

## Decisions From Experience and *Statistical Probabilities*: Why They Trigger Different Choices Than a Priori Probabilities

ROBIN HAU<sup>1\*</sup>, TIMOTHY J. PLESKAC<sup>2</sup> and RALPH HERTWIG<sup>1</sup>

<sup>1</sup>*University of Basel, Switzerland*

<sup>2</sup>*Michigan State University, USA*

### ABSTRACT

The distinction between risk and uncertainty is deeply entrenched in psychologists' and economists' thinking. Knight (1921), to whom it is frequently attributed, however, went beyond this dichotomy. Within the domain of risk, he set apart a priori and statistical probabilities, a distinction that maps onto that between decisions from description and experience, respectively. We argue this distinction is important because risky choices based on a priori (described) and statistical (experienced) probabilities can substantially diverge. To understand why, we examine various possible contributing factors to the description–experience gap. We find that payoff variability and memory limitations play only a small role in the emergence of the gap. In contrast, the presence of rare events and their representation as either natural frequencies in decisions from experience or single-event probabilities in decisions from description appear relevant for the gap. Copyright © 2009 John Wiley & Sons, Ltd.

**KEY WORDS** decisions; experience; information representation; rare events; risk and uncertainty; risky choice; sampling

### INTRODUCTION

In the early 20th century, Frank Hyneman Knight drew a conceptual map that has, to this day, shaped psychologists' and economists' thinking about the world of uncertainty in which we live. In his book *Risk, Uncertainty, and Profit*, Knight (1921, p. 224) distinguished between different types of “probability situation.” The two situations that have attracted most attention are *risk* and *uncertainty* (see e.g., Edwards, 1954; Epstein & Wang, 1994; Lopes, 1983; Luce, 2000; Luce & Raiffa, 1957). Risk has typically been used to describe situations in which “probabilities are available to guide choice,” and uncertainty to describe situations in which “information is too imprecise to be summarized adequately by probabilities” (Epstein & Wang, 1994, p. 283), or, as Tversky and Fox (1995) put it, “Following Knight (1921), decision theorists

---

\* Correspondence to: Robin Hau, Department of Psychology, University of Basel, Missionsstrasse 60/62, 4055 Basel, Switzerland.  
E-mail: robin.hau@unibas.ch

distinguish between risky (or chance) prospects where the probabilities associated with outcomes are assumed to be known, and uncertain prospects where these probabilities are not assumed to be known” (p. 270). This dichotomy has now become deeply ingrained in economic and psychological theory.

Although dichotomies promise to bring order to chaos, they do so at the cost of being simplistic (Barbe, 2001). Sometimes they are even suspected of impeding theoretical progress (Newell, 1973). Knight’s (1921) view, upon re-examination, was subtler than the simple risk versus uncertainty dichotomy (Runde, 1998). Rather than a dichotomy, he proposed a trichotomy of situations, which differ with regard to how easy or difficult it is to determine the likelihood of an event in question. Specifically, he distinguished between situations where people can use (a) *a priori probabilities*, (b) *statistical probabilities*, or (c) *estimates* (see pp. 224 and 225). *A priori probabilities* refer to situations where events result from a precisely known random mechanism, so that the probability of an event occurring can easily be assigned via mathematical calculation, as in deducing the probability of getting a six when throwing a fair die. *Statistical probabilities*, in contrast, refer to situations where probabilities cannot be calculated exactly, but assessed in an empirical manner. Consider Lopes’ (1983) example of a brewer who wishes to know how likely a particular production method yields a broken bottle. A reasonable method is for the brewer to treat this risk actuarially and approximate the probability of a breakage from an empirical sample of past events. If, out of 300 bottles, the procedure produced one broken one, the brewer would say the probability of a broken bottle using this procedure is 1/300.

Finally, Knight’s term “*estimates*” refers to situations of utter uncertainty, in which “there is *no valid basis of any kind* for classifying instances” (Knight, 1921, p. 225) or, in other words, where events are truly unique and probability cannot be meaningfully applied or empirically derived. Lopes’ (1983) example of the *Miller Brewing Company* deciding whether to produce Lite beer in the 1970s is an excellent example of such an estimate. Who knew whether beer-drinking Americans would accept low-calorie beer? Miller could merely speculate on—or, in Knight’s terms, estimate—the beer’s success.

All three situations have interesting psychological implications. Situations involving *a priori probabilities*, for instance, raise the question of whether and how outcomes and their explicit probabilities are subjectively evaluated (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) and processed (Brandstätter, Gigerenzer, & Hertwig, 2006, 2008; Tversky, 1969). Situations involving estimates highlight the role that a person’s knowledge plays in forming a belief that ultimately can carry the decision (Fox & Tversky, 1998; Tversky & Fox, 1995). Finally, the psychological analyses of situations involving statistical probabilities, the focus of this article, benefit from a framework that takes two interlocking components into account: The cognitive algorithms and mental limitations of the human mind and the statistical structure of the environment (of risky gambles) in which the mind operates. In other words, we suggest that a cognitive-ecological framework can help organize findings and theories in research on statistical probabilities and decisions from experience.

## A COGNITIVE-ECOLOGICAL FRAMEWORK AND THE DESCRIPTION–EXPERIENCE GAP

Notwithstanding Knight’s (1921, p. 215) emphasis on the ubiquity of statistical probabilities in the real world, countless investigations in psychological and economic research on *decision making under risk* have focused on choices between simple monetary gambles and probabilities akin to those involved in the roll of a die. The kinds of probabilities commonly studied in risky choice are *a priori* and not statistical in nature. People are told to consider, for instance, the choice between a prospect that offers a 10% chance to win \$32 (with a 90% chance to win nothing) and the alternative of receiving \$3 for sure. This completely specified chance event was for many decades the canonical situation in our laboratories (see Weber, Shafir, & Blais, 2004).

Recently, however, a new experimental paradigm involving statistical rather than *a priori* probabilities has attracted researchers’ attention. The simplest version of this paradigm presents respondents with two gambles

on a computer screen represented by boxes. Each gamble contains a set of outcomes that occur with some probability. People are not told anything about the properties of the gambles, but are allowed to explore each by sampling from them. Specifically, clicking on a box triggers a random draw of an outcome from the associated set of outcomes. People are encouraged to sample until they feel confident enough to decide which box is “better,” in the sense that they would prefer to draw from it during a final trial involving real monetary payoffs.

This paradigm creates a situation in which respondents can approximate the probability of each outcome occurring from their empirical samples.<sup>1</sup> In Knight’s (1921) view, they thus deal with statistical probabilities, and make, in Hertwig, Barron, Weber, and Erev’s (2004) terminology, *decisions from experience*.

In studies using this paradigm, a “description–experience gap” (Hau, Pleskac, Kiefer, & Hertwig, 2008) has emerged. The gap refers to the observation that respondents who make decisions from experience involving rare events appear to have systematically different preferences than respondents who make *decisions from description* based on the true properties of each gamble, or, in Knight’s (1921) terms, based on a priori probabilities. For example, across six problems, Hertwig et al. (2004) found that the average (absolute) difference between the choice proportions of an experience and a description group was 36 percentage points. This description–experience gap has been replicated in five other studies that used the same or similar paradigms. Table 1 shows the magnitude of the description–experience gap in those studies<sup>2</sup>, and that in the majority of problems the same key result emerged (rightmost column): Although rare events in decisions from description appear to be overweighted (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), in decisions from experience rare events receive less impact than they deserve according to their objective probabilities. This means that gambles involving a desirable rare event (like a 10% chance to win 32 points) were chosen more often in decisions from description than in decisions from experience, whereas gambles involving an undesirable rare event (like a 10% chance of not winning anything when every other outcome represents a gain) were chosen less often in decisions from description than in decisions from experience.

Why does the description–experience gap arise? From the viewpoint of a cognitive-ecological framework (e.g., Fiedler & Juslin, 2006), a number of factors can be implicated in the gap’s emergence. Specifically, preferences in decisions from experience can be shaped by transitions between (at least) three different levels of representations: (a) the population level describing the ecological but unknown (to the person making the choice) properties of each gamble; (b) the sample level describing the information about each gamble as revealed in the sample drawn by the person making the choice; and (c) the cognitive level portraying the sampled information after it was filtered through the mind’s learning, memory, and valuation processes. A potential fourth level of representation would be the neural level (see for example Sugrue, Corrado, & Newsome, 2005).

