

CHAPTER 2

The Challenge of the Description–Experience Gap to the Communication of Risks

Ralph Hertwig and Renato Frey

Introduction

Experts and the lay public are often at odds with each other when assessing risks. A common explanation for these disagreements in expert and lay opinions is that experts tend to operate on the basis of a technical (“objective”) definition of risk. This definition is generally based on a risk’s detrimental consequences (e.g., fatalities, injuries, disabilities), weighted by the probabilities of those consequences. Laypeople’s assessments of risk—and specifically their perceptions of risk—do not simply follow this metric. Instead, they include other qualitative characteristics of the hazards, such as whether exposure to the risk is voluntary, how controllable the risk is, its catastrophic potential, or its threat to future generations (Slovic, 1987). While acknowledging the importance of this insight that the concept of “risk” means different things to different people, we propose another key factor that may underlie

disagreements between risk experts and the general public. Specifically, information about risks can be acquired via explicit, convenient descriptions of outcomes and their probabilities (e.g., probabilistic weather forecasts, actuarial tables, and mutual fund brochures) or through the sequential experience of the occurrence or nonoccurrence of risky events. Before we explain this distinction (Hertwig, Barron, Weber, & Erev, 2004) in detail, let us describe three instances of expert/lay disagreements. As we will then go on to show, the description–experience distinction offers a key to understanding these disagreements.

Vaccination

Not infrequently, doctors and parents disagree on the benefits and dangers of vaccination. Consider, for example, the decision whether or

Authors’ Note: Parts of this text are based on articles by Hertwig and Erev (2009) and Hertwig (2012). We thank Susannah Goss for editing the manuscript and Elsevier and Springer for granting permission rights, and the Swiss National Science Foundation for a grant to the first author (100014–126558).

not to vaccinate a child against diphtheria, tetanus, and pertussis (DTaP). As Hertwig et al. (2004) have described,

Parents who research the side effects of the DTaP vaccine on the National Immunization Program Web site will find that up to 1 child out of 1,000 will develop high fever and about 1 child out of 14,000 will experience seizures as a result of immunization. (p. 534)

Increasing numbers of parents, after reading such information, decide not to immunize their child. In those U.S. states that permit personal belief exemptions to school and day care 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). Although doctors have the same statistics at their disposal, they also draw on information not easily available to parents—namely, their personal experience, gathered across many patients, that vaccination rarely results in side effects. Indeed, few doctors have encountered one of the rare cases (1 child in 14,000) of seizures following a vaccination. And even if they have, this experience will be dwarfed by the memory of countless immunizations without side effects. Parents, in contrast, cannot draw on this large stock of personal memories of trouble-free vaccinations.

Terrorism

In the recent past, terrorists have repeatedly targeted tourist centers. In 1997, for instance, terrorists killed 62 people outside the Temple of Queen Hatshepsut at Luxor in Egypt, all but four of them foreigners (“Bloodbath at Luxor,” 1997). In 2008, terrorist attacks struck the heart of Mumbai, India’s commercial capital, with 166 locals and foreigners being killed in machine-gun and grenade assaults (“Once more to the gallows,” 2012). Beyond the lives lost, terrorist attacks exact economic costs—through slumps in the local tourism industry, for example. Yechiam, Barron, and

Erev (2005) analyzed the economic costs of the Al-Aqsa Intifada (uprising), a wave of terrorist attacks in Israel that were “targeted towards specific civilian targets, including hotels, restaurants, cafes, and clubhouses” (p. 432). The costs were operationalized in terms of the number of overnight stays in Israeli hotels before and after the outbreak of the Intifada, separately for domestic and foreign tourists. Domestic tourists might—to use an admittedly rather loose definition—be regarded as “experts on the ground” and foreign tourists as “laypeople.” The difference between these two groups was striking: In October 2001, about a year after the onset of the Intifada, the overnight stays of foreign tourists showed an 80% decrease relative to October 2000, as compared with a 20% *increase* for domestic tourists.

Why did “experts” and “laypeople” respond so differently to the threat of terrorist attacks? Acknowledging other contributing factors, Yechiam et al. (2005) proposed that local residents behaved differently because they continued to attend public places such as cafés and markets—where their most common experience was that nothing happened. Terrorist attacks were, fortunately, rare events. Foreign tourists, in contrast, lacked the experience of countless enjoyable or at least uneventful visits to public places. Like parents whose knowledge of vaccination was gleaned from the National Immunization Program website, their primary source of information came via descriptions, here the international media coverage of the most recent terrorist attacks; they did not share the everyday experiences of Israel’s residents.

Natural Hazards

Although the most notorious eruption of Mount Vesuvius occurred in 79 CE, destroying Pompeii, the luxurious resort of wealthy Romans, it was not the largest in scope. A Bronze Age eruption around 3780 BCE buried land and villages as far as 25 km (about 16 miles) away, causing the abandonment of the entire area for centuries

(Mastrolorenzo, Petrone, Pappalardo, & Sheridan, 2006). At present, at least 3 million people live within the area that was destroyed by this Bronze Age eruption, and the periphery of Mount Vesuvius includes a significant chunk of the Naples metropolitan area (Bruni, 2003). According to the volcanologists Mastrolorenzo et al. (2006), an eruption comparable in magnitude to the Bronze Age eruption would cause total devastation and mortality within a radius of at least 12 km (about 8 miles). Moreover, volcanologists have argued that it has been roughly 2,000 years since Pompeii, and “with each year, the statistical probability increases that there will be another violent eruption” (Wilford, 2006).

In light of these dire forecasts, one might expect that local residents would be keen to move away from the danger zone. On the contrary, relocating residents has proven extremely difficult, and “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). Why the disparity between expert and lay opinions? As we have argued, personal experience or lack thereof may be the key to understanding this and other puzzling disagreements between experts and the lay public. Their personal experience tells residents living in the vicinity of Mount Vesuvius that violent eruptions are extremely rare; in fact, in most people’s lifetime, they just do not happen (the last major eruption of the Vesuvius was in 1944; for a similar phenomenon, see the residents of L’Aquila, the Italian town that “had become complacent about the seismic danger” of living near a particular type of fault; Silver, 2012, p. 144). Volcanologists have only numbers (probabilities) and descriptions of possible outcomes to counter the allure of people’s personal experience.

