**Article:**

Weiss-Cohen, L, Konstantinidis, E orcid.org/0000-0002-4782-0749, Speekenbrink, M et al. (1 more author) (2016) Incorporating conflicting descriptions into decisions from experience. Organizational Behavior and Human Decision Processes, 135. pp. 55-69. ISSN 0749-5978

https://doi.org/10.1016/j.obhdp.2016.05.005

# Incorporating conflicting descriptions into decisions from experience

Leonardo Weiss-Cohen, Emmanouil Konstantinidis,
Maarten Speekenbrink and Nigel Harvey
University College London

Abstract

Decisions in everyday life are commonly made using a combination of descriptive and experiential information, and these two sources of information frequently contradict each other. However, decision-making research has mostly focused on description-only or experience-only tasks. Three experiments show that individuals exposed to description and experience simultaneously are influenced by both, particularly in situations in which descriptions are in conflict with experience. We examined cognitive models of how people integrate their experience with descriptions of choice outcomes, with different weights given to each source of information. Experience was the dominant source of information, but descriptions were taken into consideration, albeit at a discounted level, even after many trials. Models that included the descriptive information fitted the human data more accurately than models that did not. Wider implications for understanding how these two commonly available sources of information are combined for daily decision-making are discussed.

*Keywords:* decision-making, decisions from experience, reinforcement learning, repeated decisions, warning labels

The vast majority of human decision-making research to date has been based around descriptive paradigms (Camilleri & Newell, 2009; Fantino & Navarro, 2012; Rakow, Demes, & Newell, 2008). When participants make decisions based on descriptions, they gather information about the potential outcomes of their choices and associated probabilities by

Leonardo Weiss-Cohen, Emmanouil Konstantinidis, Maarten Speekenbrink and Nigel Harvey, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, UK.
Emmanouil Konstantinidis is now at the School of Psychology, University of New South Wales, Sydney, Australia.
This research has been supported by a UCL Impact Award.
Correspondence regarding this manuscript should be addressed to Leonardo Weiss-Cohen, Department of Experimental Psychology, University College London, Gower Street, WC1E 6BT, London, UK. E-mail: l.cohen.12@ucl.ac.uk

reading complete abstract descriptions of available options (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). However in everyday life individuals are rarely presented with such detailed unambiguous descriptions and instead make decisions based on their own direct experiences in noisy environments. When making decisions based on experience, individuals learn about the potential outcomes of their choices by observing samples of outcomes over time (e.g., Bechara, Damasio, Damasio, & Anderson, 1994; Knox, Otto, Stone, & Love, 2012; Yechiam & Rakow, 2012).

Much of the research on decisions from description has dealt with factors that influence people's risky decision making. Thus, typically, experimental participants have been asked to choose between a risky option and a sure option or between a high-risk option and a low-risk option. For example, Kahneman and Tversky (1979) report studies of this type when providing their rationale for the development of Prospect Theory. Concurrently, decisions from experience have been mostly used in research on learning, exploration, exploitation, and cognitive modeling. For example, all of these processes have been extensively studied using the Iowa Gambling Task (Bechara et al., 1994), which is a frequently employed risky-choice task based on decisions from experience. Before Barron and Erev (2003), these two approaches had mostly been studied separately, with little overlap. However, from the start of the research confronting these two types of experimental paradigms, those interested in how decisions from experience differ from decisions from description have focused on the same issues. For example, Barron and Erev used experiential paradigms to explore many of the same risky choices initially presented by Kahneman and Tversky using descriptions. Thus, risky decision making provides a common thread that links classic literature on descriptive decision making with more recent research on decisions from experience.

