigm* additionally reveals the forgone payoff (i.e., the payoff that the participant would have received had he or she chosen the other option; white box). Source: Hertwig and Erev (2009). Reproduced with permission of Elsevier.

The key difference between the paradigms is the degree to which they entail an exploration–exploitation trade-off (Sutton & Barto, 1998). Exploration and exploitation are two goals associated with every choice, namely, to obtain a desired outcome (exploitation) or to gather new information about other, perhaps even better, actions (exploration). In the partial-feedback paradigm, each draw from the payoff distributions contributes to the participant's earnings (or losses); therefore, a balance needs to be struck between the simultaneous demands of exploration and exploitation. The sampling paradigm suspends this trade-off by the temporal separation of exploration and exploitation, akin to attending a free wine-tasting fair, perusing the Gault–Millau guide to select a restaurant, or checking online traffic cams before leaving home. Exploitation – for example, dining at one of the acclaimed gastronomic temples – takes place only when exploration has been terminated. The full-feedback paradigm also removes the exploration–exploitation trade-off by permitting people to exploit and to receive information about the foregone payoff at the same time.

Notwithstanding these differences, all three experiential paradigms have resulted in similar choice patterns – and in choices systematically different from decisions from description. Figure 8.2 illustrates this divergence using six decision problems employed in a choice-prediction competition reported in Erev et al. (2010). Each problem



Problem	1	2	3	4	5	6
Safe option	7	–4.1	11.5	2.2	6.8	11
Risky option						
Maximum	16.5	1.3	25.6	3	7.3	11.4
P(maximum)	0.01	0.05	0.10	0.93	0.96	0.97
Minimum	6.9	–4.3	8.1	–7.2	–8.5	1.9
Expected value	7	–4.0	9.9	2.3	6.7	11.1

Figure 8.2 The description–experience gap. Proportion of choices of the risky option as a function of the probability of the more desirable outcome in 6 of 120 problems studied in Erev et al.'s (2010) choice-prediction competition. Each presents a choice between a risky option and a safe option. The decision problems and the expected values of the risky options are displayed below the graph. Each problem was studied using the four paradigms displayed in Figure 1. Source: Hertwig and Erev (2009). Reproduced with permission of Elsevier.

involves a choice between a risky option with two outcomes and a safe option. In the risky options, either the desirable outcome or the less desirable outcome occurs with low probability (≤0.1). In the experiential paradigms, respondents tended to select the risky option when the desirable outcome occurred with high probability but chose the safe option when the desirable outcome occurred with low probability. In decisions from description, this pattern was reversed. The difference in choices can be summarized as follows: in decisions from experience, people behave as if rare events have less impact than they deserve according to their objective probabilities, whereas in decisions from description, people behave as if rare events have more impact than they deserve. Let me emphasize that the gap occurs on the level of choice and can be quantified, for instance, in terms of the mean absolute deviation in choice proportions. It can also be characterized in terms of predictable reversals of choice such as a reversed fourfold pattern of risk attitudes (Hertwig, 2012), a reversed reflection effect (Ludvig & Spetch, 2011), and a reversed common ratio and certainty effect (Barron & Erey, 2003). To what extent the description–experience gap rests on systematically different probability weighting (i.e., over- vs. underweighting of rare events) is open to debate – I will return to this question later.

Once this description–experience gap was discovered the search for its causes began. I will next review possible causes, focusing on the sampling paradigm, which has been used in most previous work. Studies investigating the potential causes of the description–experience gap have consistently replicated its existence. To the extent that there is disagreement, it is not about the gap's existence but about its causes (e.g., Camilleri & Newell, 2011a, 2011b).

Sampling Error and the Description–Experience Gap

Rare events – outcomes of small probability but typically of considerable (positive or negative) magnitude – play the leading role in the large majority of demonstrations of the description–experience gap. In Hertwig et al. (2004) we pragmatically and somewhat arbitrarily defined rare events as those with a probability of .20 or less, and thus characterized rarity in less narrow terms than did Taleb (2007), with the notion of black-swan events. In the symbolic representations on which people draw when making decisions from description, rare events are presented on a par with common events (Figure 8.1). Thus, a person cannot easily escape the reality and sway of rare events. In decisions from experience, in contrast, various factors can collude to "hide" rare events or at least make them less impactful. The factor that has received most attention is sample size.

Frugal sampling and sampling error

The sampling paradigm (Figure 8.1) permits decision makers to explore the parameters of the payoff distributions as thoroughly as they wish and then to exploit them. However, people's explorative aspirations appear modest. In Hertwig et al. (2004), for instance, we observed that respondents' typical number of draws was about seven from each payoff distribution, with a median of 15 draws in total. Each draw only takes a few seconds. Yet such limited exploration proved to be the rule rather than the exception. Subsequent studies observed similarly frugal sampling, with occasional exceptions (e.g., Study 2 in Hau, Pleskac, Kiefer, & Hertwig, 2008). Specifically, a recent meta-analysis of a total of 21 sampling-paradigm data sets (involving more than 1,000 participants and over 10,000 choices) found a median of 16 draws (Wulff, Hertwig, & Mergenthaler, 2015). Such frugal sampling exacts a price: the less exploration there is, the larger the probability that a person will fail to experience rare events and remain ignorant of their existence. Indeed, in Hertwig et al. (2004), rare events were not encountered in 44% of all sampling sequences.

Even if rare events are experienced, small samples make it more probable that they will be encountered less frequently than expected given their objective probability. Averaged across problems, Hertwig et al. (2004) observed that 78% of respondents encountered the rare event less frequently than expected (i.e., fewer than np times), whereas just 22% of respondents encountered the rare event as or more frequently than expected. The reason for the "underrepresentation" of rare events in small samples is that the binomial distribution for the number of times a particular outcome will be observed in n independent trials is skewed when p is small (i.e., the event is rare) and n is small (i.e., few outcomes are sampled). In such distributions, a respondent is more likely to encounter the rare event less frequently than expected (np) than more frequently than expected (Hertwig, 2012). For illustration purposes, suppose 1,000 people sample from a distribution in which a critical event has a probability of .1 and estimate this probability on the basis of the probability experienced in the sample. Each person samples 20 times. Yet, of the 1,000 people, only about a quarter (285) will experience the critical event twice and gauge its probability accurately. Nearly two fifths (392) will never observe the critical event or will observe it just once and therefore probably underestimate p. About a third (323) will encounter the crucial event 3, 4, 5, …, or 20 times and, based on this sampling experience, will overestimate its probability. Of course, averaged across all 1,000 people, the estimated probability of the rare event will equal its probability in the population (i.e., .1) because the sample proportion is an unbiased estimator of the proportion in the population. Yet, for small samples, the number of people who experience the rare event less frequently than expected exceeds the number of people who experience it more frequently than expected.

People's modest exploration suggests a straightforward explanation for the description–experience gap: sampling error. Experience-based choices may be at variance with description-based choices simply because they operate on the basis of different quantitative probabilities ("statistical" or sample probabilities vs. "a priori probabilities"; Knight, 1921). Indeed, Fox and Hadar (2006) and Rakow, Demes, and Newell (2008) have suggested that the sole culprit behind the description–experience gap is sampling error. But even if this were the case, a "mere" statistical cause behind the description–experience gap would not detract from its psychological significance. In many real-world environments, people may – for reasons such as lack of time and rarity of the event class – have no choice but to rely on small samples.

Is sampling error the sole cause of the description–experience gap?

Frugal sampling and the ensuing sampling error is a major determinant of the description–experience gap, but is it its sole cause? Various approaches have been taken to tackle this question. One has been to find out whether more sampling (and less error) reduces or even eliminates the gap. Hau et al. (2008) boosted sampling by raising the stakes (Study 2: median of 33 draws) and, alternatively, directing participants to sample up to a fixed sample size (Study 3: 100 draws). Larger samples decreased the description–experience gap from a baseline of 27 (Study 1) to 13 (Study 2) and 17 (Study 3) percentage points, respectively; however, the gap was not eliminated (see also Camilleri & Newell's, 2011b, sampling paradigm involving 100 draws). Another approach has been to turn one person's experienced sample probabilities into another person's described probabilities. Thus matched, experience- and description-based decisions concern the exact same probabilities and outcomes. In the first implementation of this yoked design, Rakow et al. (2008) found no gap. Subsequently, Hau, Pleskac, and Hertwig (2010) found that the gap indeed ***disappears*** with small sample sizes but *reappears* with large sample sizes. The explanation is that with small samples decisions from description can become trivial. Take, for instance, the choice between 32 with a probability of .1 (and 0 otherwise; option A) or 3 with certainty (option B). Sampling modestly from these two options may result in a sequence of 0, 0, 0, 0, 0, 0 (option A) and 3, 3, 3, 3, 3, 3 (option B). Translating these sequences into explicitly stated lotteries results in a trivial choice between "100% chance to win 0" and "100% to win 3." It is thus no surprise that one person's (yoked) description-based choice agrees with another's experience-based choice (when small samples cause yoked decisions to become trivial).