The common thread that connects the three introductory examples is that experts and laypeople disagree in their evaluations of the respective risks and that this disagreement may originate in part from the degree to which they rely on description-based versus experienced-based information about risks. Sometimes the

world affords people convenient synopsis descriptions of risky prospects—for example, the side effects detailed in drug package inserts or the risks of adverse weather events (e.g., a hurricane making landfall at a specific location) reported in the media (Gigerenzer, Hertwig, Van Den Broek, Fasolo, & Katsikopoulos, 2005). Equipped with such explicit and quantitative risk information, people can make *decisions from description* (Hertwig et al., 2004). The luxury of such quantified and stated risk information is restricted to just a few domains. In general, we humans have to navigate the perils and opportunities of our environment without tabulated risks. We have to make many consequential decisions, such as whether to marry and have children, as well as countless everyday decisions, such as whether to jaywalk, to back up our computer, or to wear a helmet when cycling, without full knowledge of the whole range of outcomes and their probabilities. To the extent that people’s past or present experiences inform their current decisions, people make *decisions from experience* (Hertwig et al., 2004).

Relying on personal experience when assessing risks has interesting implications. Often, but not always, people’s samples of personal experience are limited. An individual’s experience cannot, generally speaking, approach the scope of the collective, aggregated experience that is encapsulated in tabulated risks. Limited samples, in turn, tend to underrepresent rare (but possibly impactful) events. Therefore, when people draw on their experience, sampled across time, to make decisions involving risks, the chances are that rare events—for example, the eruption of a volcano, the burst of a housing bubble, a vaccination-induced seizure, an accident due to jaywalking, or the loss of vast amounts of data in a computer crash—have *less* impact on people’s decisions than they deserve according to their objective probability (see Hertwig & Erev, 2009). But even when a person’s immediate experience is sizable (e.g., as in the case of doctors administering hundreds of vaccinations), it may still confer less impact to rare events than explicit descriptions of rare events do (e.g., brochures that inform about the risks of

vaccination). Indeed, when people operate on the basis of symbolically described versions of risky events (e.g., distributions of possible outcomes and associated likelihoods), they appear to give rare events *more* weight than they deserve according to their objective probability. This tendency of overweighting of rare events in decisions from description is one of prospect theory's postulates (see Tversky & Kahneman, 1992), the most influential descriptive theory of risky choice; we return to one of the possible explanations of this phenomenon below.

These opposite tendencies regarding the psychological impact of rare events result in a *description–experience gap* (Hertwig et al., 2004): a robust and systematic discrepancy between experience- and description-based choices that has been observed in numerous studies, most (but not all) of them involving monetary gambles (for a review, see Hertwig & Erev, 2009). Before we review the psychology underlying the description–experience gap, let us briefly return to our introductory examples of disagreements between experts and the general public. In many cases, expert and lay decision makers can be distinguished by the degree to which they rely on either experience-based or description-based information, or on both. This difference can go in either direction. As the vaccination example illustrates, expert medical decision makers (i.e., doctors) have access to both statistics on the side effects of vaccination and their personal experience of having administered a vaccine many times. Parents, in contrast, can draw only on the statistics (and possibly the anecdotal experience of other parents, which probably represents a selective sample of experiences of rare adverse side effects; see also Berger, 2007). Similarly, local residents of Israel could draw on their “expert” personal experience as well as on statistics and media reports to gauge the risk of frequenting public places during the Intifada. Foreign tourists, in contrast, only had access to descriptions of events (e.g., newspaper reports). In the case of the Vesuvius, in contrast, it is the local residents and not expert volcanologists who can draw on personal experience of having lived

in the vicinity of the volcano—as well as on experts’ “statistical” warnings that the volcano will erupt again.

The relative indifference with which citizens and politicians sometimes consider rare but highly consequential events, such as bursting levees, floods, and eruptions of volcanoes, may be owed to the experience of their rarity. Just as people living in the shadow of Mount Vesuvius have turned down attempts to relocate them, people living in flood plains tend to turn down offers even of federally subsidized flood insurance (Kunreuther, 1984). People who lack pertinent experience have to rely, if available, on descriptions of the possible consequences of risky events and their probabilities (e.g., vaccines and their side effects). In these situations, people appear to overweight the impact of rare events and may then overreact to risks such as side effects of vaccination or the risk of contracting swine flu (H1N1). To understand how people respond to rare but high-consequence events, and to the communication of the respective risks, researchers need to take into account the psychology of people's decisions from experience and from description (Hertwig et al., 2004).

In what follows, we explain the description–experience gap and how it has typically been studied in more detail.

The Description–Experience Gap

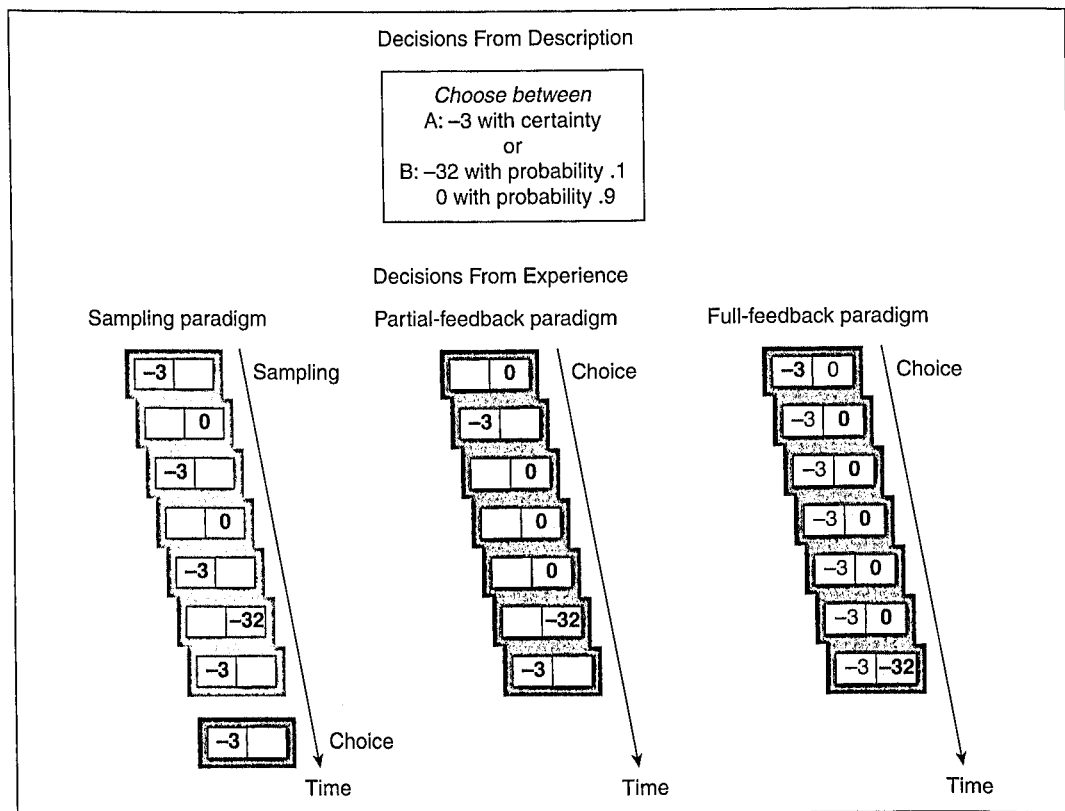
Just as biologists use the *Drosophila* as a model organism, behavioral decision researchers have used choice between monetary gambles as a model for risky choice, assuming that many real-world options have the same properties, namely, n outcomes and associated probabilities (Lopes, 1983). From this perspective, parents deciding whether not to have their child vaccinated choose between two risky “gambles.” The first is to choose vaccination and face two possible outcomes, namely, adverse side effects with a probability p_{SE} , and otherwise a healthy child. The second is to forgo vaccination and face two possible outcomes, namely, the child contracting

diphtheria, tetanus, or pertussis with a probability p_{DTaP} and otherwise a healthy child.