In what follows, we focus on three factors that may contribute to the description–experience gap and that belong to one of the three levels of representation, namely, payoff variability (population level), sampling error (sample level), and memory capacity (cognitive level). We first turn to sampling error because it has been proposed as the key determinant of the description–experience gap.

### Sampling error

An immediate cause of the gap rests in the transition from the population level to the sample level representation. Knight (1921) anticipated the imperfection of this transition when he stated that “the statistical treatment never gives closely accurate quantitative results” (p. 215). That is, the transition from the

<sup>1</sup>Barron and Erev (2003; Erev & Barron, 2005) employed a similar paradigm, but exploration and exploitation coincided so that each sampled observation carried information value as well as monetary value.

<sup>2</sup>Gottlieb, Weiss, and Chapman (2007) also observed a difference between description- and experience-based choices. Unfortunately, the majority of their problems do not allow a clear prediction about which option should be preferred depending on the impact of rare events. It is therefore not included in Table 1.

Table 1. The reality of the description–experience gap

Authors	Study/condition	No. of problems	Sample size <sup>1</sup> (median)	Description–experience gap <sup>2</sup> (%)	Correct predictions <sup>3</sup>
Hertwig, Barron, Weber, & Erev, 2004	Experiment 1 vs. Experiment 2	6	15	36	6/6
Hau, Pleskac, Kiefer, & Hertwig, 2008		6	11	27	6/6
		6	33	13	6/6
Weber, Shafir, & Blais, 2004	Experiment 1 vs. Experiment 2	6	Set to 100 <sup>4</sup>	15	5/6
Rakow, Demes, & Newell, 2008	Objective description	5	17	38	2/2
	Percentage sample description <sup>5</sup>	12	15	24	11/12
	Frequency sample description <sup>5</sup>	12	15	6	9/12
Ungemach, Chater, & Stewart, 2009	Free sampling	12	15	5	7/12
	Matched sampling <sup>6</sup>	6	19	31	6/6
		6	Set to 80 <sup>4</sup>	19	5/6

<sup>1</sup>Sample size is reported as the total number of draws from both options combined.

<sup>2</sup>The description–experience gap is measured as the mean difference between choice proportions in the description and experience conditions for those problems including at least one rare event. Differences in an unexpected direction (i.e., rare events having a larger impact on decisions from experience) are deducted.

<sup>3</sup>This column indicates how many of the differences between the description- and experience conditions are consistent with the thesis that rare events have less impact on decisions from experience than decisions from description (Hertwig et al., 2004); only problems for which a clear prediction can be derived are considered.

<sup>4</sup>Termination of search was not determined by respondents but by experimenters.

<sup>5</sup>Description and experience groups were yoked, thus eliminating sampling error as an explanation for the gap.

<sup>6</sup>Samples were chosen such as to reflect the objective probabilities, thus eliminating sampling error.

population to the sample introduces error via the sampling process itself, and consequently statistical probabilities will not be as accurate as a priori probabilities. This measurement error will, of course, be larger, the smaller the sample. Indeed, Hertwig et al. (2004) observed that the typical number of draws respondents made was approximately seven from each payoff distribution. As Table 1 shows, subsequent studies observed similarly small samples.

As Hertwig et al. (2004) pointed out, sampling error stemming from small samples can account for rare events receiving less impact than they deserve according to their objective probabilities: Owing to the skewness of the binomial distribution with low  $p$ , small  $ns$  make it more likely that a person encounters rare events less frequently than expected ( $n^*p$ ) (for details, see Hertwig, Barron, Weber, & Erev, 2006). That is, rare events are more likely to be undersampled than oversampled in small samples, and as a result the psychological impact of rare events is attenuated. In addition, with small samples, the range of probabilities that can be conveyed is restricted. For instance, a sample of size seven can never truly reflect a probability of .1.

Although sampling error undoubtedly is crucially implicated in the emergence of the description–experience gap (e.g., Rakow, Demes, & Newell, 2008), it is not the whole story. Hau et al. (2008) have provided a comprehensive investigation of the impact of sample size and undersampling of rare events. By encouraging participants to take larger samples (either via instruction or increased monetary stakes), they found the description–experience gap to be narrower with larger sample sizes. Yet, even with a sample as large as 100 draws (and very accurate probability estimates), there still remained a marked gap. Ungemach, Chater, and Stewart (2009) arrived at the same observation using a sample of 80 draws and experienced probabilities that mimicked the described probabilities (Table 1).

### Memory limits

Why does extensive experience not eliminate the gap? In the transition from the sample representation to the cognitive representation, one intermediary process is memory which can affect how—and how much of—the sampled information is fed into the choice process (e.g., Dougherty, Gettys, & Ogden, 1999; Gigerenzer, Hoffrage, & Kleinbölting, 1991; Pleskac, 2007; Schooler & Hertwig, 2005; Sedlmeier, 2002). Limits in memory (Cowan, 2001; Miller, 1956) are one obvious cause underlying the description–experience gap: If a person were able to only hold a small subset of sampled events in working memory, then this limitation—regardless of how extensive a person's experienced sample is—would result in a functionally small sample from which choice is derived. Thus, limited working memory capacity could produce the same phenomenon that genuinely small samples do: Undersampling of rare events (for a similar argument see Kareev, 2000; Kareev, Lieberman, & Lev, 1997).

Typically, samples that people hold in memory are not randomly drawn from their total experience but are skewed toward recent experiences. Such a recency effect could also contribute to description–experience gap. Previous studies, however, have found mixed evidence regarding the occurrence of a recency effect in decisions from experience: Hertwig et al. (2004) observed that the most recent half of sampled outcomes predicted ultimate choices markedly better than the first half of respondents' experiences. In contrast, Hau et al. (2008) found no strong recency effects. Rakow et al. (2008) only found recency effects when sampling was self-enacted rather than imposed on respondents. Yet, they also reported that the impact of recency was not significantly related to participants' working memory capacity. Let us finally turn to the first level of representation, the population level, which captures the objective properties of the gambles—properties that decision makers can only peek at through the window of their samples.

### Payoff variability

One objective property of monetary gambles is the variability of the payoffs, which in technical terms is defined as the pooled standard deviation of the average outcome in a particular pair of gambles. Interestingly,

several decision rules postulate that choice between gambles is co-determined by the variability of their payoffs (Allais, 1953; Busemeyer & Townsend, 1993; Erev & Barron, 2005; Markowitz, 1959). For example, the cognitive processes assumed by decision field theory (Busemeyer & Townsend, 1993) imply that the probability of choosing gamble *A* over gamble *B* is an increasing function of the ratio

$$\frac{v_A - v_B}{\sqrt{\delta_A^2 + \delta_B^2}}$$

where  $v_A$  and  $v_B$  are the expected values and  $\delta_A^2$  and  $\delta_B^2$  are the variances of the payoffs in gambles *A* and *B*, respectively. From this assumption follows that as payoff variability increases, the predicted choice proportions approach .5. Studies that have lent support to this prediction have relied on statistical (experienced) rather than a priori (described) probabilities (Busemeyer, 1985; Erev & Barron, 2005; Katz, 1962; Myers & Katz, 1962; Myers & Sadler, 1960). This raises the question of whether the effect of payoff variability is format dependent, and only influences decisions from experience but not decisions from description. If so, then this format dependency of the payoff variability effect would provide one novel candidate explanation for the description–experience gap.

For illustration, consider Problems 1 and 2 investigated by Hertwig et al. (2004):

<i>Problem 1</i>			<i>Problem 2</i>		
A [ $\sigma = 1.6$ ]	4	with $p = .80$	C [ $\sigma = 1.6$ ]	4	with $p = .20$
	0	with $p = .20$		0	with $p = .80$
B [ $\sigma = 0$ ]	3	with $p = 1.00$	D [ $\sigma = 1.3$ ]	3	with $p = .25$
					with $p = .75$

In the experience group, a majority (88%) selected the risky option *A* in Problem 1, relative to a minority in the description group (36%). In Problem 2, the minority of the experience group (44%) selected option *C*, relative to the majority in the description group (64%). Problems 1 and 2 offer gambles with identical possible earnings, but the probability of winning is scaled down (by a factor of four), thus reducing a certain option to a chance prospect. In the description group, the change in preference from Problems 1 to 2 represents a violation of the substitution axiom of utility theory and has been taken as evidence for prospect theory's assumed overweighting of rare events (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). At the same time, the opposite preference pattern in the experience group is consistent with the attenuated impact of rare events, namely, event “0” (Problem 1) and “4” (Problem 2), respectively.