If the core relevant information about outcomes is the same, in particular the quantitative information such as probabilities and values of outcomes, then there should be no differences in behavior according to how the information is presented. However this does not always appear to be the case. Recent research confronting description and experience has found significant differences between choices made from experience and those made from description when the two sources carry the same information about the outcomes (e.g., Barron & Erev, 2003; Hau, Pleskac, & Hertwig, 2010; Hertwig, Barron, Weber, & Erev, 2004). This phenomenon has been named the "description-experience gap" by Hertwig and Erev (2009), and there has been growing interest in this field recently (for reviews, see Camilleri & Newell, 2013b; Rakow & Newell, 2010). Despite strong support, some studies still failed to find any behavioral differences between decisions from description and decisions from experience (e.g., Camilleri & Newell, 2011, 2013a; Fox & Hadar, 2006; Rakow et al., 2008), raising new issues to be explored regarding the mechanisms that contribute towards the appearance of gaps. For example, one such issue is how the two different sources of information are processed and integrated together when they are both available simultaneously. However, research on the description-experience gap also kept the two paradigms separate by presenting different participants with either description-only or experience-only tasks in isolation (for a review, see Fantino & Navarro, 2012).

Decisions in everyday life are commonly made using a combination of descriptive and experiential information. For example, doctors frequently rely on readings of published literature and research, which can be considered a form of description, and combine it with their own clinical experience, when prescribing drugs or assessing the risk of a medical procedure (Dawes, Faust, & Meehl, 1989). Consumers may base their buying decisions on a combination of descriptive reviews and experiences of similar items bought in the past. Warning labels can be considered as descriptive information that is added to an individual's own experience. Limited published research so far has looked at the influence of descriptions when participants have access to both description and experience at the same time, with contradictory results. According to a study by Lejarraga and Gonzalez (2011), descriptive information is neglected when experience is also available. In contrast, Barron, Leider, and Stack (2008) showed that providing participants with descriptive information influenced behavior.

The extant "description-and-experience" research used paradigms in which the description matched the experience, with both based on the same underlying distribution of outcomes and providing participants with the same basic information. That is, the description was a verbal representation of the distribution of payoffs actually experienced by the participants. The researchers therefore had to rely on observing differences in behavior based on the existence of a robust description-experience gap and its theoretical predictions to test whether description or experience was influencing participants: Behavior consistent with underweighting of rare events would be expected from participants following experience, while overweighting would be associated with descriptive information being used. For example, Lejarraga and Gonzalez (2011) mention that, when providing participants with both description and experience simultaneously, they observed behavior consistent with underweighting of the rare event. According to the authors, this is evidence that experience was taken into account, but description was neglected: Previous research has associated the underweighting of rare events with decisions from experience (Hertwig & Erev, 2009), while the overweighting of rare events observed has subsequently been associated with decisions from description (Kahneman & Tversky, 1979).

However, recent research in this area has explored the idea that the gap is likely a product of differences between the experimental paradigms. Descriptive tasks typically rely on single-shot paradigms without feedback, while experiential tasks tend to use repeated-choice paradigms with feedback (Camilleri & Newell, 2013a; Jessup, Bishara, & Busemeyer, 2008). Therefore, the overweighting and underweighting of rare events may not be necessarily driven by the mere presence or absence of descriptions, respectively, but instead by the different nature of the paradigms used in each line of research. In both studies mentioned in the previous paragraph, the paradigms used were repeated-choice experiential tasks with feedback. We suggest that the reason why Lejarraga and Gonzalez (2011) observed behavior explicable by underweighting of rare events is not because participants neglected the descriptions, but because the paradigm used was typical of experiential research. Furthermore, the reason why the authors did not observe any differences in behavior in their experiment is not because description was neglected, but because it conveyed the same information as experience and therefore its influence on behavior was not observable. Barron et al. (2008) observed a difference in behavior in their experiment because the rare event in their description, with a chance of 1 in 1000, rarely or never occurred.