Camilleri and Newell (2011a) suggested a third approach to investigate the effect of sampling error. They "matched" experienced to objective outcome distributions by focusing on those small sets of trials in which participants' experienced distribution was within ±10% of the objective distribution. In those rare trials, experience and description resulted in very similar choices (yet, as the authors discussed, the small sets of trials were "not representative across participants and problems," p. 281). In a second study, Camilleri and Newell had people sample "blocks of trials" in which each block represented a random order of outcomes that perfectly matched the described probabilities (e.g., a rare event with a probability of .2 occurred twice in a block of 10 outcomes). Again, choice proportions in the experience and description condition were very similar, and the authors concluded that the description–experience gap is "largely the result of nonequivalent information at the point of choice" (p. 276). Ungemach, Chater, and Stewart (2009) also strove for equivalent information in experience and description. They had people sample 80 times, and they devised experienced probabilities that matched the stated probabilities seen by other respondents who made decisions from description. Unlike Camilleri and Newell, they observed a substantial gap.

It seems that the evidence on the role of sampling error is suggestive but not conclusive. Boosting sampling size, yoking description and experience (in nontrivial choices), and making them equivalent by other means has not removed the description–experience gap, with the notable exception of Camilleri and Newell's (2011a) studies. Regardless of how this debate develops, at least two important findings make it plain that sampling error is not a sine qua non for the gap. First, as illustrated in Figure 8.2 (see also Camilleri & Newell, 2011b), the gap also emerges after 100 trials with partial or full feedback. Second, Ludvig and Spetch (2011) demonstrated that the gap is not limited to options involving rare outcomes (and, by extension, sampling error).

Using hybrids of the sampling and partial-feedback paradigms they gave respondents a choice between a risky payoff distribution with equiprobable outcomes (e.g., 40 with .5, and 0 with .5) and a safe option (20 with certainty). In experience-based choices, people gambled more in the gain domain than in the loss domain, a reversal of the reflection effect that Ludvig and Spetch – and many others before them – observed in description-based choices (e.g., Kahneman & Tversky, 1979).

Recency and sampling error

A discussion of the role of sampling error is not complete without mention of the *recency* effect. Even if people enjoy perfect equivalence of experienced and described information at the moment of choice, recency may sneak (internal) sampling error in through the mnemonic backdoor, thus destroying equivalence. Recency refers to the phenomenon that observations made late in a sequence may receive more weight than they deserve ($\geq 1/n$). Frequently observed in memory, belief updating, and judgments (e.g., Hogarth & Einhorn, 1992), recency reduces a large sample from the environment to a smaller mnemonic sample of recent events, thus reintroducing the aforementioned skew into large samples of experience. In research on decisions from experience, the evidence on recency is mixed, with some studies finding traces of recency (e.g., Camilleri & Newell, 2011a, Experiment 1, random group; Camilleri & Newell, 2011b; Hertwig et al., 2004) and others not (Camilleri & Newell, 2011b, Experiment 2; Hau et al., 2010; Rakow et al., 2008; Ungemach et al., 2009). Note, however, that two of the most successful models of experience-based choices, the explorative sampler with recency (Erev et al., 2010) and instance-based learning theory (IBLT; Gonzalez & Dutt, 2011) include a recency parameter.

To conclude, the debate concerning the role of sampling error in the description–experience gap has been important. Nevertheless, several other factors have also been identified as contributing to the gap. In my view, these factors are at least as captivating as the role of sampling error. One concerns the way people explore options in the world.

Search Policies and the Description–Experience Gap

An explorer can adopt one of two idealized ways of exploring payoff distributions. She can oscillate between payoff distributions (e.g., options A and B), each time drawing the smallest possible sample (piecewise sampling; Figure 8.3a). Alternatively, she can sample extensively from one distribution and then extensively from the other (comprehensive sampling). Taking these two sampling strategies as a starting point – many empirical strategies will, of course, fall on the continuum between them – Hills and Hertwig (2010) hypothesized that the way people explore foreshadows their decisions. Specifically, the choices of individuals who tend to sample piecewise are more likely to be consistent with a roundwise decision strategy: they determine which option yields better rewards in each round of sampling and ultimately choose the one that wins the most rounds (Figure 8.3b). In contrast, people who sample comprehensively from one distribution before advancing to the other are more likely to make decisions

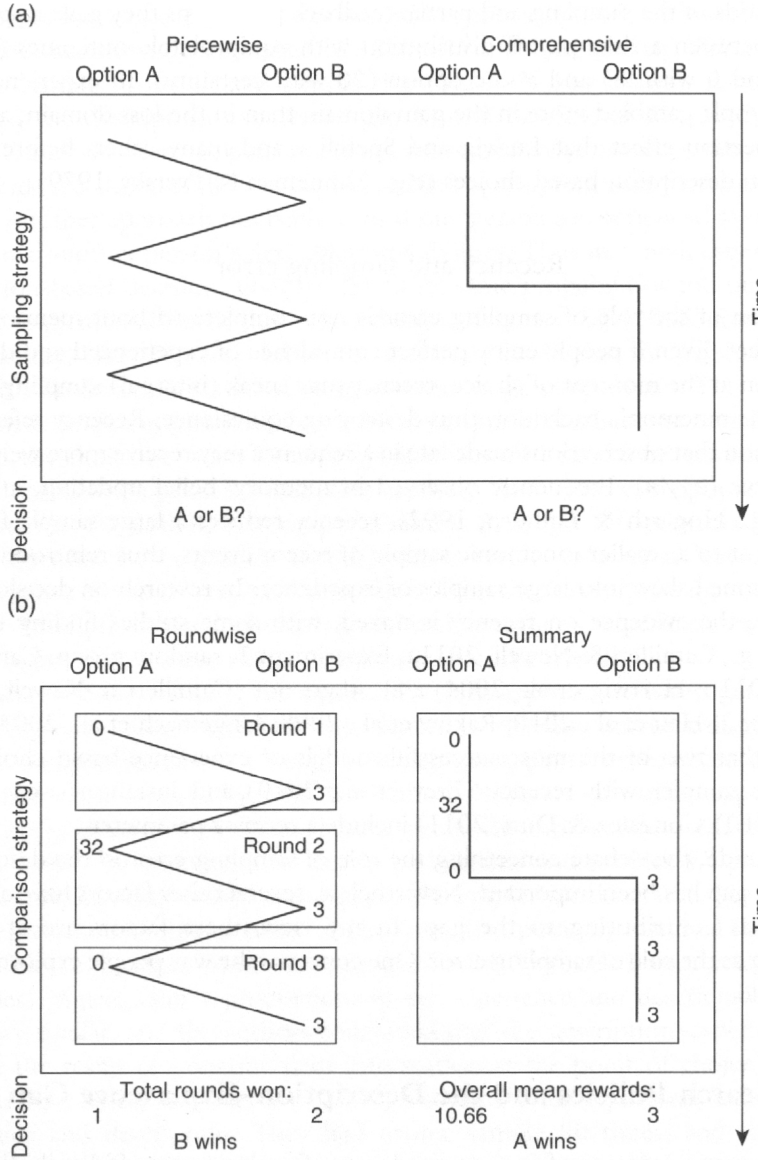


Figure 8.3 The coupling of sampling and decision strategies. Two idealized sampling strategies (a) and correlated decision strategies (b). A piecewise sampling strategy alternates back and forth between payoff distributions, whereas a comprehensive sampling strategy takes one large sample from each distribution in turn. Following sampling, participants make a decision about which distribution they prefer. A roundwise decision strategy compares outcomes (gains and losses) over repeated rounds and chooses the distribution that yields higher rewards in most of the rounds. A summary decision strategy calculates the mean reward per distribution and chooses the option with the higher value. Source: Hills and Hertwig (2010). Reproduced with permission of SAGE Publications.

consistent with a summary decision strategy: they evaluate the average reward and then choose the distribution whose reward promises to be higher. Search and decisions may be dependent because the two different ways of exploring the world facilitate comparisons across different scales of information (i.e., rounds relative to summaries of rewards). In the most extreme case, as shown in Figure 8.3b, summary and roundwise decision strategies will lead to different choices, even though they operate on the same information.