Alternatively, however, the change in preference in the experience group might have been triggered by the increased payoff variability in Problem 2 (relative to Problem 1):  $\sqrt{1.6^2 + 1.3^2} = 2.1$  relative to  $\sqrt{1.6^2 + 0^2} = 1.6$ . As payoff variability has been assumed to render choice more random, choices in the experience group may have simply become more random from Problem 1 to Problem 2 (88 to 44%). Payoff variability can also explain the change in preference between Problem 5 (32 with a probability of .1 or 3 for certain) and Problem 6 (32 with a probability of .025 and 3 for certain) in Hertwig et al. (2004). Several investigations of decisions from experience (Fox & Hadar, 2006; Hau et al., 2008; Jessup, Bishara, & Busemeyer, 2008; Rakow et al., 2008; Ungemach et al., 2009) have used the same set of problems or included some or all of the problems investigated by Hertwig et al. (2004) in their problem set. Consequently, payoff variability may be a previously unsuspected culprit behind a larger set of findings in research on decisions from experience.

## OUR STUDIES

The description–experience gap has been explained in terms of sampling error and memory limits (e.g., recency). **We have introduced another possible explanation that is related to an objective property of gambles,**



namely, payoff variability. The goal of our investigations is to analyze the relative contribution of each of these factors to the gap. In Study 1, we manipulated the occurrence of rare events and payoff variability of the gambles independently. We could do so easily, by slightly changing the previously used decisions from experience paradigm. Respondents were presented with one described certain outcome, and an unknown gamble from which they could sample (see also Busemeyer, 1985). This set-up had the additional advantage of giving us greater control over a person's sampling process: It allowed us to study how a person's choices alter as a function of the increasing experience about *one* gamble, and to aggregate choices from different respondents because they acquired the same amount of experience.

Still another goal of Study 1 is to investigate the effect of memory limits by removing the need to keep one's sampled experience in memory—a condition that we call *decisions from records*. If decisions from records deviate from decisions from experience (given otherwise identical information), this may imply that memory limits contribute to the description–experience gap. In Study 2, we addressed another factor that may contribute to the gap and that is located in the cognitive pillar of the cognitive-ecological framework. Even if information in decisions from description and experience were mathematically equivalent, it might not be so psychologically. Specifically, we examined the question to what extent a priori probabilities and statistical probabilities are psychologically different, given that the latter tend to be experienced in terms of frequencies and not probabilities.

## STUDY 1: DO PAYOFF VARIABILITY AND MEMORY LIMITS MATTER?

### Method

#### Design

Study 1 had four independent variables: Representation format (experience, records, and description) varied between participants, payoff variability (3 levels), probability of winning (4 levels), and sample-size (5 levels) varied within participants.

#### Participants

A total of 120 students from the University of Basel participated. Forty participants were each randomly assigned to one of three groups: *Experience*, *records* and *description*.

#### Stimuli

We constructed 12 decision problems (Table 2) that allowed us to systematically increase payoff variability in one option (the risky option), while keeping it constant in the second option (the certain option). The risky option in each problem yields  $x$  Swiss Francs with probability  $p$  and 0 with probability  $1-p$ . Problems represent all combinations of four levels of probability  $p$  (.1, .2, .8, .9) and three levels of payoff variability ( $\sigma = 1.60, 4.50$ , and  $9.60$  Swiss Francs). The probabilities are chosen such that either gain  $x$  or outcome “0” represents a low-probability event.

The certain option tracked the expected value of the risky option, except that for half of the participants the value of the certain option was reduced by 0.20 Swiss Francs and for the other half of participants the value was augmented by 0.20 Francs. Respondents received course credit and, in addition, were informed that they would be paid according to one of their decisions, randomly selected after all choices were completed. Respondents' payoffs ranged between 0 and 32 Swiss Francs. The stimuli were presented on a PC running e-prime 1.1.

Table 2. The 12 choice problems used in studies 1 and 2

No.	$p_{\text{win}}^1$	$X_{\text{win}}^1$	$\sigma^1$	Certain option <sup>2</sup>
1	10	5.30	1.60	0.30/0.70
2	20	4	1.60	0.60/1.00
3	80	4	1.60	3.00/3.40
4	90	5.30	1.60	4.60/5.00
5	10	15	4.50	1.30/1.70
6	20	11.30	4.50	2.10/2.50
7	80	11.30	4.50	8.80/9.20
8	90	15	4.50	13.30/13.70
9	10	32	9.60	3.00/3.40
10	20	24	9.60	11.80/12.20
11	80	24	9.60	19.00/19.40
12	90	32	9.60	28.60/29.00

<sup>1</sup>Each problem consists of a risky option—represented by a deck of cards—paying  $X_{\text{win}}$  Swiss Francs with probability  $p_{\text{win}}$ , 0 otherwise, coupled with a certain option. Payoff variability is indicated by the standard deviation  $\sigma$  of the risky option.

<sup>2</sup>For half of the participants, the certain option was the expected value of the risky option augmented by 0.20 Swiss Francs; for the other half, it was reduced by the same amount; 1 Swiss Franc = US \$0.95.

### Procedure

In all three groups, respondents selected their preferred option by pressing one of two buttons on a standard keyboard. Options were displayed on the monitor, divided by a central vertical line, with the position of risky and certain option randomly determined. The order of the problems was also randomized. In the description group, respondents were presented with a description of the risky option (e.g., 4 with a probability of 80%, 0 otherwise) displayed opposite to that of the certain option (e.g., 3 with certainty).

In the experience group, respondents initially did not know the outcomes and associated probabilities in the risky option. Before choosing between the risky and certain option, they were instructed to explore the risky option—graphically represented in terms of a deck of cards—by sampling from it. Pressing a key elicited the sampling of a card from this deck. The card, with the outcome printed on it, was then displayed on the screen until the next card replaced it. After the first five draws, the deck of cards was displayed opposite a written description of the certain option (i.e., 3 with certainty). Participants were asked to make a choice. Subsequently, they were instructed to continue sampling. Again, after five draws, a choice was requested. In total, participants made five choices per problem, having sampled 5, 10, 20, 35, and 50 cards, respectively. This manipulation enabled us to track—within each person and problem—how choice changes as a function of growing experience. The sequences of cards respondents experienced were constructed as follows: For each deck, we created 20 random sequences of 50 cards using a computerized random-generator. Each sequence was employed for two respondents, one whose certain option was augmented, and one whose certain option was reduced, relative to the expected value of the risky option.

The records group was identical to the experience group, except for two differences. First, each sampling from the deck yielded an episode comprising several outcomes rather than one. Thus, prior to each choice a respondent would receive all new information simultaneously. Second, prior to their choices (as in the experience group, five per problem) all outcomes that had been sampled up to that point were displayed in front of them. That is, records of all past experiences were openly accessible, thus reducing—or even eliminating—the burden on memory during decision making. For illustration, Figure 1 depicts a shot of the screen encountered by a respondent who had experienced 35 outcomes from the risky deck of cards in Problem 9 (Table 2), and who on the next screen would be asked to make a choice. The records group saw exactly the same random sequences as the experience group.



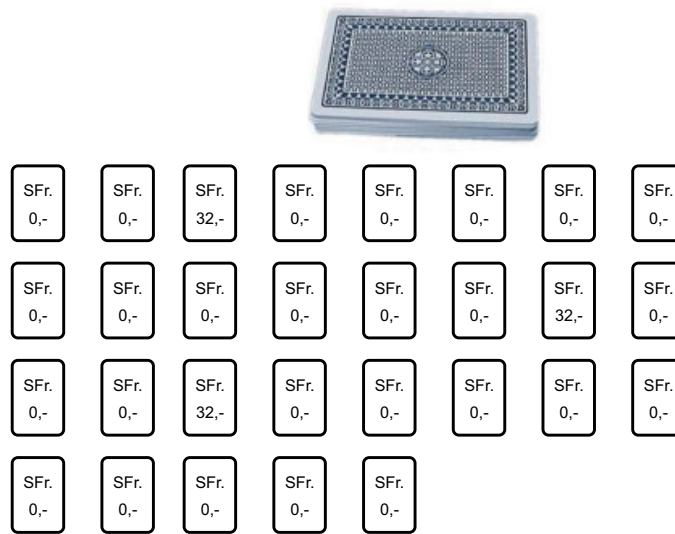


Figure 1. Screenshot displaying a sample of 35 cards in Problem 9 as seen by a respondent in the records group

## Results and discussion

Although the statistical analyses reported here were conducted using all five levels of sample size, for reasons of simplicity we focus on the choices made after 5 and 50 draws, respectively. All subsequent conclusions drawn from comparing these two extreme sample sizes, however, are consistent with the trends that emerge for the intermediary sample sizes (see Appendix for a detailed summary of choices across all sample sizes).