We present an experimental paradigm in which description conflicts with experience, which will allow us to verify how different sources of information influence behavior. If each source provides different information to participants then, by analyzing the choice patterns, we can determine which one has been used in the decision process. These situations of conflicting information are likely to be representative of typical day-to-day decision making in a dynamic world. In such ever changing environments, the more adaptive short-term nature of experience, which tends to rely on small samples (Hertwig & Pleskac, 2010), compared to the relatively more static long-term nature of description, which tend to rely on large samples (e.g., published results from randomized control trials), would naturally lead to the two diverging over time [1]. Experience allows for continuous learning of the environment; this is not the case with descriptions, which typically take longer to be updated and can quickly become out of date, leading to negative impacts on choices made in changing environments (Rakow & Miler, 2009). Sampling biases can also create mismatches between description and experience, in particular when rare events are involved (Fox & Hadar, 2006; Hertwig et al., 2004).

Even with large samples, the representative set behind a description can differ from an individual's particular experience, depending on the source of the description. Glasgow et al. (2006) and Kamal and Peppercorn (2013) discuss the external validity of medical research findings, which are typically used as reference points for decision-making, but are not always applicable to a doctor's more localized clinical experience. This is especially true for doctors who have to deal with patient populations that are not representative of the reference population in the standard description. Rakow, Vincent, Bull, and Harvey (2005) showed how mortality risk assessments based on reference research conducted in the Unites States differed from personal experience of doctors at a selected hospital in the United Kingdom. Other examples can result from the overzealous usage of warning signs which misrepresent risks, for example by describing a risk as likely when in reality it is rarely experienced. Carson and Mannering (2001) mention the overuse of road traffic ice warning signs in locations where ice is rarely observed.

If such mismatches between description and experience are encountered frequently, understanding how individuals deal with these situations is crucial for ecologically valid research with real life practical implications. For example, warning labels can be considered descriptive information that conflicts with experience, since they typically present rare events that are not observed directly by the majority of individuals. Research "suggests that the warning labels' impact on behavioral compliance is not as clear as expected" (Argo & Main, 2004, p. 193), with a potential explanation for warning label ineffectiveness being that "[i]f the information in a warning contradicts one's existing beliefs, the warning information might be discounted" (Rogers, Lamson, & Rousseau, 2000, p.130).

The experiments and cognitive models in this study were designed to investigate the hypothesis that descriptions are not neglected by individuals when both description and experience are available, and that descriptions are partially discounted. The allocation of different weights to description and experience has been suggested before (Barron et al., 2008; Newell & Rakow, 2007; Shlomi, 2014). In support of these suggestions, research has shown that descriptions can be overwhelmed by experience (Jessup et al., 2008), and decision makers seem to prefer experiential information (Lejarraga, 2010). Concordantly,

[1] We thank an anonymous reviewer for highlighting that in a new world of more dynamic on-line information sharing such divergences can also occur in the opposite direction. For instance, when considering customer reviews on web pages, reviews constantly accumulate, affecting the overall mean rating of a product, leading to more dynamic descriptive information. Conversely, experiences might remain static if a person is simply no longer exposed to similar situations in the future.

experience is easier to process cognitively (Glöckner, Fiedler, Hochman, Ayal, & Hilbig, 2012), with personal experiences evoking strong emotional and visceral reactions, vis-à-vis statistical descriptions, which lead to more detached analytic considerations (Weber, 2006). We also introduce a Bayesian-updating cognitive model which combines description and experience with different weights, allowing the weighting to change over time according to the plausibility of the descriptive information. The combination of behavioral observations and cognitive modeling of the results from our experiments helps to shed additional light on how description and experience are integrated during the decision-making process.

## Experiment 1

### Method

**Design.** The first experiment had a 4 × 2 between-subjects design: four types of information presentation and two levels of probabilities for the risky option. Information was presented in one of the following: the description-only (D) condition; the experience-only (E) condition; the description-experience-same (DES) condition; and the description-experience-conflict (DEC) condition. The two levels of probabilities referred to the risky option: the 80% probability condition provided participants with a reward 80% of the time (and no reward otherwise), and the 20% probability condition had a 20% chance of providing the reward. Each participant was presented with only one type of information presentation and only one level of probability.