Such a coupling between search policy and choice strategy can also contribute to the description–experience gap. If frequent oscillation between options co-occurs with a roundwise decision strategy then this strategy will weigh each round equally, ignore the magnitude of wins and losses, and ultimately act as if it underweights rare outcomes. The strategy that chooses on the basis of the average reward, in contrast, will be less inclined to undersell rare events (for its treatment of rare events, only sampling error matters). Reanalyzing a large set of choices taken from past studies, Hills and Hertwig (2010) indeed found a link between search behavior and the magnitude of the gap (Figure 8.4). Specifically, people who switched frequently between distributions (piecewise sampling) and were thus likely to be using a roundwise strategy made more choices consistent with giving less weight to rare events relative to infrequent switchers and, even more so, relative to people responding to descriptions.

Apart from sampling error and search policy, other factors contribute to the description–experience gap. Two of them concern our theoretical conceptions of the processes behind experience-based versus description-based choice. Before turning to

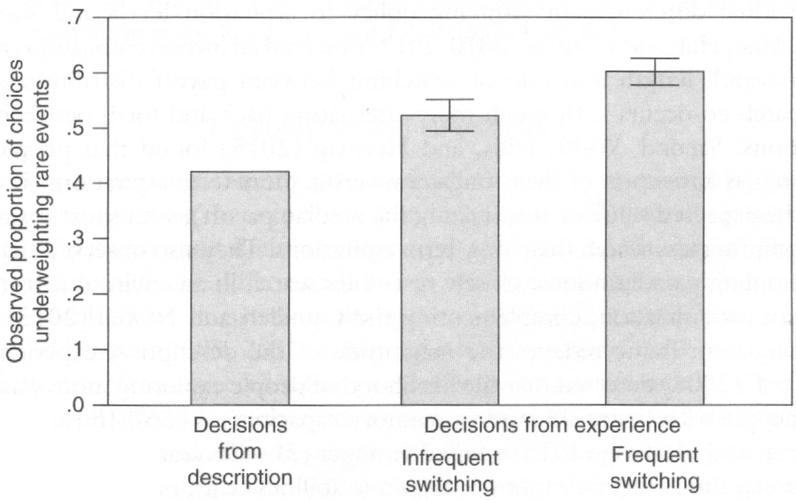


Figure 8.4 Exploration policy and the description–experience gap. Observed proportions of choices consistent with rare events receiving less impact than they deserve (relative to their objective probability) among infrequent switchers (comprehensive sampling), frequent switchers (piecewise sampling), and in the corresponding decisions from description (for details, see Hills & Hertwig, 2010). Error bars represent standard errors of the mean. Source: Hills and Hertwig (2010). Reproduced with permission of SAGE Publications.

these factors, however, I will highlight a methodological benefit that researchers of decisions from experience enjoy, and briefly review some of the insights it has already facilitated.

The Anatomy of Search in the Sampling Paradigm

Decisions from experience are a worthy object of investigation far beyond the description–experience gap – among other reasons, because they offer a great methodological advantage relative to decisions from description. Specifically, the experimental paradigms (Figure 8.1) lay open what is otherwise difficult to observe: people's search for information. As a result, Hills and Hertwig (2010) have been able to investigate how choices hinge on search policies, and Pachur and Scheibehenne (2012), how predecisional patterns of information search give rise to the endowment effect. Thanks to this transparency, researchers have possibly learned, in just a few years, more about the cognitive and ecological factors shaping search in decisions from experience than in decades of study of decisions from description, where resourceful methods are needed (Schulte-Mecklenbeck, Kühberger, & Ranyard, 2011) to make observable what little is left of search when all the necessary information is displayed in front of people. Here are some of the things we have learned so far.

Variations in explorative efforts

Although average exploration efforts are modest, there is nevertheless variability in people's inclination to search. Several factors appear to be in play here, ranging from interindividual differences in sampling policy to aspiration levels and short-term memory. First, Hills and Hertwig (2010, 2012) observed an inverse correlation ($r = -.44$) between search length and rate of switching between payoff distributions: more frugal search co-occurs with much more alternating back and forth between payoff distributions. Second, Wulff, Hills, and Hertwig (2015) found that people adjust their search as a function of their long-term versus short-term aspirations (i.e., maximizing the expected value vs. maximizing the median payoff), with short-term aspirations prompting less search than long-term aspirations. They also observed that search in the sampling paradigm most closely resembles search in an environment inducing short-term maximization, complementing the Camilleri and Newell (2013) finding that a long-term frame reduces the magnitude of the description–experience gap. Rakow et al. (2008) detected that the likelihood of people exploring more was greater among people with larger short-term memory capacity ($r = .36$). In the same vein, Frey, Mata, and Hertwig (2015) studied younger ($M = 24$ years) and older ($M = 71$ years) adults, the latter having lower cognitive abilities (e.g., processing speed) than the former, and they found that search became (slightly) more frugal with age, possibly because of declining working memory capacity (see also Spaniol & Wegier, 2012). Finally, explorative efforts are also moderated by numeracy (the ability to comprehend and transform probability numbers; Peters, 2012), with more numerate people tending to draw larger samples. The same holds for people who regard themselves as good rational thinkers (Lejarraga, 2010).

Choice ecology and search

In decisions from experience, people are the masters of their information search to the extent that they control how much to explore and when to exploit. But some control may also reside in the environment and in interactions between the environment and psychological constructs such as loss aversion and risk aversion (see also Chapter 13 of this handbook). Lejarraga, Hertwig, and Gonzalez (2012) analyzed (a subset of) existing data sets to find out how two key ecological properties influence exploration: domain of choice (gains vs. losses) and experienced variance (variance vs. no variance). The prospect of a loss invoked more search than the prospect of a gain, with a relative average increase in search in the range of 25% to 29% (but with substantial interindividual variability). Does this behavior signal loss aversion? In research on description- and experience-based risky choice, this factor's ability to explain choice in monetary gambles has been challenged (see Erev, Ert, & Yechiam, 2008; Yechiam & Hochman, 2013). At the same time, Yechiam and Hochman found that, despite the lack of behavioral loss aversion in experience-based decisions, losses (relative to gains) involve distinct consequences including higher arousal and frontal cortical activation. The (small) effects of search observed by Lejarraga et al. may thus be part of a larger set of effects that emerge when people face the prospect of losses, even in the absence of loss aversion in choice. Like losses, experienced variance (i.e., the experience of more than one outcome in a payoff distribution) also appears to prompt more exploration. With the experience of variance, people sampled 3.7 observations more from the risky than the safe payoff distribution (relative to .5 observations for people who failed to experience variance when exploring the risky distribution). Last but not least, there are indications that people learn the structure of choice problems over time (e.g., that a given ecology includes a safe and a risky distribution and that the risky one offers a high and a low payoff) and that they allocate their exploration accordingly (see also Hadar & Fox, 2009).

Exploration and amplification

Why do people quite consistently rely on small samples in the sampling paradigm? Several factors may cause them to terminate their search early, including opportunity costs, lack of patience, and limits in short-term memory. Another possibility is that small samples amplify the difference between the expected earnings associated with the payoff distributions, thus making the options more distinct and facilitating choice. Hertwig and Pleskac (2008, 2010) described a simple mathematical proof according to which the absolute expected difference between the sample means of two payoff distributions will always be as large as or larger than the expected (or description-based) difference. In a simulated ecology of 1,000 pairs of randomly generated gambles they also quantified this amplification effect. Figure 8.5 plots the expected (and the median) absolute values of the experienced differences as a function of sample size. The straight line is the average difference between payoff distributions (assuming the objective or described parameters). Small samples amplify the difference between gambles. With two draws from each distribution, the average (expected) experienced difference is 1.5 times larger than the description difference. With increasing sample size this difference shrinks, and with 25 draws per distribution, it is nearly zero.