Behavioral decision researchers investigating choice between monetary gambles as a model for risky choice have grown accustomed to presenting

their respondents with one particular genus of the fruit fly: gambles in which all outcomes and their probabilities are stated, and respondents make a single choice, as illustrated in the upper panel of Figure 2.1. And, indeed, parents who

Figure 2.1 A Description-Based Paradigm and Three Experiential Paradigms for Studying Choice



Source: Based on Hertwig and Erev (2009).

Note: How to study decisions from descriptions and experience? The choice task in decisions from description (*upper panel*) typically consists of two monetary gambles with explicitly stated outcomes and their probabilities. In research on decisions from experience (*lower panel*), three paradigms have been employed: The *sampling paradigm* consists of an initial sampling stage (here represented by seven fictitious draws) in which the participant explores two payoff distributions (that offer various monetary outcomes, each one associated with a certain probability) at no cost by clicking on one of two buttons on a computer screen (shown here outlined in light gray). The buttons chosen by the participant are shaded. After terminating sampling, the participant sees a choice screen (here shown outlined in dark gray) and is asked to draw once for real. The *partial-feedback paradigm* collapses sampling and choice; thus, each draw represents an act of both exploration and exploitation. The participant receives feedback on the obtained payoff after each draw (shaded box). The *full-feedback paradigm* is identical to the partial-feedback paradigm, except that it also provides feedback on the forgone payoff (i.e., the payoff that the participant would have received, had he or she chosen the other option; white box).

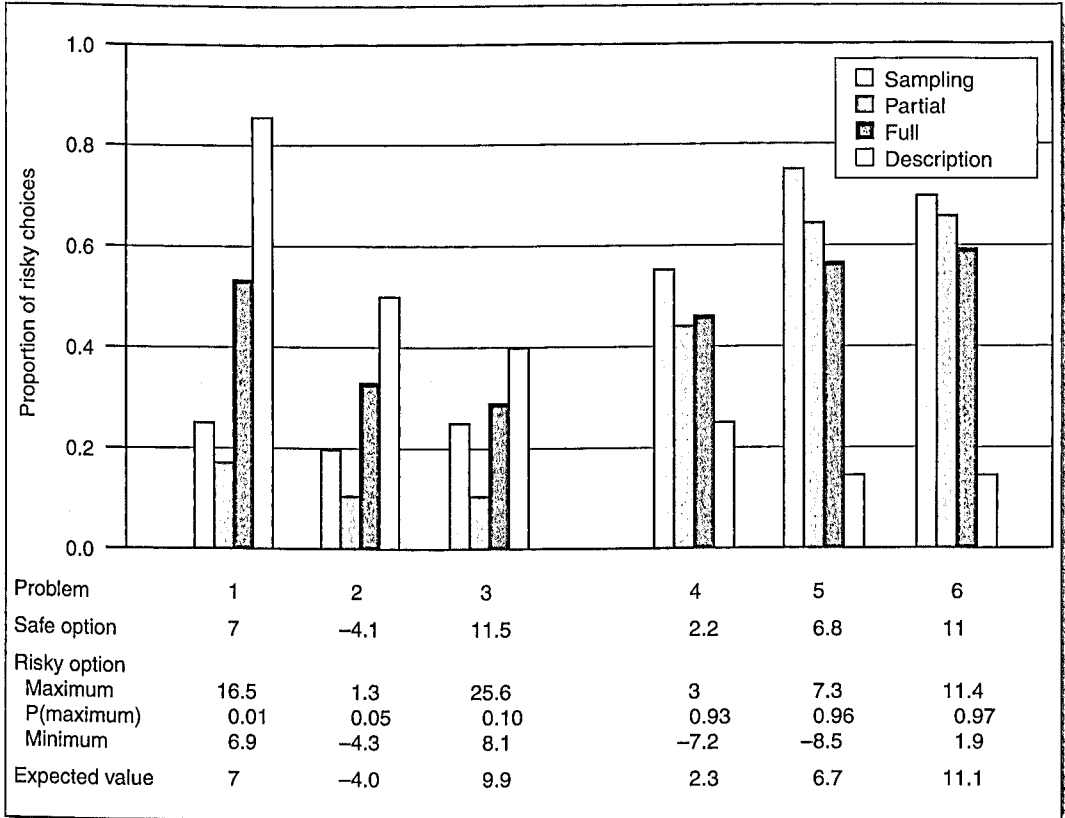
choose for or against vaccination enjoy the convenience of explicitly described probabilities and outcomes. This is a rare exception, however. In everyday life, people rarely have access to such descriptions of probability distributions. When people decide whether to take out a mortgage or contemplate the success of a first date, there are no tabulated risks to consult. Instead, they need to rely on their previous experience—if existent—of these options. Decisions from experience and decisions from description can be understood as located at opposite ends of a continuum of uncertainty.

In the first years of the new millennium, the observation of systematic and robust differences between decisions based on experience and decisions based on description has drawn decision scientists' interest back to decisions from experience. Before we turn to their findings, let us briefly explain how researchers investigate decisions from experience. In general, they employ a simple experimental tool, a "computerized money machine." Respondents see two buttons on a computer screen, each one 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 been employed (lower panel of Figure 2.1). In the *sampling paradigm*, participants first sample as many outcomes as they wish and only then decide from which distribution to make a single draw for real (Hertwig et al., 2004; Weber, Shafir, & Blais, 2004). In the *full-feedback paradigm*, each draw contributes to participants' earnings, and they receive draw-by-draw feedback on the obtained and forgone payoffs (i.e., the payoff they would have received had they selected the other option; Yechiam & Busemeyer, 2006). The *partial-feedback paradigm* is identical to the full-feedback paradigm, except that participants are only informed about the obtained payoffs (Barron & Erev, 2003; Erev & Barron, 2005). In contrast to the sampling and full-feedback paradigms, respondents face an exploitation–exploration trade-off partial-feedback paradigm as they negotiate between the two goals associated with every choice: to obtain a desired outcome (exploitation)

or to gather new information about other, perhaps better, actions (exploration; Cohen, McClure, & Yu, 2007).