**Participants.** 172 participants (67 females; age: $M = 32$ years, $SD = 10$ years) were recruited on-line using Amazon's Mechanical Turk service. Participation was restricted to individuals whose location was defined as in the United States. There were 28 participants in the $D_{20\%}E_{80\%}C$ condition, 24 in the $D_{80\%}$ condition, and 20 each in the remaining six conditions ($D_{80\%}E_{20\%}C$, $D_{80\%}E_{80\%}S$, $D_{20\%}E_{20\%}S$, $E_{80\%}$, $E_{20\%}$, and $D_{20\%}$; subscripts indicate the probability levels used in the description *D* and experience *E*, whilst *C* stands for *conflicting* and *S* stands for *same*). No participants were excluded from the analysis. Participants were paid a fixed amount of US$ 0.25 for participating and an additional bonus according to the outcomes of the choices they made during the experiment (Bonus: $M =$ US$ 0.71, $SD =$ US$ 0.12)[2].

**Task.** Participants were initially presented with an instructions screen with information about the task. They were told that the task involved choosing between two on-screen buttons, with each button associated with a gamble paying rewards with a certain chance. The idea of conflicting descriptions was introduced as follows:"Because of the way that computers generate random numbers, sometimes the actual frequency that you will experience of winning rewards might not be the same as the one indicated. Ideally, it should be the same, but sometimes it can fluctuate both up and down. It is up to you to assess how attractive each button is based on the actual rewards you get from clicking it. Choosing wisely between the two gambles, in order to maximize your points, will help you increase your bonus."

[2]Bonus amount in Experiment 1 was not influenced by the probability level condition (80%: 0.71; 20%: 0.70; $F(1, 164) = 0.587, p = .44$), but it was influenced by the information presentation condition, with a significantly lower overall bonus in the DEC condition (D: 0.69; E: 0.73; DES: 0.74; DEC: 0.67; $F(3, 164) = 4.111, p = .008$).

After reading the instructions, participants were then presented with two buttons side by side on screen: one button provided the participant with the sure outcome of two points 100% of the time, and the other button was a risky gamble which gave participants five points either 20% or 80% of the time, depending on the experimental condition, and zero points otherwise. Safe and risky button locations were counterbalanced between participants. Choices were made using the mouse. All of the participants' choices between the two options were financially consequent and accumulated towards their final pay. Points were converted to money at a rate of US$ 0.20/100 points in the 80% condition and US$ 0.40/100 points in the 20% condition[3]. Accumulated amounts in points and US dollars were shown on-screen and updated after each choice was made. Participants completed the task in an average of 7.0 minutes (SD = 3.5 minutes).

Table 1
*Button labels according to condition in Experiment 1*

| Experimental conditions | Safe choice button label | Risky choice button label |
|---|---|---|
| $E_{80\%}$, $E_{20\%}$ | (blank) | (blank) |
| $D_{80\%}$, $D_{80\%}E_{80\%}S$, $D_{80\%}E_{20\%}C$ | 2 points with 100% probability. | 5 points with 80% probability; Zero otherwise. |
| $D_{20\%}$, $D_{20\%}E_{20\%}S$, $D_{20\%}E_{80\%}C$ | 2 points with 100% probability. | 5 points with 20% probability; Zero otherwise. |

In the description-only (D) conditions, each button had a label that provided participants with a description of the underlying distribution of outcomes, as detailed in Table 1. Participants were told to choose one button once, and that their selection would be repeated by the computer 100 times, each time drawing from the underlying distribution of the option chosen, to calculate their total bonus (Camilleri & Newell, 2013a).