As a first step, we checked whether we were able to replicate two findings of previous studies. First, the risky option of our Problem 3 (4; .8) has been studied by Hertwig et al. (2004) and Hau et al. (2008). In both studies, the majority of respondents making decisions from experience preferred the risky option to a certain option of similar expected value. We found similar preferences in our experience and records groups: 55 and 62% respondents chose the risky option in the experience and records groups, respectively. Secondly, Hertwig et al. and Hau et al. found the risky option of Problem 9 (32; .1) to be less preferred than the certain option. Echoing this result, we found that merely 27 and 32% of respondents in the experience and records groups chose the risky option. **These findings suggest that our new method yielded results similar to those observed previously.**

We now turn to the extent to which more experience reduces the description–experience gap and then examine the impact of payoff variability, and memory limits. We should point out that we conducted all of our statistical analyses by conditioning on one effect. So, to examine the effect of rare events we formed choice proportions by averaging across the different levels of variability. To examine the effect of payoff variability, we averaged across different levels of probability to win.

### *The role of more experience*

Respondents in the experience and records groups, respectively, chose between the risky and certain option after an experience of 5, 10, 20, 35, and 50 draws. As can be expected from the law of large numbers, a larger  $n$  resulted in a more veridical experience of the outcomes' probabilities in the risky option. When, for example, people sampled  $n = 5$  times, the mean absolute difference between the objective and experienced probabilities was 12.3 (median = 10) percentage points, relative to 8.2 ( $n = 10$ , md = 10), 5.5 ( $n = 20$ , md = 5), 4.8 ( $n = 35$ , md = 4.3), and finally 4.1 percentage points (md = 4) in the case of  $n = 50$  sampled outcomes.

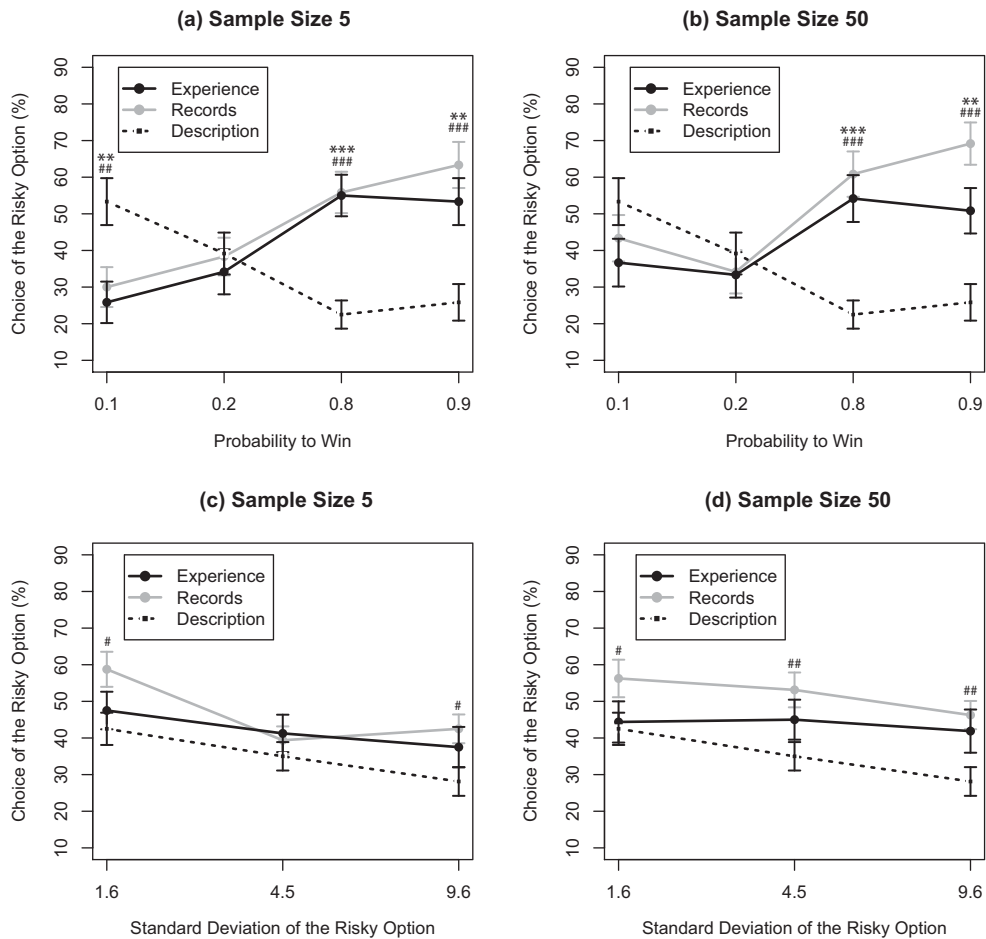


Figure 2. Mean proportion of choices of the risky option in the description, experience, and records groups as a function of probability to win based on a sample of size 5 (Figure 2a) and size 50 (Figure 2b), respectively, and as a function of payoff variability based on a sample of size 5 (Figure 2c) and 50 (Figure 2d), respectively. Error bars indicate  $\pm 1$  SE of the mean. For every level of sample size and at every level of probability or payoff variability, significant differences of choice proportions between decisions from experience and decisions from description are marked by asterisks ( $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ ); significant differences between decisions from records and decisions from description are marked by hash marks ( $#p < .05$ ;  $##p < .01$ ;  $###p < .001$ )

Does more experience alter choice? The upper row of Figure 2 shows the proportion of individuals that selected the risky option as a function of probability  $p$  and of sample size (Figure 2a: 5 draws; Figure 2b: 50 draws) averaging across levels of payoff variability. The lower row shows the choice proportion as a function of payoff variability and sample size (Figure 2c: 5 draws; Figure 2d: 50 draws) averaging across probability  $p$  (we return to the effects of probability shortly). These figures illustrate that the impact of additional experience was small. The mean absolute difference for choice proportions between the experience and the description groups was 24, 23, 23, 21, and 20 percentage points, respectively, for  $n = 5$ ,  $n = 10$ ,  $n = 20$ ,  $n = 35$ , and  $n = 50$  samples. That is, in this new set of problems we again found a large and systematic experience–description gap that shrinks with more experience. The reduction, however, was small given that the difference between objective and experienced probabilities shrank from

12.3 to 4.1 percentage points. Although sampling error is an important contributor to the gap, other factors play a role as well.

Two ANOVAs shed light on the gap: A  $5 \times 3 \times 4$  “rare event ANOVA” and a  $5 \times 3 \times 3$  “payoff variability ANOVA.” Both ANOVAs had sample size (5 levels, within-subjects) and group (3 levels, between-subjects) as factors. The rare-event ANOVA treated probability to win as a factor (4 levels, within-subjects) while averaging across variability, whereas the payoff-variability ANOVA treated payoff variability (3 levels, within-subjects) as a factor while averaging across levels of probability. In both ANOVAs, the main effect of sample size was not significant ( $F_{\text{rare events}}(4, 468) = 1.06, p = .38, \eta^2 = .009$ ;  $F_{\text{payoff variability}}(4, 468) = 0.86, p = .49, \eta^2 = .007$ ).

### *The role of rare events*

The differences in choices in Figures 2a and 2b are consistent with the implied opposite impact of the rare event in decisions from description and experience, observed by Hertwig et al. (2004). Respondents in the experience and records groups preferred the certain to the risky option when the probability to win was small ( $p = .1$ ), suggesting that this desirable but unlikely gain had less impact than in the description group, which arrived at the opposite preference. The pattern reversed when  $p$  became large ( $p = .8$  and  $p = .9$ ), and by extension, the rare event proved to be the undesirable “0” outcome. Then, the description group preferred the certain option, whereas the experience group was more likely to choose the risky option, although the strength of their preference for the risky option was modest (i.e., close to 50%). The only exception from this general pattern of significantly different choice proportions between description and experience occurred at  $p = .2$ , for which all three groups expressed a similar preference.