Across all three experiential paradigms, a robust and systematic *description–experience gap* has emerged in numerous studies. Figure 2.2 illustrates this gap in six decision problems (Erev et al., 2010), each of which offers 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 (probability of .1 or less). In all three experiential paradigms, respondents tend to select the risky option when the desirable outcome occurs with high probability but tend to select the safe option when the desirable outcome occurs with low probability. This tendency is reversed in decisions from description. The general pattern 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. By way of illustration, let us consider Problem 1, in which the risky option offers a relatively large gain of 16.5 with a small probability of 1% (and 6.9 otherwise). If the psychological weight of this rare outcome amounts to less than its objective probability, then the overall value of the risky option will become less attractive than the safe bet of 7. Consequently, people will be more likely to choose the safe than the risky option in decisions from experience, and they indeed do (Figure 2.2). If, however, the rare outcome's psychological weight exceeds its objective probability, then the overall value of the risky option will become more attractive, relative to the safe bet of 7. Consequently, people will be more likely to choose the risky than the safe option in decisions from description. The same logic applies to rare events that involve negative consequences (losses), except that people will now be drawn to the safe option in decisions from description and to the risky option in decisions from experience. Figure 2.2 demonstrates this reversal of preferences.

Figure 2.2 The Description–Experience Gap: Proportion of Choices of the Risky Option as a Function of Paradigm (Description vs. Experience) and as a Function of Decision Problems in Which the Rare Event Is Desirable (Maximum) or Undesirable (Minimum)



Source: Based on Hertwig and Erev (2009).

Note: Each decision problem 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. Each problem was studied using the four paradigms shown in Figure 2.1. Participants (20 per paradigm) were paid (in shekels) the value of one of their choices (which was randomly selected). The partial- and full-feedback paradigms involved 100 choices per problem, and the reported proportions are the means over these choices and participants.

Before we turn to the causes of this reversal, let us briefly review what behavior emerges when respondents receive description *and* experience at the same time. Jessup, Bishara, and Busemeyer (2008) explored how feedback influences repeated decisions from description in a monetary binary choice task. Respondents made 120 repeated choices in response to descriptive information. One group received feedback (indicating

their winnings on the previous trial); another group received no feedback. Based on the observed trajectory of choices, the authors concluded, “Apparently, feedback overwhelms descriptive information” (p. 1019), and “individuals who received feedback underweighted small probabilities relative to their no-feedback counterparts” (p. 1019), with the former moving toward linear objective weighting of probabilities.

Similarly, Yechiam et al. (2005) observed that experience limits the effect of description. Using a similar experimental setup as Jessup et al. (2008), Lejarraga and Gonzalez (2011) reached at an even stronger conclusion about the relative impact of description and experience: “Our results suggest that decision makers . . . overlook the descriptive information in an attempt to simplify the cognitive decision process” (p. 289). Whatever the reasons for the observed dominance of experience relative to description, it has potential implications for risk communication; we return to these later.

Reflections for Theory and Research

In this section, we review several factors that have been suggested as contributing to the description–experience gap.

Small Samples

Reliance on small samples has been proposed as one factor that contributes to the attenuated impact of rare events (Hertwig et al., 2004). Across numerous studies employing the sampling paradigm, respondents have typically proved restrained in their information search, with the median number of samples per choice problem typically ranging between 11 and 19 (reviewed in Hau et al., 2010). The chances are that a respondent drawing such small samples will not experience rare events. Even if they do so, the rare event will be encountered less frequently than expected (given its objective probability). This is because the binomial distribution for the number of times a particular outcome will be observed in n independent trials is markedly skewed when p is small (i.e., the event is rare) and n is small (i.e., few outcomes are sampled). For example, let us assume that each of 1,000 players draws 20 times from a distribution in which an attractive outcome occurs with a small probability of .1. Of the

1,000 players, 285 will encounter the event 2 times and could thus estimate its probability accurately. Another 323 will experience the event 3, 4, 5, . . . , or 20 times and would thus most likely overestimate its probability. But 392 players—almost two fifths of the total—will not encounter the event at all, or will encounter it only once, and would thus most likely underestimate its probability.

Interestingly, reliance on small samples has also been discussed as a potential explanation for bumblebees’ underweighting of rare events: Having studied the foraging decisions of bees in a spatial arrangement of flowers that promised different amounts of nectar with varying probabilities, Real (1991) concluded that “bumblebees underperceive rare events and overperceive common events” (p. 985). Real (1992) explained this distortion in bees’ probability perception as a consequence of their sampling behavior—“bees frame their decisions on the basis of only a few visits” (p. 133)—and suggested that such reliance on small samples can be adaptive:

Short-term optimization may be adaptive when there is a high degree of spatial autocorrelation in the distribution of floral rewards. In most field situations, there is intense local competition among pollinators for floral resources. When “hot” and “cold” spots in fields of flowers are created through pollinator activity, then such activity will generate a high degree of spatial autocorrelation in nectar rewards. If information about individual flowers is pooled, then the spatial structure of reward distributions will be lost, and foraging over the entire field will be less efficient. In spatially autocorrelated environments (“rugged landscapes”), averaging obscures the true nature of the environment. (p. 135)

In other words, Real (1992) suggested that in environments in which a set of features is clustered together in space, reliance on a small sample is adaptive. Could there be any advantage to

frugal sampling in humans' decisions from experience? Hertwig and Pleskac (2008, 2010) proposed one possible advantage that is rooted in the notion of amplification. Unlike Real, however, they argued that amplification offers a cognitive rather than an evolutionary benefit. Through mathematical analysis and computer simulation, Hertwig and Pleskac (2010) showed that small samples amplify the difference between the options' average rewards. That is, drawing small samples from payoff distributions results in experienced differences of sample means that are larger than the objective difference. Such amplified absolute differences make the choice between gambles simpler, thus explaining the frugal sampling behavior observed in investigations of decisions from experience.