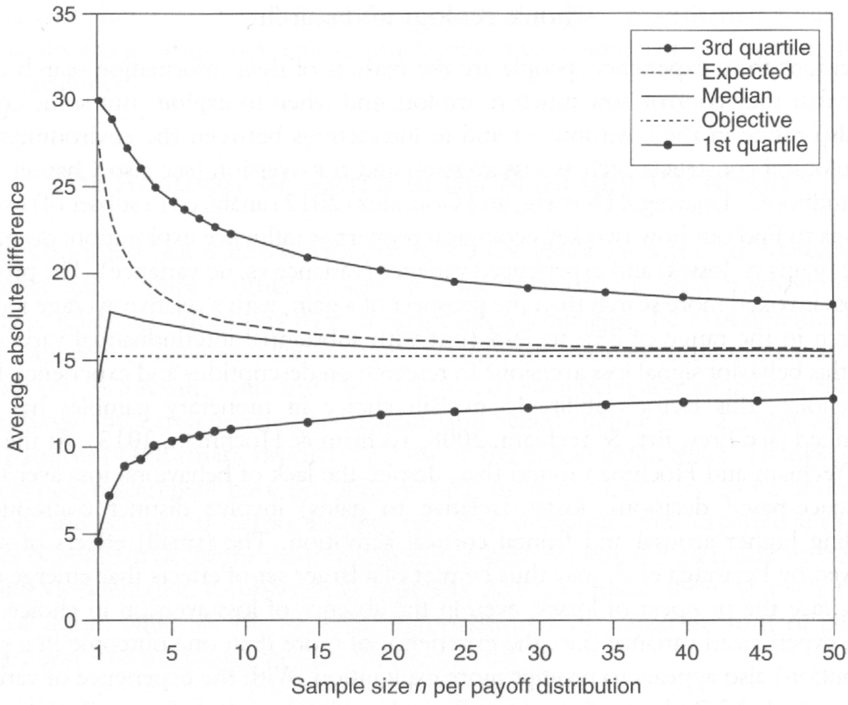


Figure 8.5 How small samples foster discriminability (amplification effect). Experienced differences across 1,000 pairs of gambles as a function of sample size (per payoff distribution). The curves represent (a) the mean of the expected absolute difference, (b) the median of the experienced absolute difference, and (c) the first and third quartiles of the experienced absolute difference. The straight horizontal line represents the average description difference (the objective difference) based on the expected value (15.2) in the simulated ecology. Source: Hertwig and Pleskac (2010). Reproduced with permission of Elsevier.

How costly is it for searchers to enjoy the simultaneous advantages of small search costs, small opportunity costs, and amplified differences? It is surprisingly inexpensive, at least in this choice ecology. With a sample as tiny as one draw, the chance of selecting the better distribution (higher expected value) is as good as 60%; with seven draws, it rises to 81% – after that, additional search offers marginally increasing benefits in accuracy.

To conclude, explorative efforts in the sampling paradigm are shaped by a number of factors. More search is associated with or even prompted by greater short-term (or working) memory capacity, higher numeracy, a comprehensive sampling policy, the prospect of losses, the experience of variance, a long-term frame and, to make this list complete, magnitude of payoff (Hau et al., 2008) and "vigilant" emotions such as fear (relative to happiness; Frey, Hertwig, & Rieskamp, 2014). The robust finding across studies that search is often quite frugal may reflect a smart strategy that harvests accuracy gain while easing choice difficulty (see the empirical evidence in Hertwig & Pleskac, 2010). After this excursion into the cognitive and ecological factors impacting search, let us return to the factors that contribute to the description–experience gap and consider the possibility that description-based and experience-based decisions trigger qualitatively different cognitive processes.

Models of Decisions From Experience

The prevailing premise in many normative and descriptive theories of choice under risk is that people behave as if they multiply some function of probability and value, and then maximize. The value function determines the subjective value of each possible outcome, and it is either an identity function, as in expected value theory, or a nonlinear function reflecting diminishing sensitivities to payoffs. A probability weighting function expresses the weight assigned to each consequence or outcome's value. It is assumed to be either an identity function of the (subjective) probabilities, as in the case of (subjective) EU theory (Savage, 1954; von Neumann & Morgenstern, 1947), or a nonlinear function reflecting the psychological impact of probabilities on choice – as, for instance, in prospect theory (Edwards, 1954; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). These theories have typically been employed in the context of exact, stated probabilities. However, they could also be used to model experience-based choice – and thus situations in which exact probabilities cannot be easily deduced but need to be assessed empirically. When applied to decisions from experience, however, these traditional theories require that a person develop some kind of representation of the probabilities experienced in the sample that are then entered into the theories' functional framework. Some have argued that these theories are capable of accommodating decisions from experience (e.g., Fox & Hadar, 2006; Fox & Tversky, 1998). However, at least two classes of descriptive theories assume that individuals do without any event probabilities in experience-based choice or that the probabilities are "filtered" through the decision maker's cognitive architecture and thus are subject to memory dynamics such as recency and forgetting.

Cognitive heuristics

Choice heuristics can be separated into two classes (Brandstätter, Gigerenzer, & Hertwig, 2006). *Outcome heuristics* draw solely on information about outcomes and ignore probabilities. Examples are the maximax and the minimax heuristics (Savage, 1954), which were originally proposed as models for decisions in situations where no information about probabilities is available. *Dual heuristics*, in contrast, use at least rudimentary probability information. For instance, the *lexicographic heuristic* chooses by first determining the most likely outcome of each payoff distribution and then selecting the distribution whose most likely outcome offers the highest monetary payoff (Payne, Bettman, & Johnson, 1993).

Typically, heuristics for risky choice have been examined in the context of decisions from description (e.g., Brandstätter et al., 2006; Payne et al., 1993). Yet they could, of course, also be recruited in experience-based choice. Hertwig and Pleskac (2010; see also Hau et al., 2008) suggested an outcome heuristic that may describe (some) people's choice in the sampling paradigm. According to the natural-mean heuristic, a person calculates the natural mean of outcomes for both payoff distributions by summing, separately for each one, all experienced outcomes and then dividing the respective sums by the number of outcomes. The person then chooses the payoff distribution with the larger natural mean (i.e., the distribution with the best mean reward in the sampling phase).

This heuristic was originally proposed in the context of *n*-armed bandit problems as a simple method for estimating the values of actions (e.g., pulling one of a slot machine's levers): "The true value of an action is the mean reward received when the action is selected. One natural way to estimate this is by averaging the rewards actually received when the action was selected" (Sutton & Barto, 1998, p. 27). The natural-mean heuristic has two attractive properties: (a) it readily processes sequentially encountered outcomes, and (b) it predicts the same choice as the expected value calculus, assuming that the latter is fed with the experienced outcomes and probabilities in the sample. The decision process, however, is different from that assumed in the calculation of the expected value (i.e., multiplication of all outcomes by their respective probabilities and then summing). Relative to the multiplication process underlying the calculation of the expected value, the heuristic eases the demand on memory and computational capacity, in particular when the payoff distribution involves more than two distinct outcomes.

Associative learning models

Models in this class conceptualize choice as a learning process of behavior–outcome contingencies (e.g., Erev & Barron, 2005; March, 1996; Sutton & Barto, 1998). A good experience following the choice of an alternative boosts the propensity to choose it in the future; a poor outcome diminishes it. Associative learning models have been proposed to explain choice in the sampling paradigm (Erev, Glozman et al., 2008; Hertwig, Barron, Weber, & Erev, 2006; Weber et al., 2004), in the partial-feedback paradigm (e.g., Barron & Erev, 2003; Denrell 2007; Erev & Barron, 2005; Lejarraga, Dutt, & Gonzalez, 2012), and in both paradigms (Gonzalez & Dutt, 2011). Let us briefly consider three representatives of this class of theories: the *value-updating model*, the *explorative sampler model*, and the *instance-based learning model* (the formal framework of each model can be found in the articles referenced).