The rare-event ANOVA confirms that rare events influenced the groups’ choices differentially. The interaction between probability to win and group is significant ( $F_{\text{rare events}}(6, 351) = 13.46, p < .001, \eta^2 = .187$ ). This pattern is not significantly influenced by sample size as the non-significant interaction between the probability to win  $\times$  group  $\times$  sample size implies,  $F_{\text{rare events}}(24, 1404) = 1.00, p = .46, \eta^2 = .017$ .<sup>3</sup>

### *The role of payoff variability*

Figures 2c and 2d plot the choice proportions of the risky option as a function of payoff variability. Across all three groups, there is a tendency to choose the certain option the larger the payoff variability in the risky option, regardless of sample size (compare Figures 2c and 2d). This main effect is confirmed by the payoff-variability ANOVA ( $F_{\text{payoff variability}}(2, 234) = 10.31, p < .001, \eta^2 = .081$ ), which finds no interaction of payoff variability and group ( $F_{\text{payoff variability}}(4, 351) = 0.05, p = .48, \eta^2 = .015$ ) or of payoff variability, group and sample size ( $F_{\text{payoff variability}}(16, 936) = 0.04, p < .43, \eta^2 = .017$ ). This lack of interaction suggests that payoff variability is not likely to be the source of the description–experience gap: It does not influence the two kinds of decisions differentially. Therefore, and also because rare events did in fact have a differential impact, we will focus on the probability to win in all further analyses.<sup>4</sup>

<sup>3</sup>Note that the change in preference with rare events in the experience and records groups is also consistent with Weber et al.’s (2004) phenomenological use of the coefficient of variation to predict risk taking. Specifically, as the coefficient of variation increases people become less risk seeking and choose the certain option. Due to the procedures we used to construct our gambles a change in coefficient of variation is perfectly correlated with a change in the rare event. Because our focus is on the process level, we interpret the effect in terms of the effect of rare events. However, future research should experimentally investigate these different concepts.

<sup>4</sup>The conclusion that the rare event—not payoff variability—matters for the description–experience gap is further corroborated by an ANOVA in which we collapsed across all sample sizes and included payoff variability, probability to win and group as factors. There were no significant variability  $\times$  group ( $F(4, 234) = 0.79, p = .54, \eta^2 = .013$ ), variability  $\times$  probability ( $F(6, 702) = 1.30, p = .25, \eta^2 = .011$ ), or variability  $\times$  probability  $\times$  group ( $F(12, 702) = 0.07, p = .53, \eta^2 = .015$ ) interactions, but there was a significant probability  $\times$  group interaction ( $F(6, 351) = 13.46, p < .001, \eta^2 = .187$ ).

### *The role of memory limits*

The experience and the records groups encountered the same sequences of outcomes. The key difference between the groups was that people in the records group did not need to rely on their memory to evaluate the risky option. Yet, choices in both groups are surprisingly similar (Figure 2). This may not be surprising when the sample size is small (Figure 2a). The pattern, however, also holds when people had to commit a large sequence of 50 draws to their memory (Figure 2b; see Appendix for similar conclusions about the intermediate sample sizes). There is only one large, significant difference between both groups; it occurs when the probability of winning equals .9 ( $t(78) = 2.16, p = .034$ ). Interestingly, however, this difference is such that the records group has an even stronger preference than the experience group for the risky option, thus making the gap with the description group even larger. **These results suggest that memory limits are not a major source of the description–experience gap.**

### *Choice inertia*

**One possible reason why the description–experience gap does not disappear with larger samples is simple choice inertia (Biele, Erev, & Ert, 2009). That is, people might just stick with the option they initially chose.** To test this hypothesis, we determined the average overlap between subsequent choices (i.e., first and second decision, second and third, and so on). The mean overlap was 84% in both the experience and records groups. Although this may seem like an instance of inertia, one has to keep in mind that choices are bound to overlap simply because earlier, smaller samples represent subsets of later, larger samples. To quantify the influence of this dependency, we simulated the choices of respondents who face the same samples that our participants encountered, and employed the natural-mean heuristic (Hau et al., 2008; Hertwig & Pleskac, 2008): They choose the risky option if their sample has a better average outcome than the certain option. The average overlap among these simulated choices was 79%, and the observed overlap is only slightly larger than that predicted by the natural-mean heuristic. **Consequently, choice inertia is unlikely to explain the substantial description–experience gap even for large samples (Figure 2b).**

**To conclude, we found a large description–experience gap, and thereby observed choices that are consistent with the thesis that rare events have more impact in description-based than in experience-based choices, as suggested by Barron and Erev (2003), Hertwig et al. (2004), and Weber et al. (2004). In addition, the gap cannot be explained in terms of limited memory, choice inertia, or payoff variability—the first two factors belonging to the cognitive pillar of the cognitive-ecological framework and payoff variability belonging to the ecological pillar. Finally, sampling error matters but the gap did not decrease in equal measure as the decreasing error: Specifically, we found that 10 times as much experience (50 vs. 5 draws) slightly reduced but by no means eliminated the gap. Sampling error is not the only game in town.**

## STUDY 2: DOES REPRESENTATION FORMAT MATTER?

Statistical probabilities and a priori probabilities differ insofar as the latter can be deduced from logical or other principles, and the former must be inferred from experience. Moreover, both kinds of probability also differ with regard to their representation format. The large majority of studies investigating risky choice represent the chance of described outcomes in terms of *single-event probabilities* (or *percentages*). In contrast, the reliance on experience in generating statistical probabilities leaves it to decision makers to

mentally represent the statistical probabilities in any format they wish, ranging from raw frequencies to single-event probabilities. This begs the question: Does the format of statistical probabilities influence the effect of rare events on decisions from experience?

Indeed, Rakow et al. (2008) suggested that the effect of rare events—via sampling error—is so pervasive that the format of the information does not matter and that the description–experience gap is tantamount to a population–sample gap. In terms of a cognitive–ecological framework, they thus argued that the description–experience gap is entirely explained by the transition from the latent population representation to the sample representation. A corollary of their thesis is that a person’s cognitive representation of the properties of the gambles is either identical in decisions from description and decisions from experience, or if different, inconsequential. Some evidence from studies on probabilistic reasoning, however, conflict with this corollary. Specifically, Gigerenzer and colleagues found evidence that many cognitive illusions—biases in probabilistic reasoning—disappear or are reduced when uncertainty is communicated in terms of evolved representations of information such as natural frequencies rather than in terms of single-event probabilities (e.g., Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; Gigerenzer et al., 1991; Hertwig & Gigerenzer, 1999; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000). These authors’ findings and conclusions have produced a heated and ongoing debate (e.g., Barbey & Sloman, 2007; Kahneman & Tversky, 1996; Sloman, Over, Slovak, & Stibel, 2003).

In the context of risky choice, Gottlieb, Weiss, and Chapman (2007) indeed found that single-event probabilities are processed differently than other representation formats. Our Study 2 builds on this work and asks: To what extent does risky choice depend on how information about gambles is represented, and can the description–experience gap in part be attributed to the fact that experience-based and description-based choices invoke different representations? To this end, we adopt the same yoking procedure that Rakow et al. (2008) used, and from which they concluded that the description–experience gap is a population–sample gap. Specifically, we take the same statistical probabilities as experienced by our respondents in Study 1, and represent them in terms of percentages. Thereby, respondents in both studies receive mathematically equivalent information, merely represented in different formats. One group encountered the information in terms of sequentially experienced outcomes (experience group in Study 1), the other in terms of described probabilities (yoked description group in Study 2). If, notwithstanding mathematical equivalence, different choices emerge, then the description–experience gap is not solely a function of the transition from a latent population to a sample—the population–sample gap—but also a result of the transition from a sample to its cognitive representation.

## Method

### *Participants*

Forty students of the University of Basel took part in this study, recruited from the same pool as used in Study 1.

### *Procedure*

Each participant was matched up with a “twin” in the experience group of Study 1, and was exposed to the information experienced by that twin. Consequently, each participant had to make five repeated choices for each of the twelve decision problems. For example, a participant in Study 1 may have seen the sequence 32, 0, 0, 32, 32 prior to choosing between the risky option and the certain option of 29 in Problem 12. The

corresponding choice described to the participant's twin in Study 2 was a "60% chance to win 32" or a "100% chance to win 29."

To avoid confusion, participants were informed that the risky option could change between each of five subsequent choices<sup>5</sup>; that they might have to make the same choice more than once; and that some choices might be trivial (e.g., when, due to sampling error, the risky option dominated the certain option). The payoff scheme was identical to that in Study 1, with one randomly selected decision played out for money.