This explanation of the description–experience gap in terms of small samples has been the subject of critical debate (Fox & Hadar, 2006). One question is whether the gap observed in the sampling paradigm can in fact be fully attributed to sampling error. Indeed, small samples on average cause the probability of rare events to be underestimated (as illustrated by the above example of 1,000 people and their estimates) and, on average, the smaller the sample, the larger the error. If sampling error were the sole culprit, however, reducing the error by increasing the sample size would attenuate and eventually eliminate the gap. Yet increasing sample sizes substantially (up to 50 and 100 draws per choice problem) reduced, but did not eliminate, the gap (Hau et al., 2010; Hau, Pleskac, Kiefer, & Hertwig, 2008). Were sampling error the sole cause of the gap, moreover, removing the error by aligning the experienced probabilities with the objective probabilities should eliminate it. Yet empirical findings have shown that this is not the case (Ungemach, Chater, & Stewart, 2009). Furthermore, if sampling error were solely to blame, then presenting respondents in the description condition with exactly the same information that others experienced (“yoking”) should eliminate the gap. In one study, it did (Rakow, Demes, & Newell, 2008); in another, it did for small but not for large samples

(Hau et al., 2010; see these authors' discussion of “trivial choices” as one possible explanation for the mixed results obtained). Finally, the gap persisted even when people were presented with both descriptions and experience, rather than just descriptions (Jessup et al., 2008).

At this point in the research process, the reality of the description–experience gap across the three experiential paradigms is unchallenged—its cause, however, is debated. Some researchers have argued that the gap in the sampling paradigm is statistical in nature (Fox & Hadar, 2006; Hadar & Fox, 2009; Rakow et al., 2008); others have proposed that the sampling error is not the sole cause (Hau et al., 2008; Hau et al., 2010; Hertwig et al., 2004; Ungemach et al., 2009). Regardless of how this debate advances, it is informative to go beyond the sampling paradigm. Reliance on small samples cannot be the reason behind the description–experience gap in the full-feedback paradigm (see Figure 2.1), for example, in which the impact of rare events is attenuated even after hundreds of trials with perfect feedback. Beyond sampling error, what psychological factors might come into play?

Recency

A psychological factor proposed to contribute to the description–experience gap is *recency* (Hertwig et al., 2004). Ubiquitously observed in memory, belief updating, and judgments (Hogarth & Einhorn, 1992), the recency effect describes the phenomenon that observations made late in a sequence receive more weight than they deserve (i.e., more than $1/n$). Recency is closely related to reliance on small samples: the small sample of recent events can reintroduce the aforementioned skew into large samples of experience. Although the original finding was that people give more weight to outcomes that occurred recently in the flow of their experience than to previous outcomes (Hertwig et al., 2004), later studies showed no or little impact of recency (Hau et al., 2010; Rakow et al., 2008; Ungemach et al., 2009).

Estimation Error

In theory, the description–experience gap could also be the consequence of people systematically *underestimating* the frequency of the rare event that they experienced in the sample (Fox & Hadar, 2006). However, a large stock of studies of frequency and probability assessments report the opposite tendency, namely, overestimation of rare events when people are asked to estimate, for instance, the frequency of lethal events (Hertwig, Pachur, & Kurzenhäuser, 2005; Lichtenstein, Slovic, Fischhoff, Layman, & Combs, 1978). Moreover, studies recording people's estimates of rare events in the sampling paradigm found them to be well calibrated or a little too high relative to the experienced frequency (Hau et al., 2008; Ungemach et al., 2009). That is, people do not systematically estimate rare things to be even rarer.

Reliance on Selective Past Experience

Another factor potentially underlying the description–experience gap, especially in the feedback paradigm, is that people recruit recent and past experiences in similar situations when making decisions (for related notions, see Gilboa & Schmeidler, 1995; Gonzalez & Dutt, 2011). This tendency is likely to be ubiquitous in the wild (Klein, 1999). For example, when firefighters need to predict the behavior of a fire, they appear to retrieve from memory similar instances from the past. Recruiting similar past experiences implies recency and reliance on small sampling to the extent that similarity decreases with time. Furthermore, in dynamic environments (e.g., restless bandit problem; Whittle, 1988), reliance on similar experiences is an efficient heuristic (Biele, Erev, & Ert, 2009). Below, we consider how the process of contingent sampling can be modeled.

Spatial Search Policies

Like any organism, humans can sample information from payoff distributions (e.g., flowers,

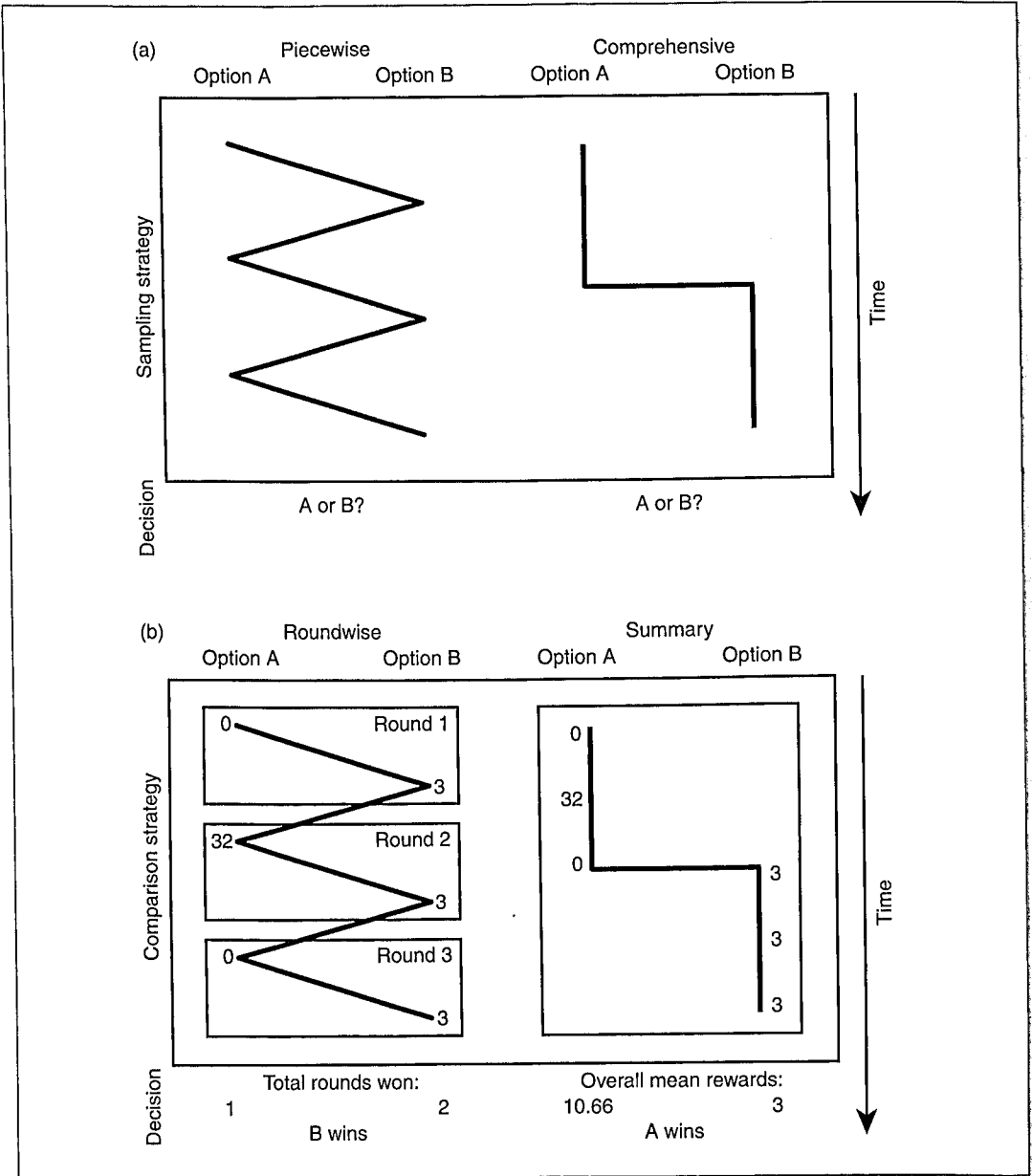
ponds, other people, gambles) in at least two very different ways. Figure 2.3 depicts two paradigmatic sequential-sampling strategies from two options. In *piecewise* sampling, the searcher oscillates between options, each time drawing (in the most extreme case) the smallest possible sample. In *comprehensive* sampling, in contrast, the searcher samples extensively from one option and only then turns to the other.