Value-updating model. This model (Hertwig et al., 2006) assumes that a person engaged in exploration in the sampling paradigm updates the estimated value of a payoff distribution after each new draw from it. The value is the weighted average of the previously estimated value and the value of the most recent outcome. The weight accorded to the most recently drawn outcome can vary and depends on whether the decision maker's memory is subject to recency, primacy, or neither. Once sampling is terminated, the person selects the payoff distribution with the highest value for the final incentivized draw. In a recent model competition involving 70 participants tendering choices in 84 decision problems (no previous study has had respondents make more sampling-based decisions from experience), Frey et al. (2015) found this model, among a set of six learning models, to be the runner-up model for the choices of older participants.

Explorative sampler model. This model, which aims to predict choice in the partial-feedback paradigm, rests on three assumptions (Erev, Ert, & Yechiam, 2008). First, people either explore or exploit. At the outset, the probability of exploration is 1.0. With experience (past outcomes), this probability will decrease, and the smaller the expected number of trials in the sequence the faster this decrease will be. Second, once exploitation begins, a person retrieves from memory a sample of past experiences with each distribution. Third, based on this memory sample, the person estimates the average value of each alternative in a given trial and then selects the alternative with the highest estimated value. The explorative sampler has been shown to outperform reinforcement models in the partial-feedback paradigm, and it was the best performing baseline model in a large-scale model comparison (Erev, Ert, & Yechiam, 2008; Erev et al., 2010).

The instance-based learning model (IBL). Unlike the previous models, the IBL model (Gonzalez & Dutt, 2011; Lejarraga, Dutt, & Gonzalez, 2012) was designed to account for choice in both the sampling and the partial-feedback paradigms. In some important respects, it builds on the ACT–R cognitive architecture (Anderson & Lebiere, 1998, 2003). The model assumes that a choice is a function of the accumulated (through experience) value (i.e., blended value) for each of the two payoff distributions (options). This *blended* value is a function of a payoff distribution's associated outcomes and the probability of retrieving corresponding instances from memory. The activation of an instance in memory, in turn, corresponds to the frequencies and recency with which corresponding outcomes have been observed, including forgetting in memory (decay). Thus, the IBL model, unlike the learning models reviewed above, assumes probabilities; however, they are retrieval probabilities and reflect regularities both of the environment and of human cognitive architecture.

The IBL model offers a single learning process to account for behavior observable in both the sampling and the partial-feedback paradigm. A particularly interesting observation is that behavior in both paradigms gradually progresses from exploration to exploitation. That is, even when search imposes no immediate costs (sampling paradigm) people will take a similar explorative–exploitative path as manifested in costly search (partial-feedback paradigm): initial exploration that is increasingly replaced by exploitation (i.e., low rate of switching between distributions). Whether this path is indeed isomorphic in both paradigms has been debated (Gonzalez & Dutt, 2012; Hills & Hertwig, 2012). Irrespective of this issue, however, Gonzalez and Dutt's model is a general representation of the cognitive process in experiential choice. The model competes with, and can even outperform, the best models created for each paradigm separately (based on the data collected in the Technion prediction tournament; Erev et al., 2010) – including variants of cumulative prospect theory, the natural-mean heuristic, and the explorative sampler model.

To conclude, animals and humans alike often have no choice but to navigate initially unknown reward distributions. Learning via sequential sampling is one strategy that, if possible, empowers them to escape the gloom of uncertainty. Sometimes this learning is subject to an exploration–exploitation trade-off; sometimes the trade-off is suspended. One key theoretical issue is to what extent people form an explicit representation of the probabilities associated with reward distributions, multiply some function of value and probability, and maximize. Conventional models of risk choice, normative and descriptive alike, require this explicit representation of objective or subjective probabilities and weighting of risk and reward. Based on the existing model competitions (e.g., Erev et al., 2010; Gonzalez & Dutt, 2012), however, I believe experience-based choice is better captured by a sequential-learning process in which each individual forms an impression of the value of the options. Learning models differ in the extent to which this value formation process is subject to memory constraints (e.g., recency, decay, random noise). The IBL model is currently the most versatile account of this learning process, even though it may not always outperform simpler (fewer parameters), specialized choice models and although some conjectures (the isomorphism of costless and costly search; Hills & Hertwig, 2012) have been criticized. It will be exciting, with an ever-growing number of data sets involving decisions from experience and search sequences, to carry on the process of model competition and to further refine this class of choice models, in which learning finally takes the central stage. Eventually, the goal will be to develop models that will predict both search (e.g., sample size) and choice.

Probability Weighting and the Description–Experience Gap

With the enormous impact of prospect theory (Kahneman & Tversky, 1979; see also Chapter 2 of this handbook), the assumption of an inverse S-shaped probability weighting seems to have been accepted as a cut-and-dried fact in many researchers' minds. It is perhaps against this background that the following conjecture, complemented by a call for two different theories of risk, has piqued researchers' interest:

> Differences in choices were consistent with the assumption that in decisions from experience, rare events had less impact than they deserved on the basis of objective probability (and in decisions from description, rare events had more impact than they deserved). (Hertwig et al., 2004, p. 535; see also Weber et al., 2004)

This conclusion was not derived by fitting a weighting function to the data and testing for systematic differences in weighting. It originated from the observable choices themselves. In Hertwig et al. (2004), we selected lotteries such that if rare events were accorded less weight in experience than in description systematically different patterns of choices should result. But we were aware that the terminology of over- and underweighting of small probabilities is an interpretation of people's choices based on the assumptions (a) of some (explicit) representation of probabilities (a premise not readily shared by the learning models discussed before), (b) that the way people weight probabilities deviates from (normative) linear weighting, and (c) of a multiplication calculus. Overweighting and underweighting of small probabilities interprets cognition within a Bernoullian framework of choice (but qualitatively other cognitive processes, heuristic ones, may lie beneath probability weighting; Suter, Pachur, & Hertwig, 2013). We were not at all persuaded by such a framework in the context of decisions from experience, and our subsequent modeling made this clear (e.g., Erev & Barron, 2005; Hertwig et al., 2006; Weber et al., 2004). Yet, with our emphasis on the weighting of rare events, we have, unsurprisingly, nudged the debate in this direction.

It is, of course, pertinent to ask to what extent the robust gap in choices corresponds with differential probability weighting of rare events in experience-based and description-based choice. But let me emphasize that the models described above account for the gap, without assuming any over- or underweighting, and that factors such as memory-order effects (recency, primacy), sampling error, or selective reliance on past experiences (i.e., frequent and recent events of the past) can give rise to choices that are consistent with specific weighting patterns. Probability weighting is thus not (necessarily) the mental process at the origin of the gap, but a paramorphic representation of the process (Hoffman, 1960).

A number of (published) analyses have addressed the issue of differential probability weighting in experience and description. Hau et al. (2008, Figure 7) fitted an optimal set of parameters for cumulative prospect theory's (CPT) weighting function to decisions from experience, and found nearly linear weighting. Using a similar fitting procedure, Ungemach et al. (2009, Figure 1) found that "underweighting of small probabilities is also reflected in the best-fitting parameter values obtained when prospect theory … is applied to the data" (p. 473). Camilleri and Newell (2011b) estimated CPT weighting-function parameters using the actually experienced (in the sample) rather than objective probabilities and concluded that the sampling paradigm produced similar degrees of fit across a wide range of weighting-function parameters; in other words, no clear weighting pattern emerged. In contrast, "the regions with the best fit for the Partial and Full Feedback groups … [imply] underweighting of small probabilities" (p. 381). Coupling description with repeated choice and feedback, Jessup et al. (2008) observed that "feedback drove individuals' decision weights toward objective probability weighting" (p. 1015). Abdellaoui et al. (2011) "measured" CPT's parameters (using the certainty equivalence method and at the level of individual people) and concluded that "decision weights exhibit similar qualitative properties across contexts" (sampling paradigm vs. description), and yet their data "suggest that, for gains at least, the subjective treatment of uncertainty in experience-based and description-based decisions is significantly different" (p. 1879). Specifically, they found a less pronounced overweighting of small probabilities in experienced than in described probabilities.