## Results and discussion

Does representation format matter? Figure 3 reports participants' choice proportions for the risky option in Study 2 along with those of their twins in Study 1. Let us first turn to choices based on *five* draws (Figure 3a). The twins prove to be twins. The patterns of choices in both the experience group and Study 2's yoked description group track each other. Both groups prefer the certain option at low probabilities and the risky option with high probabilities of winning. To provide a simplified analysis of this trend, we compared the average choices for the risky option for low (.1 and .2) and high probabilities (.8 and .9). The difference is significant both for the experience group ( $t(39) = 4.37, p < .001$ ) and for their twins in the yoked description group ( $t(39) = 5.02, p < .001$ ). In addition, as Figure 3a shows, the trend in the yoked description group runs opposite to that of the description group in Study 1 (negative trend,  $t(39) = -3.07, p = .004$ ). At this point, one cannot help but conclude that there is no systematic description–experience gap once respondents receive mathematically equivalent information. **That is, representation format in risky choice does not matter and the gap is indeed a population-sample gap.**

This conclusion, however, proves premature once results in the remaining panels of Figure 3b–e enter the picture. Here, choices are based on increasingly larger sample sizes, up to a maximum of 50. Twins ain't twins no more. **With more experience, the yoked description group becomes increasingly dissimilar to the experience group.** At sample size 50, there is a substantial difference between the yoked description (Study 2) and the experience groups (Study 1). This gradual change is confirmed by the results of a  $5$  (sample size)  $\times 2$  (group, experience and yoked description)  $\times 4$  (probability to win) ANOVA. The sample size  $\times$  group  $\times$  probability interaction is significant ( $F(12, 936) = 2.50, p = .003, \eta^2 = .031$ ), confirming that rare events drive apart decisions from experience and yoked decisions from description as sample size increases. Furthermore, at sample size 50 there is no significant difference between the average choices for the low and high probabilities evident in the yoked description group any more ( $t(39) = -0.42, p = .68$ ).

Last but not least, because the experience and records groups in Study 1 received identical information, the twins of Study 2 were actually triplets. The information that the respondents in Study 2 received matched that encountered by respondents in the experience as well as in the records group in Study 1. Therefore, we also conducted all the analyses reported above, comparing the records group (Study 1) and the yoked description group (Study 2). We found nearly identical results.

**These findings imply that mathematical equivalence does not necessarily constitute psychological equivalence. Representation format does matter and the gap cannot be simply reduced to a population-sample gap. Perhaps, the most puzzling result from Study 2 is that decisions from experience and yoked decisions from description coincide for small sample sizes and diverge for large ones. Why is that? There is an interesting and testable answer that we outline in the following discussion.**

<sup>5</sup>In Study 1, respondents were informed that from one trial to the next, they would learn more information about a stable underlying risky option. This instruction would have been incomprehensible in Study 2.



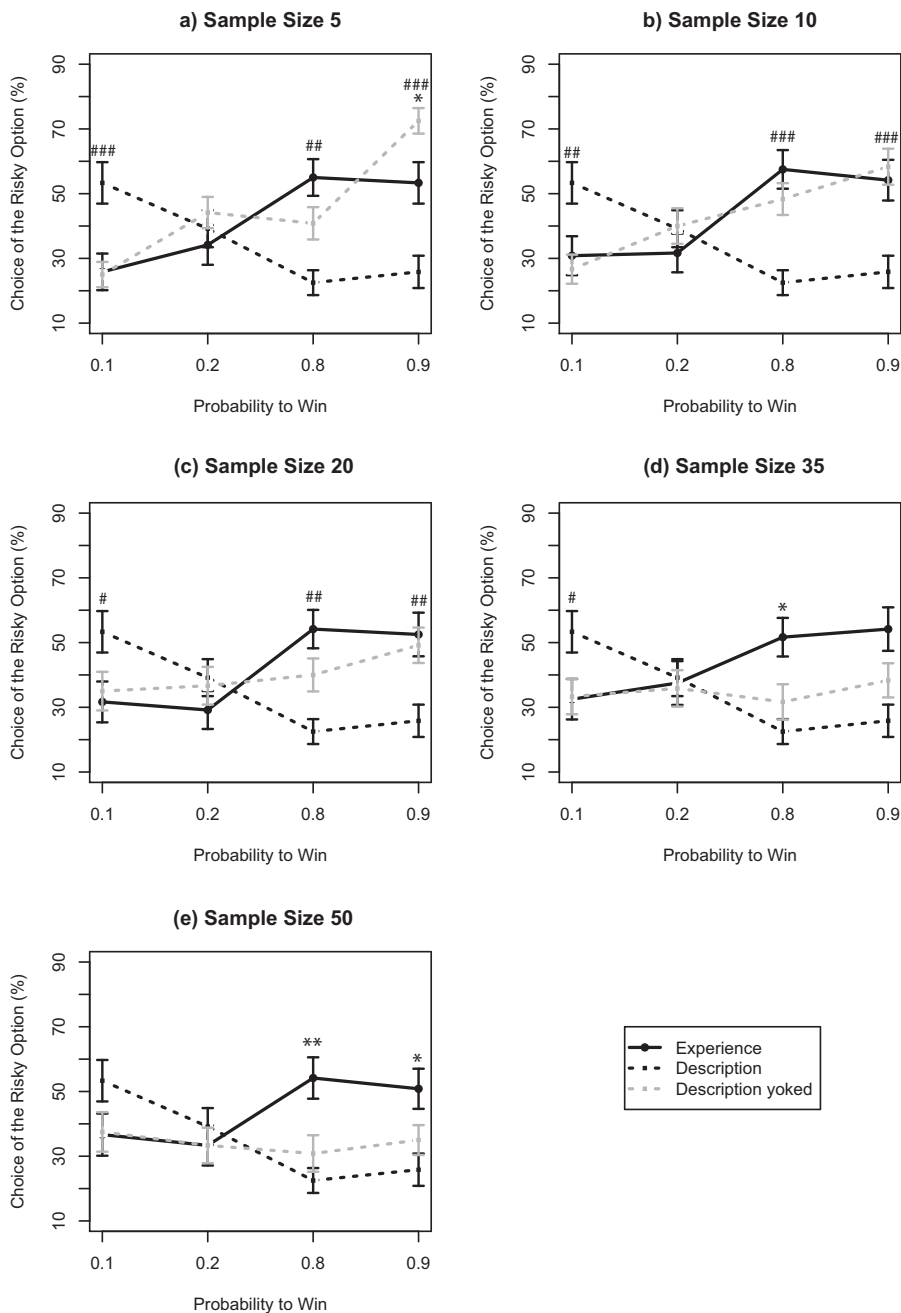


Figure 3. Mean proportion of choices of the risky option in the experience (Study 1) and yoked description (Study 2) groups as a function of probability to win and sample size (5 in Figure 3a to 50 in Figure 3e). For comparison, we also included choices of the description group in Study 1 in all figures. Error bars indicate  $\pm 1$  SE of the mean. For every level of sample size and at every level of probability, significant differences between decisions from experience and yoked decisions from description are marked by asterisks (\* $p < .05$ ; \*\* $p < .01$ ); significant differences between decisions from description with objective probabilities and yoked decisions from description are marked by hash marks (# $p < .05$ ; ## $p < .01$ ; ### $p < .001$ )

## GENERAL DISCUSSION

Dichotomies are seductively simplistic. We began by pointing out that the distinction between situations of risk and uncertainty, typically attributed to Knight (1921), was originally a trichotomy of probability situations: If decision makers are in risky situations they can either have a priori probabilities or statistical probabilities and if they are in uncertain situations they will only have their estimates (or beliefs) to act on. From there, it is easy to see how the distinction between decisions from description and decisions from experience (e.g., Barron & Erev, 2003; Hertwig et al., 2004; Weber et al., 2004) map onto Knight's distinction between a priori and statistical probabilities. But even Knight's trichotomy is simplistic. As Knight himself pointed out, perhaps a better conceptualization is of a continuum of different decision-making situations anchored by three different poles of decisions made from description, decisions made from experience, and decisions made from estimates or beliefs (see p. 199).