Taking these two sampling strategies as a starting point, Hills and Hertwig (2010) suggested that the (spatial) method of sampling foreshadows how people make their final decision. Specifically, they predicted that someone who samples piecewise will make decisions like a judge scoring each round of a boxing match: He or she will determine which option yields the better reward in each round of sampling and will ultimately pick the one that wins the most rounds. In contrast, someone using a comprehensive sampling strategy will gauge the average reward for each option and then choose the one promising the larger harvest. Piecewise and comprehensive sampling strategies thus foster comparisons across different scales of information: rounds versus summaries, respectively. Determining which option is ahead in most rounds versus which yields the largest expected reward can lead to different choices, even when both decision makers experience the exact same information. This is because the piecewise strategy weights each round equally, ignores the magnitude of wins and losses, and thus underweights rare but consequential outcomes. Indeed, Hills and Hertwig (2010) found that individuals who frequently oscillated between options, relative to those who rarely switched, were more likely to choose the roundwise winning options and were also more likely to make choices as if they underweighted rare events.

The Mere-Presentation Effect: Analogical Versus Propositional Representations

All potential causes of the description–experience gap listed so far concern decisions

Figure 2.3 Two Paradigmatic Search Policies in Decisions From Experience



Source: Based on Hills and Hertwig (2010).

Note: (a) Representations of the sampling patterns associated with piecewise and comprehensive sampling strategies. Piecewise strategies repeatedly alternate back and forth between options. Comprehensive sampling strategies take one larger sample from each option. Following the sampling phase, the participants make a decision about which option they prefer. (b) Representations of the comparison strategies associated with roundwise and summary strategies for a set of hypothetical outcomes. Roundwise strategies compare outcomes over repeated rounds and choose options that win the most rounds. Summary strategies compare final values (here, the overall expected value) and choose options with the better final value.

from experience; none deal with the tendency to overweight rare events in decisions from description. But why does this overweighting of rare events occur? The phenomenon has been postulated within prospect theory, the most influential *descriptive* account of how people choose between risky options. Prospect theory deals with empirical violations of the most important *normative* theory of risky choice, namely, expected utility. According to expected utility theory, people making choices behave as if they were multiplying some function of the outcomes' subjective value (utility) with the outcomes' probabilities and then maximizing (i.e., choosing the option that promises the highest expected utility). Prospect theory was proposed in response to experimental evidence showing that people systematically violate expected utility theory in their choices. It proposed several modifications of expected utility theory to address these empirical violations. One relates to how people respond to stated probabilities in general and to extreme probabilities in particular. The originators of prospect theory, Kahneman and Tversky (1979) asserted, "Because people are limited in their ability to comprehend and evaluate extreme probabilities, highly unlikely events are either neglected or overweighted, and the difference between high probability and certainty is either neglected or exaggerated" (p. 283). In more formal terms, Kahneman and Tversky argued that people do not take stated probabilities at face value when choosing but that these probabilities enter choices via decision weight. These weights are obtained from the objective probabilities by a nonlinear, inverse S-shaped weighting function that overweights small probabilities and underweights moderate and large ones (resulting in an inverse S shape).

But why should rare events be overweighted when explicitly stated? Hertwig, Barron, Weber, and Erev (2006) and Erev, Glozmann, and Hertwig (2008) suggested that a mere mention of those events lends them weight—a phenomenon they referred to as the *mere-presentation effect*. The *propositional* (symbolic) representations of options in decisions from description—for instance, "32

with probability .1; 0 otherwise"—may put more equal emphasis on the possible outcomes than the actual probabilities of occurrence warrants. If attention translates into decision weight, as some research suggests (Weber & Kirsner, 1996), then, other things being equal, the weights of rare and common events will draw closer together than they should. Decisions from experience, in contrast, 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. Information regarding the relative frequency of the option's outcomes can thus be read off directly. Moreover, to the extent that more attention is automatically allocated to the processing of the event that occurs more frequently (i.e., 0) than the rare event (i.e., 32), the resulting weights may more accurately reflect the sample probabilities (and, by extension, the objective probabilities).

In sum, having previously focused on people's responses to descriptions of events, behavioral decision research has more recently turned to decisions from experience, using three experiential paradigms to study how experience affects risky choice. A consistent picture has emerged: Description- and experience-based decisions can drastically diverge, especially when rare events are involved (for the demonstration of the gap with common events, see Ludvig & Spetch, 2011). Several factors have been proposed (e.g., sampling error, recency, reliance on selective past experience, propositional representation) as causing people to give rare events less impact than they deserve (according to their objective probabilities) in decisions from experience but more impact than they deserve in decisions from description. We now discuss the implications of the description–experience gap when extrapolated to the domain of risk communication.