There are also analyses pitting perceptuo-motor decision making, a form of experience-based decision making in which probabilities are not explicitly stated but can be gauged through sampling, against described lotteries. Wu, Delgado, and Maloney (2009, 2011) found that in choices between lotteries people overweight small probabilities and underweight large probabilities. In an equivalent motor-decision task the reverse weighting pattern emerged, with small probabilities being underweighted. Using a visual decision making task (firing bullets at rectangles of varying widths several hundred times), Glaser, Trommerhäuser, Mamassian, and Maloney (2012) found, on average, linear weighting in the equivalent described lotteries but "marked overweighting of small probabilities and underweighting of large probabilities" (p. 425) in the visual lotteries. Finally, in a very interesting article, Jarvstad, Hahn, Rushton, and Warren (2013) matched the returns and likelihoods of described lotteries to those of two experiential tasks, namely, a perceptuo-motor task (i.e., pointing at and hitting one of two targets on a computer screen) and a mental-arithmetic task (i.e., summing four numbers). The authors found, on average, underweighting of low probabilities for the pointing and arithmetic tasks, but overweighting for the lottery task (as did Frey et al., 2015, when determining the best fit probability weighting function across 84 decision problems). Interestingly, however, this weighting pattern was reversed when they fitted participants' subjective probabilities; these probabilities were extracted from participants' ratings of their ability to hit the targets, which they consistently underestimated.

The heterogeneity of findings and conclusions should not come as a surprise. Researchers have implemented the sampling paradigm in different ways. For instance, Camilleri and Newell (2011b) did not give respondents control over their sampling efforts but required them to draw 100 samples (thus reducing the gap and increasing the likelihood that rare events were experienced multiple times; see also Hau et al., 2008). Abdellaoui et al. (2011) provided participants with an exhaustive list of outcomes (including outcomes they had not necessarily experienced) immediately after sampling, thus examining a mixture of description and experience. Similarly, researchers have used rather diverse approaches to determine the probability weighting parameters (e.g., measurement, best-fitting parameters). In light of this methodological diversity, disparity in results and conclusions is only to be expected. In my view, the final word on probability weighting of rare events has not yet been spoken, but it seems fair to conclude that most results suggest that patterns of probability weighting in experience and from description are not the same, with rare events receiving less weight in experience than in description.

I would like to conclude with one more thought on this issue. In explaining the description–experience gap most effort has been dedicated to investigating what makes experience special. But the starting point for explaining the gap could equally be description – perhaps description is the anomaly. Why, in decisions from description, do people behave as if they overweight rare events? A truly convincing answer to this question is still lacking. Here is a speculation. Hertwig et al. (2006) and Erev et al. (2008) suggest that the *propositional* (symbolic) representation of options in decisions from description – for instance, "32 with probability .1; 0 otherwise" – is not without consequence. The mere mention of a rare event (32) lends it weight (a mere-presentation effect); furthermore, presenting the rare event on a par with the common event (0) channels more equal attention to the two events than is warranted by their actual probabilities. To the extent that attention translates into decision weights, as some research suggests (Weber & Kirsner, 1996), the weights of rare and common events will draw closer together than they should. In contrast, decisions from experience rest on an *analogical* representation. For instance, 10 draws from the option "32 with probability .1; 0 otherwise" can be experienced as 0, 0, 0, 0, 0, 32, 0, 0, 0, 0. In this stream of experience, the frequency of the option's events can be read off directly. Moreover, to the extent that attention is allocated as a function of experienced frequency, the resulting decision weights may veridically reflect the sample probabilities.

Beyond Monetary Gambles and Beyond a Simple Dichotomy

With few exceptions, investigations of the description–experience gap have concerned monetary lotteries. There is, however, no reason to assume that the gap is restricted to this domain. Remember the introductory discussion of research on Bayesian reasoning, and the Gestalt switch from conservatism to base-rate neglect. These starkly different conclusions could at least partly be caused by the difference in representations – one requiring repeated updating of probability estimates (Phillips & Edwards, 1966); the other requiring little to no learning (Kahneman & Tversky, 1973). There is evidence that learning indeed matters to the handling of base rates, and may lie behind a description–experience gap in Bayesian reasoning. In his review of Bayesian reasoning studies, Koehler (1996) concluded that "when base rates are directly experienced through trial-by-trial outcome feedback, their impact on judgments increases" (p. 6) relative to summary statistics (descriptions). Possible reasons he entertained are that indirectly experienced base rates may be better remembered, more easily accessed, more meaningful, or more trustworthy.

In theory, a description–experience gap may emerge in any domain in which outcome and probability information can be conveyed either through summary statistics (descriptions) or through explorative and exploitative sampling (experience). The existence of rare events and frugal sampling may be sufficient conditions for the gap but not necessary ones. Furthermore, the gap may pertain not only to games against nature (monetary lotteries) but also to social games (Hertwig, Hoffrage, & the ABC Research Group, 2013) in which uncertainty arises not from the operation of a dispassionate chance device but from the decisions of other people. Take, for instance, a social game such as the repeated ultimatum game (Avrahami, Güth, Hertwig, Kareev, & Otsubo, 2013). In its simplest form, the ultimatum game involves two people playing a single round in which one player, the proposer, suggests how to split a fixed monetary pie. This split represents a take-it-or-leave-it offer (an ultimatum) that the other player, the responder, must accept or reject. In a repeated ultimatum game, the same two players face off repeatedly or one player is paired with ever-changing opponents. Typically, players learn about others' behavior in terms of descriptions (i.e., summary statistics of, say, median offers and median acceptance thresholds; see Avrahami et al., 2013). Alternatively, however, they could learn about the social payoff distribution by sampling from it, thus learning about specific offers and their acceptance or rejection, respectively. Indeed, in the games we play in the real world we rarely enjoy access to summary statistics of "social risks," instead learning sequentially from our experience and from the experience of others (vicarious experience).

In searching for description–experience gaps in other domains, it is worth keeping a few things in mind. First, all dichotomies are overly simplistic. In reality, the description–experience dichotomy is more like a continuum of situations, with pure experience and pure descriptions as its poles. There is also an important class of situations in which people are unable to recruit either experience or description. In such situations of utter uncertainty, in which "there is no valid basis of any kind for classifying instances" (Knight, 1921, p. 225), events are truly unique and probabilities cannot be empirically derived. Important singular life events – such as the decision to marry, the decision to have a child, and the choice of a profession – belong in this class. Or consider the gamble to invest in a start-up; investors in start-ups cannot consult descriptions of past performance, nor do they know to what extent other start-ups they have experienced constitute cases that are similar enough to be included in the same reference class, thus offering a basis for deriving empirical probabilities.

In my view, the description–experience dichotomy and the experiential paradigms offer cognitive and decision scientists a powerful tool to explore human cognition. Of course, any tool has its limits, but the strengths of this one are evident in the slew of questions it raises. Many questions relating to the description–experience gap remain open. For instance, is the experience of others (vicarious experience) the same as personal experience? When and why do people fail to learn from experience (e.g., Brehmer, 1980)? What is the essence of experience (e.g., psychological distance, reliability, trustworthiness) that distinguishes it from description? And how does experience inform decision making in fast-changing worlds for which descriptions are unlikely to exist?

The Description–Experience Gap and Risk Communication

Experts and the general public are not infrequently at odds with each other when reckoning with risks. One often cited reason is that they operate on the basis of different risk constructs (Slovic, 2000). Experts' definition of risk, designated to be the "objective" one, involves the enumeration of the risk's detrimental consequences (e.g., fatalities, injuries, disabilities) weighted by their probabilities of occurrence. Citizens' perceptions of risk do not map one-to-one on this metric. Rather, they include other qualitative characteristics of the hazards, such as how voluntary and controllable exposure to them is, their catastrophic potential, or their threat to future generations.

The existence of the description–experience gap highlights another potential source of expert–layperson disagreement. Expert and lay assessments can be distinguished by the extent to which they rely on either experienced- or description-based information or on both. A surgeon, for instance, has access to statistics on the side effects of a medical intervention *and* the personal experience of having, for example, replaced hundreds of hips or inserted hundreds of stents into blocked arteries. Patients or parents making decisions on behalf of their children can often only rely on statistics. For instance, "parents who research the side effects of the DTaP vaccine on the National Immunization Program website will find that up to one child out of 1,000 will develop high fever and about one child out of 14,000 will experience seizures as a result of immunization" (Hertwig et al., 2004, p. 534). A growing number of parents, having encountered such information, appear to decide against immunization. In U.S. states that permit personal belief exemptions to school immunization requirements, the mean exemption rate increased, on average, by 6% per year, from 0.99% in 1991 to 2.54% in 2004 (Omer et al., 2006). Of course, doctors have access to the same statistics as parents. But they can also draw on their experience, gathered across many patients, that vaccination rarely results in side effects. Few doctors will have encountered one of the rare cases (1 in 14,000) of a vaccine causing a seizure. And even if they have done so, this experience will be dwarfed by the memory of countless immunizations without side effects.