We pursued two goals. First, we tested the robustness of the description–experience gap by using choice problems not previously studied. The problems consisted of a safe option and a risky option. Like in Busemeyer (1985), respondents merely sampled the risky option, as the safe option was explicitly described as such. In this paradigm, which combines one “described option” and one “experienced option,” we again found a substantial description–experience gap (Figures 2a and 2b). That is, the gap generalizes to tasks that combined a priori and statistical probabilities. The second goal was to further our understanding of the causes behind the gap, using a cognitive-ecological framework that organizes potential causes. What did we find in Study 1? By pitting the impact of rare events against the possible impact of payoff variability (both ecological factors), we observed that the latter could hardly explain the gap. Specifically, choices in both the description and experience groups were influenced by payoff variability, but its effect failed to explain differences between the groups. Moreover, we also could absolve capacity limits in memory (a cognitive factor) of responsibility for the gap. Decisions from records, in which participants saw the total sequence of draws when making a decision and in which memory was not taxed, produced a description–experience gap that, if anything, was even larger than that for decisions from experience (Figures 2a and 2b). Last but not least, the fact that the gap was not eliminated after substantial experience in both the experience and records group indicates that sampling error—although undoubtedly a contributory factor to the gap (Hertwig et al., 2004; Hau et al., 2008)—is insufficient to account for the whole story.

By combining a priori and statistical probabilities, our experimental task also allowed us to eliminate another possible explanation for the description–experience gap. In decisions from description involving a priori probabilities, the *certainty effect* (Kahneman & Tversky, 1979) refers to the theoretical interpretation that people overweight outcomes that are considered certain, relative to outcomes that are merely probable, and to the empirical observation that certain gains are preferred over their probabilistic counterparts with similar expected values. In decisions from experience involving statistical probabilities, certainty is unattainable. If this lack of certainty had contributed to the description–experience gap in previous studies, our respondents in the experience group would have flocked to the described, certain option. We found no such flight into certainty. In fact, our respondents took considerable risks, predominately in problems where the risky option promised only slightly higher gains than the certain option (i.e., a 90% chance to win 32 Swiss Francs vs. a certain option of 29 Swiss Francs).

In Study 2, we adopted a clever procedure that Rakow et al. (2008) employed. Specifically, we created a twin experience by representing the statistical probabilities encountered in Study 1's experience group to another group; this time, however, in terms of a priori probabilities. Thereby, respondents in Study 2 relived the other group's experience but in a different representation format. The results were striking. With small samples we observed no differential impact of representation format. With large samples, however, we found that mathematically equivalent representation formats produced different patterns of choice. It appears that the frequency–probability divide, so powerful in research on statistical reasoning (e.g., Gigerenzer, 1991), is also of importance in risky choice. In what follows, we address two issues. First, Figure 3 (Study 2) shows the

puzzling effect that with small sample sizes there is no gap, but, paradoxically, with large samples it reappears. Second, why does representation format—a priori probabilities versus statistical probabilities—matter in risky choice?

### The amplification effect: When choices become trivial

Recently, Hertwig and Pleskac (2008) showed that small samples amplify the difference between the options' average rewards. They suggested that this *amplification effect* might ease the difficulty of choosing between options, thus explaining why respondents in decisions from experience studies are content with relatively small samples (Table 1). According to the amplification effect, small samples tend to make the differences between gambles on average larger than the population level difference thereby making choices less difficult. To appreciate the magnitude of the amplification effect as a function of sample size, we calculated, for all participants and problems in Study 1, the absolute difference between the value of the certain option and the sample mean (i.e., the mean of all outcomes in a sample) of the risky option. The average median differences were 1.30 and 0.44 Swiss Francs for sample sizes 5 and 50, respectively. Recall the gambles were constructed so that the objective difference in expected values was 0.20 Swiss Francs. That is, the median difference between the two options based on a sample of 5 was about 300% larger than that based on a sample involving 50 draws (and 650% larger than the objective differences in expected values).

The amplification effect can perhaps help us understand some of the results in Study 2 (Figure 3). Recall the yoked-description group saw described (a priori) probabilities in Study 2 that were completely identical to the experienced (statistical) probabilities in Study 1. Focusing on small samples, it is not surprising that we found no difference between the yoked description and experience groups. For instance, a person who read the description of a “100% chance to win 0” (vs. a certain payoff) is likely to concur with a person who sampled a sequence of 0, 0, 0, 0, 0. The amplification effect also applies to Rakow et al.'s (2008) finding of no description–experience gap when respondents saw a description of the probabilities experienced by another group. In fact, we reanalyzed their data (which they graciously provided), adapting the analysis above. Again, we found an amplification effect: Based on a median sample of 15 outcomes, the median absolute difference between the sample means was 2, a 1500% increase from the median difference between the objective expected values of 0.125 (presented in their decisions from description condition). They and we thus observed the impact of the amplification effect.

Increasing sample size and thus reducing sampling error, however, had a differential impact on decisions from experience and yoked decisions from description. As Figure 3 shows, with larger samples, the yoked description group became more similar to the description group and less similar to the experience group. The choices in the experience group remained largely unchanged as a function of sample size. One might have expected that with increasing sample size, the amplification effect gets smaller and thereby the description–experience gap should get smaller as well. This, however, only holds if the amplification effect and the associated sampling error were the only causes underlying the gap. Results in Hau et al. (2008) and Study 1 show that this is not the case.

To conclude, the amplification effect and sampling error contribute to the description–experience gap, because decisions from description are devoid of sampling error. For small samples, the amplification effect is so large that when we introduced the sampling error of decisions from experience into yoked decisions from description, then choices became trivial in both representation formats and removed any difference in choice. But, when the amplification effect dissipated in larger samples, the gap reappeared. This suggests two things. First, the gap should only be expected when choices are not trivial. Second, in these non-trivial choices, factors other than sampling error—some of which we discussed in this article—contribute to the gap.

### Risky choice: Why does representation format matter?

Figure 3e shows a description–experience gap, notwithstanding the fact that choices are based on mathematically equivalent information. The information only differed in the representation format, statistical

probabilities versus a priori probabilities. How can these different representation formats trigger different choices? We can think of two reasons: one concerns the weighting of monetary outcomes, the other concerns triggering of cognitive heuristics. In Hau et al. (2008; Figure 7, Table 6; see also Jessup et al., 2008), we found that prospect theory's S-shaped decision-weighting function (Tversky & Kahneman, 1992), embodying overweighting of low- and medium probability events and underweighting of common events, is not well suited to model choices based on experienced frequencies. The optimal, fitted parameters proved very close to unity. That is, to the extent that monetary outcomes are explicitly weighted by their likelihood in decisions from experience, the weights equal the experienced frequencies. In decisions from description, in contrast, Edwards (1954) and Tversky and Kahneman (1992; Kahneman & Tversky, 1979) suggested a weighting pattern that systematically deviates from the stated probabilities.

A second explanation, also suggested by Hau et al. (2008), is that decisions from experience and description trigger different decision strategies. Some models that have been proposed to account for experience-based choices are not applicable to decisions from description. For instance, associative learning models as well as simple strategies such as the natural-mean heuristic require no representation of probabilities, nor do they invoke the multiplication of some function of monetary outcomes by some function of their likelihoods as assumed by Neo-Bernoullian models entertained in decisions from description such as prospect theory (Kahneman & Tversky, 1979), cumulative prospect theory (Tversky & Kahneman, 1992), the transfer-of-attention-exchange model (Birnbbaum & Chavez, 1997), or decision affect theory (Mellers, 2000).

## CONCLUSION

In 1983, Lola Lopes, reflecting on the psychological concept of risk, wrote:

After 30 years or more of research on risk, we know a lot about how people make decisions about simple lotteries, but we know remarkably little about decision under uncertainty, possibly because we have not had a good laboratory model of uncertainty. (p. 138)

Not too long ago, one could have rightfully extended this assessment to the “extremely common” (Knight, 1921; p. 215) situations in which we need to base our decisions on statistical, rather than a priori, probabilities. Now, however, we have some simplified laboratory models of these choices involving statistical probabilities (for a review see Erev & Haruvy, in press; Hertwig & Erev, in press; see also Pleskac, 2008). These paradigms allow us to investigate key cognitive processes often absent in the study of simple lotteries such as search for information, termination of search, memory retrieval, and mental tabulation of risks. Moreover, these models enable comparisons between decisions made in response to lotteries involving described probabilities and decisions involving experienced probabilities. These comparisons have revealed a substantial description–experience gap. In the process of understanding the causes and boundaries of this gap, researchers have the chance, so we believe, to develop a richer conception of not only the cognitive, but also the ecological aspects underlying risky decisions in the real world.

## ACKNOWLEDGEMENTS

Ralph Hertwig was supported by Swiss National Science Foundation Grant 100014-118283. We are grateful to Laura Wiles for editing the manuscript and to Eui-Jee Hah for his help and commitment in conducting the studies.