Recommendations for Practice

The psychologist Paul Slovic (2000) described the following experience as the starting point of his influential research program on risk perception:

In 1970, I was introduced to Gilbert White, who asked if the studies on decision making under risk that Lichtenstein and I had been doing could provide insight into some of the puzzling behaviors he had observed in the domain of human response to natural hazards. Much to our embarrassment, we realized that our laboratory studies had been too narrowly focused on choices among simple gambles to tell us much about risk-taking behavior in the flood plain or on the earthquake fault. (p. xxxi)

We are sympathetic to Slovic's sober assessment of the limited value that studying choices among simple gambles has for understanding people's behavior under risk. However, we believe that the discovery of the description–experience gap has led to interesting insights that, if they generalize beyond simple monetary gambles, may immensely benefit our understanding of potential obstacles to successful risk communication. So what are the potential implications of the description–experience gap? Our starting premise is that risk communication typically involves *descriptive* information.

The domains portrayed in our introductory examples differ in the extent to which receivers of risk communication have had the opportunity to personally experience the occurrence and nonoccurrence of the risk in question. In the domain of vaccination, parents—or, more generally, people considering vaccination—typically have no direct experience of the side effects of vaccination. They are thus blank slates, and communication about side effects and the associated probabilities will not compete with direct personal experience. In contrast, most residents of Naples have lifelong experience of living safe and sound in the shadow of the Vesuvius. They will evaluate warnings about the impending danger of a violent eruption against the backdrop of this experience. To predict the impact of a risk warning, communicators therefore need to take into account the receiver's degree of experience and the extent to which the warning may be at odds with this experience.

The issue is not just whether or not people have personal experience of a risk, however, but the rarity of that risk and, relatedly, the scope of personal experience. If the risk in question is a “black swan” (Taleb, 2007), a highly improbable event that occurs on average once in, say, 50 years, then a person may not know of its existence or may consider it to be less likely than it is. If, in contrast, a person has had opportunity to experience a very rare event—or, equally importantly, its nonoccurrence—through, for instance, administering thousands of vaccinations, his or her experienced sampled frequencies (“natural frequencies”; Gigerenzer & Hoffrage, 1995) may reflect the risks' objective frequencies quite accurately. But even considerable experience may result in underrepresentation of the rare risk due to recency (as mentioned above). Because improbable events are less likely to have occurred recently as compared with anytime during one's whole life, recency lessens the event's impact on people's choices. But the opposite also holds. Once a rare, high-impact event has recently occurred, recency is likely to give it a larger impact on behavior than its probability warrants. For instance, after four planes crashed in the terrorist attacks of September 11, 2001, many people stopped flying, at least temporarily, and instead drove some of the miles not flown. According to Gigerenzer's (2004) analysis (see also Gaissmaier & Gigerenzer, 2012), this response caused more people to lose their lives on the road by avoiding the risk of flying than were killed in the four fatal flights on 9/11.

Risk warnings do not operate in a vacuum. Sometimes people have experienced many safe encounters prior to receiving a warning (e.g., the repeated experience of unprotected sex without contracting a sexually transmitted disease); sometimes they receive the warning right after disaster has struck for the first time; sometimes they are blank slates with no experience at all. In all likelihood, how risk communication affects behavior depends on people's past experience. The few available studies that have investigated (focusing on monetary gambles) the relative impact of

description and experience when they co-occur suggest that experience tends to overpower description (Jessup et al., 2008; Lejarraga & Gonazlez, 2011). Without understanding the interplay of description and experience, scientists and policymakers will continue to be surprised by how ineffectual risk communication can be (see Barron, Leider, & Stack, 2008).

To date, we know of only one systematic investigation of this interplay that explicitly addressed the context of risk warnings and communication (Barron et al., 2008). The authors concluded that, even after being adequately warned, some people may continue to take risks simply because they have experienced good outcomes after making the same choice in the past (consistent with Jessup et al., 2008; Lejarraga & Gonazlez, 2011). One of their illustrations of the allure of positive experiences is the case of Vioxx, a nonsteroidal, anti-inflammatory drug developed by Merck & Co. When the drug was found to increase the risk of heart attack, an alert was first added to the package insert, and the drug was ultimately taken off the global market. Yet more than 2 million people continued to take the drug until their prescription ran out; presumably because they had taken it for quite some time without adverse consequences. Clearly, behavioral decision research has only just begun to appreciate that in many situations people can recruit descriptive or experienced-based information. What is now needed is a better understanding of the conditions under which these information sources are contradictory and of which source then gets the upper hand.

Let us briefly sketch one important line of future scientific inquiry. Personal experience is a powerful consultant (Weinstein, 1989). But some evolving 21st-century risks, such as the risks brought about by climate change, are “virtually impossible to detect from personal experience, amid the noise of random fluctuation around the central trend” (Weber & Stern, 2011, p. 318). Under such circumstances, personal experience offers the wrong advice and is possibly one key

driver behind the public (e.g., in the United States) lacking the willingness to take action against climate change. One key task for future research is therefore to find out whether and to what extent it may be possible to retrain the powerful consultant of personal experience through experience garnered in virtual realities (Weber & Stern, 2011). These learning environments offer the opportunity to experience various possible future climates—including catastrophic events that, although unlikely, deserve being taken into account: Today’s black swans may turn lighter in the future.

Conclusions

Really good things and really bad things happen infrequently. Most of us experience only one true love. Few of us get to graduate from Stanford or Oxford and become rocket scientists, brain surgeons, or CEOs; even fewer make it as movie stars. By the same token, few of us lose our life’s savings in a stock market crash, are rendered quadriplegic in an accident, or have a debilitating birth defect. The events of our brightest dreams and darkest nightmares tend to happen rarely. We can learn about such rare opportunities and dangers in at least two ways: (1) through symbolic representations or (2) through personal experience. Research on how people make risky choices has recently arrived at an important insight: These two types of information, description and experience, can prompt qualitatively different choices. Although the original findings pertained to choices among simple gambles, the description–experience gap also applies to the world outside decision scientists’ laboratories. We suggest that the description–experience gap is one key to a better understanding of why experts and the general public are often at odds with each other when reckoning with risks. It also explains why risk communications and expert warnings clothed in numbers (probabilities) and descriptions of possible outcomes lack persuasive power.

Suggested Additional Readings

- Hertwig, R. (in press). Decisions from experience. In G. Keren & G. Wu (Eds.), *Blackwell handbook of decision making*. Oxford, England: Blackwell.
- Weber, E. U. (2013). Doing the right thing willingly: Behavioral decision theory and environmental policy. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 380–397). Princeton, NJ: Princeton University Press.