Access to description *and* experience is not always the privilege of experts, however. Take, for instance, the experience (or lack thereof) of natural disasters such as volcanic eruptions. In 3780 BC, an eruption of Mount Vesuvius, the still-active volcano looming over Naples, buried land and villages as far as 25 km away, causing the abandonment of the entire area for centuries. According to Mastrolorenzo, Petrone, Pappalardo, and Sheridan (2006), a comparable eruption today would cause total devastation and mortality within a radius of at least 12 km – that is, a significant chunk of the Naples metropolitan area. The last eruption was in 1944, and volcanologists have recently cautioned that "with each year, the statistical probability increases that there will be another violent eruption" (Wilford, 2006). In light of these expert warnings (description based), one might expect residents to be keen to leave the danger zone. But relocation has proven extremely difficult: "In the shadow of Vesuvius, those residents have cultivated a remarkable optimism, a transcendent fatalism and a form of denial as deep as the earth's molten core" (Bruni, 2003). One key to understand such puzzling disagreements between expert and public opinion (Hertwig & Frey, 2015) is to analyze both description-based risk warnings and their recipients' actual experience. Experience tells residents in the vicinity of Mount Vesuvius that violent eruptions are rare and, in fact, nonexistent in most people's lifetime. Against the backdrop of having lived safe and sound at the foot of mainland Europe's only remaining active volcano, expert warnings, clothed in numbers (probabilities) and descriptions of possible outcomes, lack persuasive power (see also Lejarraga & Gonzalez, 2011).

The general point is this: Risk warnings do not operate in a vacuum. Sometimes people have experienced numerous safe encounters with a risk event prior to obtaining a warning (e.g., the repeated experience of unprotected sex without contracting a sexually transmitted disease). Sometimes people receive the warning right after disaster has struck for the first time; sometimes they are blank slates with no experience at all. How risk communication affects behavior is likely to depend on people's past and present experience. Without understanding the intricate interplay of description and experience, we will continue to be surprised by how ineffectual warnings can be (see Barron, Leider, & Stack, 2008). The combination of descriptions of risks with simulated experience in virtual realities may prove a valuable tool to convey transparent and persuasive risk information in many key domains (e.g., investment; Kaufmann, Weber, & Haisley, 2012; see also Lejarraga, 2010) – especially in domains such as climate change, in which gradual changes are "virtually impossible to detect from personal experience, amid the noise of random fluctuation around the central trend" (Weber & Stern, 2011, p. 318).

Let Us Not Give Descriptions Short Shrift

One of the greatest cultural inventions of all time is communication with others through written symbols. By agreeing on the meaning of specific symbols we all take advantage of a powerful form of self-expression and are able to draw on the accumulated wisdom of others, their strokes of genius, and knowledge acquired through trial and error or lucky accidents. Unlike language, typically understood as a biological adaptation to communicate information, writing is a cultural adaptation. Over several millennia, simple three-dimensional symbols (tokens such as a cone) evolved into complex written symbols (Schmandt-Besserat, 1996), and the representation of the world's three-dimensional reality in terms of a small set of graphemes has played a revolutionary role in human affairs. Barely any aspect of modern life – from technology, science, commerce, literary arts to news media and the World Wide Web – is conceivable without reading and writing of symbolic descriptions. In the words of Schmandt-Besserat: "speech, the universal way by which humans communicate and transmit experience, fades instantly: before a word is fully pronounced it has already vanished forever. Writing, the first technology to make the spoken word permanent, changed the human condition" (p. 1).

The ability to extensively communicate in symbolic description is one of the qualities that distinguish humans from other animals, and the description–experience gap is, in all likelihood, a phenomenon unique to humans. One way to understand the recent tide of interest in experience-based choice is that it is somehow the worthier or richer subject of investigation. This conclusion would be wrong, and I can think of myriad unanswered questions about the psychology behind descriptions. For instance, why is it that people greatly overestimate relatively rare risks once they are explicitly stated (e.g., the risk of contracting lung cancer from smoking; Viscusi, 1990) and, relatedly, what are the driving forces behind the mere-presentation effect (Erev et al., 2008)? Why does unpacking or repacking the description of a hypothesis change an event's probability estimate (Rottenstreich & Tversky, 1997)? These are just a few of the many interesting questions concerning decisions based on descriptions to which full answers are still missing. The description–experience gap represents a new point of entry for research. It holds the promise, when combined with a "do-it-both-ways" heuristic (Hertwig & Ortmann, 2001), of more rapid progress in understanding the psychology and rationality of description *and* of experience.

Conclusions

In the early twentieth century, the economist Knight (1921) drew a conceptual distinction between risk and uncertainty – a distinction that had enormous impact on economists' and psychologists' theorizing about decision making. Decisions under risk refer to situations in which the probability distribution over the possible outcomes is known. Decisions under uncertainty concern situations in which events' probabilities are not known or in which the events can hardly be conceived of in advance, such as a meteor crash injuring more than 1,000 people or the first papal resignation in almost six centuries – two events of the recent past. One way to interpret the description–experience gap is as a risk–uncertainty gap. In decisions from experience, people can never be certain of being aware of all possible (e.g., even extremely rare) events in the payoff distributions, no matter how much they search. In decisions from description, in contrast, people are typically informed about all outcomes and their probabilities. Admittedly, this mapping is not perfect – for instance, descriptions can be incomplete. Yet, thus interpreted, research on the description–experience gap has led to genuinely new insights into the distinction between risk and uncertainty, and unraveling decisions from experience can reveal a great deal about how humans reckon with uncertainty.

Acknowledgments

I am grateful to Nick Chater, Ido Erev, Renato Frey, Cleotilde Gonzalez, Gideon Keren, Tomas Lejarraga, Thorsten Pachur, Tim Pleskac, Rui Mata, and George Wu for their comments. I also thank Susannah Goss for editing the manuscript and the Swiss National Science Foundation for grant CRSII1_136227.

References

Abdellaoui, M., L'Haridon, O., & Paraschiv, C. (2011). Experienced vs. described uncertainty: Do we need two prospect theory specifications? *Management Science, 57*(10), 1879–1895.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine [Rational man's behavior in face of risk: Critique of the American School's postulates and axioms]. *Econometrica, 21*(4), 503–546.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Anderson, J. R., & Lebiere, C. (2003). The Newell test for a theory of cognition. *Behavioral and Brain Sciences, 26*(5), 587–601.

Avrahami, J., Güth, W., Hertwig, R., Kareev, Y., & Otsubo, H. (2013). Learning (not) to yield: An experimental study of evolving ultimatum game behavior. *The Journal of Socio-Economics, 47*, 47–54. doi:10.1016/j.socec.2013.08.009

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

Barron, G., Leider, S., & Stack, J. (2008). The effect of safe experience on a warnings' impact: Sex, drugs, and rock-n-roll. *Organizational Behavior and Human Decision Processes, 106*(2), 125–142.

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

Brehmer, B. (1980). In one word: Not from experience. *Acta Psychologica, 45*, 223–241.

Bruni, F. (2003, August 26). San Giuseppe Vesuviano Journal: Who's afraid of Vesuvius? (Pompeii is history). *The New York Times.* Retrieved from http://www.nytimes.com

Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed-sample, and sequential-sampling models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(3), 538–564.

Camilleri, A. R., & Newell, B. R. (2011a). Description- and experience-based choice: Does equivalent information equal equivalent choice? *Acta Psychologica, 136*(3), 276–284.

Camilleri, A. R., & Newell, B. R. (2011b). 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*(2), 377–384.

Camilleri, A. R., & Newell, B. R. (2013). The long and short of it: Closing the description–experience "gap" by taking the long-run view. *Cognition, 126*(1), 54–71.