## APPENDIX

All choice proportions of Studies 1 and 2

Problem <sup>1</sup>				Choice of the risky option (%)															
				Description	Experience (Study 1) Sample size					Records (Study 1) sample size					Yoked description (Study 2) sample size				
					5	10	20	35	50	5	10	20	35	50	5	10	20	35	50
No.	$p_{\text{win}}$	$X_{\text{win}}$	$\sigma$		5	10	20	35	50	5	10	20	35	50	5	10	20	35	50
1	10	5.30	1.60	67.5	37.5	40.0	40.0	37.5	42.5	47.5	45.0	40.0	45.0	47.5	40.0	47.5	50.0	42.5	47.5
2	20	4	1.60	42.5	35.0	40.0	30.0	37.5	32.5	60.0	45.0	40.0	32.5	42.5	57.5	55.0	50.0	42.5	45.0
3	80	4	1.60	32.5	57.5	57.5	50.0	55.0	55.0	60.0	70.0	57.5	65.0	57.5	47.5	45.0	25.0	35.0	27.5
4	90	5.30	1.60	27.5	60.0	55.0	52.5	52.5	47.5	67.5	75.0	77.5	67.5	77.5	85.0	67.5	62.5	45.0	40.0
5	10	15	4.50	52.5	20.0	35.0	30.0	27.5	30.0	17.5	22.5	37.5	37.5	40.0	17.5	25.0	37.5	37.5	37.5
6	20	11.30	4.50	42.5	32.5	32.5	25.0	37.5	32.5	30.0	25.0	35.0	27.5	35.0	42.5	40.0	37.5	35.0	30.0
7	80	11.30	4.50	17.5	60.0	62.5	57.5	52.5	57.5	50.0	60.0	62.5	67.5	67.5	40.0	50.0	45.0	32.5	27.5
8	90	15	4.50	27.5	52.5	55.0	60.0	62.5	60.0	60.0	62.5	72.5	72.5	70.0	67.5	60.0	47.5	37.5	35.0
9	10	32	9.60	40.0	20.0	17.5	25.0	32.5	37.5	25.0	30.0	32.5	30.0	42.5	17.5	7.5	17.5	20.0	27.5
10	20	24	9.60	32.5	35.0	22.5	32.5	37.5	35.0	25.0	22.5	30.0	25.0	25.0	32.5	25.0	22.5	30.0	25.0
11	80	24	9.60	17.5	47.5	52.5	55.0	47.5	50.0	57.5	65.0	62.5	47.5	57.5	35.0	50.0	50.0	27.5	37.5
12	90	32	9.60	22.5	47.5	52.5	45.0	47.5	45.0	62.5	52.5	55.0	57.5	60.0	65.0	47.5	37.5	32.5	30.0

<sup>1</sup>Each problem consists of a risky option—represented by a deck of cards—paying  $X_{\text{win}}$  Swiss Francs with probability  $p_{\text{win}}$ , 0 otherwise, coupled with a certain option. Payoff variability is indicated by the standard deviation  $\sigma$  of the risky option. For half of the participants, the certain option was the expected value of the risky option augmented by 0.20 Swiss Francs, for the other half, it was reduced by the same amount; 1 Swiss Franc = US \$ 0.95

## REFERENCES

- Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, 21, 503–546.
- Barbe, K. (2001). The dilemma with dichotomies. *Language and Communication*, 21, 89–103.
- Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, 30, 241–254.
- 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.
- Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology*, 53, 155–167.
- Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: Violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes*, 71, 161–194.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review*, 113, 409–432.
- Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky choice with heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, and Willemsen (2008), and Rieger and Wang (2008). *Psychological Review*, 115, 281–290.
- Bussemeyer, 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, 538–564.
- Bussemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, 100, 432–459.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–114.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions of the literature on judgment under uncertainty. *Cognition*, 58, 1–73.
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106, 180–209.



- Edwards, W. (1954). The theory of decision making. *Psychological Bulletin*, 51, 380–417.
- Epstein, L. G., & Wang, T. (1994). Intertemporal asset pricing under Knightian uncertainty. *Econometrica*, 62, 283–322.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112, 912–931.
- Erev, I., & Haruvy, E. (in press). Learning and the economics of small decisions. In J. H. Kagel, & A. E. Roth (Eds.), *The handbook of experimental economics* (2nd ed.). Princeton: Princeton University Press.
- Erev, I., & Hertwig, R. (in press). The description–experience gap in risky choice. *Trends in Cognitive Sciences*.
- Fiedler, K., & Juslin, P. (2006). Taking the interface between mind and environment seriously. In K. Fiedler, & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 3–29). New York: Cambridge University Press.
- 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.
- Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science*, 44, 879–895.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases.” In W. Stroebe, & M. Hewstone (Eds.), *European Review of Social Psychology* (Vol. 2, pp 83–115). Chichester, England: Wiley.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684–704.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Gottlieb, D. A., Weiss, T., & Chapman, G. B. (2007). The format in which uncertainty information is presented affects decision biases. *Psychological Science*, 18, 240–246.
- Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, 21, 493–518.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, 15, 534–539.
- Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). The role of information sampling in risky choice. In K. Fiedler, & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 72–91). New York: Cambridge University Press.
- Hertwig, R., & Gigerenzer, G. (1999). The ‘conjunction fallacy’ revisited: how intelligent inferences look like reasoning errors. *Journal of Behavioral Decision Making*, 12, 275–305.
- 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.
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290, 2261–2262.
- Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback produces divergence from prospect theory in descriptive choice. *Psychological Science*, 19, 1015–1022.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582–591.
- Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review*, 107, 397–402.
- Kareev, Y., Lieberman, I., & Lev, M. (1997). Through a narrow window: Sample size and the perception of correlation. *Journal of Experimental Psychology: General*, 126, 278–287.
- Katz, L. (1962). Monetary incentive and range of payoffs as determiners of risk taking. *Journal of Experimental Psychology: General*, 64, 541–544.
- Knight, F. H. (1921). *Risk, uncertainty, and profit*. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Company.
- Lopes, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 137–144.
- Luce, R. D. (2000). *Utility of gains and losses: measurement-theoretical and experimental approaches*. Mahwah: Lawrence Erlbaum Associates.
- Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: Dover Publications, Inc.
- Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification of investments*. New York: Wiley.
- Mellers, B. A. (2000). Choice and the relative pleasure of consequences. *Psychological Bulletin*, 126, 910–924.
- Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- Myers, J. L., & Katz, L. (1962). Range of payoffs and feedback in risk taking. *Psychological Reports*, 10, 483–486.
- Myers, J. L., & Sadler, E. (1960). Effects of range of payoffs as a variable in risk taking. *Journal of Experimental Psychology: General*, 60, 306–309.



- Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). New York: Academic Press.
- Pleskac, T. J. (2007). A signal detection analysis of the recognition heuristic. *Psychonomic Bulletin & Review*, 14, 379–391.
- Pleskac, T. J. (2008). Decision making and learning while taking sequential risks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 167–185.
- 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.
- Runde, J. (1998). Clarifying Frank Knight's discussion of the meaning of risk and uncertainty. *Cambridge Journal of Economics*, 22, 539–546.
- Schooler, L. J., & Hertwig, R. (2005). How forgetting aids heuristic inference. *Psychological Review*, 112, 610–628.
- Sedlmeier, P. (2002). Associative learning and frequency judgements: The PASS model. In P. Sedlmeier, & T. Betsch (Eds.), *Etc.: Frequency processing and cognition* (pp. 137–152). Oxford: Oxford University Press.
- Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decisions Processes*, 91, 296–309.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, 6, 363–375.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31–48.
- Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, 102, 269–283.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.
- Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychological Science*, 20, 473–479.
- 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.

#### *Authors' biographies:*

**Robin Hau** is a Post-Doctoral Researcher in the Department of Psychology at the University of Basel, Switzerland. His research focuses on experience-based decisions and cognitive modeling.

**Timothy J. Pleskac** is an Assistant Professor of Applied Cognitive Science in the Department of Psychology at Michigan State University. His research focuses on developing formal cognitive models in the area of cognitive decision theory.

**Ralph Hertwig** is a Professor of Cognitive and Decision Sciences in the Department of Psychology at the University of Basel, Switzerland. His research focuses on models of bounded rationality, social intelligence, and methodology of the social sciences.

#### *Authors' addresses:*

**Robin Hau** and **Ralph Hertwig**, Department of Psychology, University of Basel, Missionsstrasse 64a, 4055 Basel, Switzerland.

**Tim Pleskac**, Department of Psychology, Michigan State University, East Lansing, MI, 48824, USA.