References

- Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215–233. doi:10.1002/bdm.443
- Barron, G., Leider, S., & Stack, J. (2008). The effect of safe experience on a warning's impact: Sex, drugs, and rock-n-roll. *Organizational Behavior and Human Decision Processes*, *106*(2), 125–142. doi:10.1016/j.obhdp.2007.11.002
- Berger, C. R. (2007). A tale of two communication modes: When rational and experiential processing systems encounter statistical and anecdotal depictions of threat. *Journal of Language and Social Psychology*, *26*, 215–333. doi:10.1177/0261927X06303453
- Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, *53*, 155–167. doi:10.1016/j.jmp.2008.05.006
- Bloodbath at Luxor. (1997, November 20). *The Economist*. Retrieved from <http://www.economist.com/node/106278>
- 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/2003/08/26/world/san-giuseppe-vesuviano-journal-who-s-afraid-of-vesuvius-pompeii-is-history.html>
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 933–942. doi:10.1098/rstb.2007.2098
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, *112*, 912–931. doi:10.1037/0033-295X.112.4.912
- 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. doi:10.1002/bdm.683
- Erev, I., Glozman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty*, *36*, 153–177. doi:10.1007/s11166-008-9035-z
- Fox, C. R., & Hadar, L. (2006). “Decisions from experience” = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making*, *1*, 159–161.
- Gaissmaier, W., & Gigerenzer, G. (2012). 9/11, Act II: A fine-grained analysis of regional variations in traffic fatalities in the aftermath of the terrorist attacks. *Psychological Science*, *23*(12), 1449–1454. doi:10.1177/0956797612447804
- Gigerenzer, G. (2004). Dread risk, September 11, and fatal traffic accidents. *Psychological Science*, *15*(4), 286–287. doi:10.1111/j.0956-7976.2004.00668.x
- 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*, 623–629.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704. doi:10.1037/0033-295X.102.4.684
- Gilboa, I., & Schmeidler, D. (1995). Case-based decision-theory. *Quarterly Journal of Economics*, *110*, 605–639. doi:10.2307/2946694
- Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review*, *118*, 523–551. doi:10.1037/a0024558
- Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*, 317–325.
- Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making*, *23*, 48–68. doi:10.1002/bdm.665
- Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description–experience gap in risky choice: The role of sample size and experienced

- probabilities. *Journal of Behavioral Decision Making*, 21, 493–518. doi:10.1002/bdm.598
- Hertwig, R. (2012). The psychology and rationality of decisions from experience. *Synthese*, 187, 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, 534–539. doi:10.1111/j.0956–7976.2004.00715.x
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). Decisions from experience: Sampling and updating of information. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition*. New York, NY: Cambridge University Press.
- Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13, 517–523. doi:10.1016/j.tics.2009.09.004
- Hertwig, R., Pachur, T., & Kurzenhäuser, S. (2005). Judgments of risk frequencies: Tests of possible cognitive mechanisms. *Journal of Experimental Psychology: Learning Memory and Cognition*, 35, 621–642. doi:10.1037/0278–7393.31.4.621
- 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. 209–235). Oxford, England: Oxford University Press.
- Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, 115, 225–237. doi:10.1016/j.cognition.2009.12.009
- Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science*, 12, 1787–1792. doi:10.1177/0956797610387443
- Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology*, 24, 1–55. doi:10.1016/0010–0285(92)90002-J
- Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science*, 19, 1015–1022. doi:10.1111/j.1467–9280.2008.02193.x
- Kahneman, D., & Tversky, A. (1979). Prospect theory: Analysis of decision under risk. *Econometrica*, 47(2), 263–291. doi:10.2307/1914185
- Klein, G. A. (1999). *Sources of power: How people make decisions*. Cambridge: MIT Press.
- Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, 116(2), 286–295.
- Lichtenstein, S., Slovic, P., Fischhoff, B., Layman, M., & Combs, B. (1978). Judged frequency of lethal events. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 551–578. doi:10.1037/0278–7393.4.6.551
- Lopes, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 137–144. doi:10.1037/0096–1523.9.1.137
- 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), e20262. doi:10.1371/journal.pone.0020262
- Mastrolorenzo, G., Petrone, P., Pappalardo, L., & Sheridan, M. F. (2006). The Avellino 3780-yr-B.P. 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, 4366–4370. doi:10.1073/pnas.0508697103
- Omer, S. B., Pan, W. K., Halsey, N. A., Stokley, S., Moulton, L. H., Navar, A. M., . . . Salmon, D. A. (2006). Nonmedical exemptions to school immunization requirements. *JAMA*, 296(14), 1757–1763. doi:10.1001/jama.296.14.1757
- Once more to the gallows. (2012, November 24). *The Economist*. Retrieved from <http://www.economist.com/news/asia/21567104-renewed-unwelcome-enthusiasm-execution-once-more-gallows>
- Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes*, 106, 168–179. doi:10.1016/j.obhdp.2008.02.001
- Real, L. A. (1991). Animal choice behavior and the evolution of cognitive architecture. *Science*, 253, 980–986. doi:10.1126/science.1887231
- Real, L. A. (1992). Information processing and the evolutionary ecology of cognitive architecture. *American Naturalist*, 140, 108–145. doi:10.1086/285399
- Silver, N. (2012). *The signal and the noise: Why so many predictions fail—but some don't*. New York, NY: Penguin Books.
- Slovic, P. (1987). Risk perception. *Science*, 236(4799), 280–285. doi:10.1126/science.3563507
- Slovic, P. (2000). *The perception of risk*. London, England: Earthscan.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. New York, NY: Random House.

- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. doi:10.1007/BF00122574
- Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, 20, 473–479. doi:10.1111/j.1467-9280.2009.02319.x
- Weber, E. U., & Kirsner, B. (1996). Reasons for rank-dependent utility evaluation. *Journal of Risk and Uncertainty*, 14, 41–61. doi:10.1023/A:1007769703493
- Weber, E. U., Shafir, S., & Blais, A.-R. (2004). Predicting risk-sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review*, 111, 430–445. doi:10.1037/0033-295X.111.2.430
- Weber, E. U., & Stern, P. C. (2011). Public understanding of climate change in the United States. *American Psychologist*, 66(4), 315–328. doi:10.1037/a0023253
- Weinstein, N. D. (1989). Effects of personal experience on self-protective behavior. *Psychological Bulletin*, 105(1), 31–50. doi:10.1037/0033-2909.105.1.31
- Whittle, P. (1988). Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25A, 287–298. doi:10.2307/3214163
- Wilford, J. N. (2006, March 7). Long before burying Pompeii, Vesuvius vented its wrath. *The New York Times*. Retrieved from <http://www.nytimes.com/2006/03/07/science/07volc.html>
- Yechiam, E., Barron, G., & Erev, I. (2005). The role of personal experience in contributing to different patterns of response to rare terrorist attacks. *Journal of Conflict Resolution*, 49, 430–439. doi:10.1177/0022002704270847
- Yechiam, E., & Busemeyer, J. R. (2006). The effect of foregone payoffs on underweighting small probability events. *Journal of Behavioral Decision Making*, 19, 1–16. doi:10.1002/bdm.509