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological Review, 114*(1), 177–187.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51*(4), 380–417.
Edwards, W. (1961). Probability learning in 1000 trials. *Journal of Experimental Psychology, 62*(4), 385–394.
Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment* (pp. 17–52). New York, NY: Wiley.
Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75*(4), 643–669.
Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*(4), 912–931.
Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., … Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23*, 15–47.
Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*(5), 575–597.
Erev, I., Glozman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty, 36*(2), 153–177.
Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88*(4), 848–881.
Fox, C. R., & Hadar, L. (2006). "Decisions from experience" = sampling error plus prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making Journal, 1*(2), 159–161.
Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science, 44*(7), 879–895.
Frey, R., Hertwig, R., & Rieskamp, J. (2014). Fear shapes information acquisition in decisions from experience. *Cognition, 132*, 90–99.
Frey, R., Mata, R., & Hertwig, R. (2015). The role of cognitive abilities in decisions from experience: Age differences emerge as a function of choice set size. *Cognition 142*, 60–80.
Gigerenzer, G., Hertwig, R., Van den Broek, E., Fasolo, B., & Katsikopoulos, K. V. (2005). "A 30% chance of rain tomorrow": How does the public understand probabilistic weather forecasts? *Risk Analysis, 25*(3), 623–629.
Glaser, C., Trommershäuser, J., Mamassian, P., & Maloney, L. T. (2012). Comparison of the distortion of probability information in decision under risk and in an equivalent visual task. *Psychological Science, 23*(4), 419–426.
Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118*(4), 523–551.
Gonzalez, C., & Dutt, V. (2012). Refuting data aggregation arguments and how the instance-based learning model stands criticism: A reply to Hills and Hertwig (2012). *Psychological Review, 119*(4), 893–898.
Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making, 4*(4), 317–325.
Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making, 23*(1), 48–68.
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*(5), 493–518.
Hertwig, R. (2012). The psychology and rationality of decisions from experience. *Synthese, 187*(1), 269–292.

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*(8), 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., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517–523.

Hertwig, R., & Frey, R. (2015). The challenge of the descripton–experience gap to the communication of risks. In H. Cho, T. O. Reimer, & K. A. McComas (Eds.), *The Sage handbook of risk communication* (pp. 24–40). Thousand Oaks, CA: Sage.

Hertwig, R., Hoffrage, U., & the ABC Research Group (2013). *Simple heuristics in a social world*. New York, NY: Oxford University Press.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences, 24*, 383–451.

Hertwig, R., & Pleskac, T. J. (2008). The game of life: How small samples render choice simpler. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for Bayesian cognitive science* (pp. 209–236). Oxford, UK: Oxford University Press.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition, 115*(2), 225–237.

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21*(12), 1787–1792.

Hills, T. T., & Hertwig, R. (2012). Two distinct exploratory behaviors in decisions from experience: Comment on Gonzalez and Dutt (2011). *Psychological Review, 119*(4), 888–892.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. *Psychological Bulletin, 57*(2), 116–131.

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

Jarvstad, A., Hahn, U., Rushton, S. K., & Warren, P. A. (2013). Perceptuo-motor, cognitive, and description-based decision making seem equally good. *PNAS, 110*, 16271–16276. doi:10.1073/pnas.1300239110

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from Prospect Theory in descriptive choice. *Psychological Science, 19*(10), 1015–1022.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*(3), 430–454.

Kahneman, D., & Tversky, A. (1973). Psychology of prediction. *Psychological Review, 80*(4), 237–251.

Kahneman, D., & Tversky, A. (1979). Prospect theory: Analysis of decision under risk. *Econometrica, 47*(2), 263–291.

Kaufmann, C., Weber, M., & Haisley, E. C. (2012). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*. Advance online publication. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1616186

Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston, MA: Houghton Mifflin.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Normative, descriptive and methodological challenges. *Behavioral and Brain Sciences, 19*, 1–53.

Lejarraga, T. (2010). When experience is better than description: Time delays and complexity. *Journal of Behavioral Decision Making, 23*(1), 100–116.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making, 25*(2), 143–153.

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*(3), 334–342.

Locke, J. (1690/1959). *An essay concerning human understanding*. A. C. Fraser (Ed.). New York, NY: Dover.

Luce, R. D., & Suppes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 3, pp. 249–410). 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*(6) (Article e20262).

March, J. G. (1996). Learning to be risk averse. *Psychological Review, 103*(2), 309–319.

Mastrolorenzo, G., Petrone, P., Pappalardo, L., & Sheridan, M. F. (2006). The Avellino 3780-yr-BP catastrophe as a worst-case scenario for a future eruption at Vesuvius. *Proceedings of the National Academy of Sciences of the United States of America, 103*(12), 4366–4370.

Omer, S. B., Pan, W. K. Y., Halsey, N. A., Stokley, S., Moulton, L. H., Navar, A. M., … Salmon, D. A. (2006). Nonmedical exemptions to school immunization requirements: Secular trends and association of state policies with pertussis incidence. *Journal of the American Medical Association, 296*(14), 1757–1763.

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.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.

Peters, E. (2012). Beyond comprehension: The role of numeracy in judgments and decisions. *Current Directions in Psychological Science, 21*(1), 31–35.

Phillips, L. D., & Edwards, W. (1966). Conservatism in a simple probability inference task. *Journal of Experimental Psychology, 72*(3), 346–354.

Pleskac, T. J., & Hertwig, R. (2014). Ecologically rational choice and the structure of the environment. *Journal of Experimental Psychology: General, 143*, 2000–2019.

Quetelet, L. A. J. (1842/1969). *A treatise on man and the development of his faculties*. Translated by R. Knox. Facsimile edition of translation 1969. Introduction by S. Diamond. Gainesville, FL: Scholars' Facsimiles & Reprints.

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*(2), 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), 1–14.

Rottenstreich, Y., & Tversky, A. (1997). Unpacking, repacking, and anchoring: Advances in support theory. *Psychological Review, 104*(2), 406–415.

Savage, L. J. (1954). *The foundations of statistics*. New York, NY: John Wiley & Sons.

Schmandt-Besserat, D. (1996). *How writing came about*. Austin, TX: University of Texas Press.

Schulte-Mecklenbeck, M., Kühberger, A., & Ranyard, R. (Eds.). (2011). *A handbook of process tracing methods for decision research: A critical review and user's guide*. New York, NY: Taylor & Francis.

Slovic, P. (2000). *The perception of risk*. London, UK: Earthscan.

Smith, V. L. (2001). From old issues to new directions in experimental psychology and economics. *Behavioral and Brain Sciences, 24*(3), 428–429.

Spaniol, J., & Wegier, P. (2012). Decisions from experience: Adaptive information search and choice in younger and older adults. *Frontiers in Neuroscience, 6*, Article 36.

Suter, R., Pachur, T., & Hertwig, R. (2013). How does prospect theory reflect heuristics' probability sensitivity in risky choice? In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Cooperative minds: Social interaction and group dynamics. Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 1408–1413). Austin, TX: Cognitive Science Society.

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

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable.* New York, NY: Random House.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review, 102*(2), 269–283.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.

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

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted, when rare outcomes are experienced (rarely)? *Psychological Science, 20*(4), 473–479.

Viscusi, W. K. (1990). Do smokers underestimate risks? *Journal of Political Economy, 98*(6), 1253–1269.

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

Weber, E. U., & Kirsner, B. (1996). Reasons for rank-dependent utility evaluation. *Journal of Risk and Uncertainty, 14*, 41–61.

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*(2), 430–445.

Weber, E. U., & Stern, P. C. (2011). Public understanding of climate change in the United States. *American Psychologist, 66*(4), 315–328.

Wilford, J. N. (2006, March 7). Long before burying Pompeii, Vesuvius vented its wrath. *The New York Times.* Retrieved from http://www.nytimes.com

Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2009). Economic decision making compared with an equivalent motor task. *Proceedings of the National Academy of Sciences of the United States of America, 106*(15), 6088–6093.

Wu, S.-W., Delgado, M. R., & Maloney, L. T. (2011). The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision making. *Journal of Neuroscience, 31*(24), 8822–8831.

Wulff, D., Hertwig, R., & Mergenthaler, M. (2015). *Ten years of the description–experience gap. A meta-analysis.* Manuscript submitted for publication.

Wulff, D. U., Hills, T. T., & Hertwig, R. (2015). How short- and long-run aspirations impact search and choice in decisions from experience. *Cognition, 144*, 29–37.

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

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.