The experiential conditions (E, DES and DEC) involved 100 repeated individual choices. After each trial, participants were given full feedback, with both the earned and foregone outcomes displayed in the relevant buttons, and asked to choose again. In the E conditions, the two buttons were blank. In the DEC and DES conditions, the buttons contained descriptive labels, as detailed in Table 1. In the DES conditions, the description matched the experience: the outcomes of each choice were drawn from the same distribution as that described in the button. In the DEC conditions, the description for the risky choice showed a probability level opposite to the one used to draw the experiential outcomes after each choice. For example, participants in the $D_{80\%}E_{20\%}C$ condition were shown a risky button with a description (D) that indicated an 80% probability of winning five points. However the actual gains experienced (E) by participants for the risky choice were drawn with a 20% probability distribution. In this condition, the conflicting description made the choice more attractive than it was in reality. The situation was reversed in the $D_{20\%}E_{80\%}C$ condition, with the conflicting description making the risky choice appear less attractive (D=20%, E=80%).

[3]We used different exchange rates according to condition to ensure that all subjects could earn a similar amount of money for their participation in our task, by keeping the financial expected value of the highest earning gamble equal to US$0.80 regardless of condition, across all experiments in this study.

In order to avoid sampling biases, samples were pseudo-randomized in groups of 10 outcomes each, for each participant. Within each 10 outcomes, the samples were yoked to perfectly represent the exact appropriate level of reward events expected in the underlying distribution, either eight or two observations (80% and 20% conditions respectively), in a randomized order (Camilleri & Newell, 2011).

## Results and Discussion

The main dependent variable was the proportion of risky choices (R-rate). The R-rate was calculated as the average proportion of times that participants selected the risky option for each block of 20 trials.

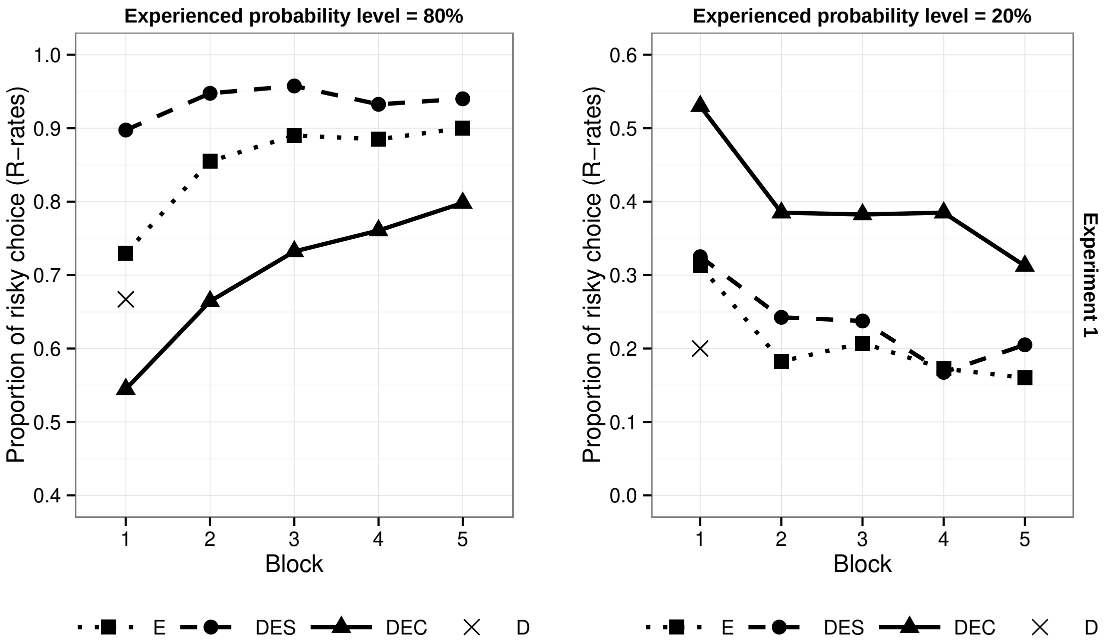


*Figure 1*. Evolution of the risky choice rate (R-rate) for each block of 20 trials for Experiment 1. The left panel shows the results when the risky choice paid a reward of five points 80% of the time, while in the right panel the same reward was paid 20% of the time. The lines refer to the different descriptions presented to participants (E: experience only, no description; DES: description and experience same; DEC: description and experience conflicting). The X refers to the description-only condition, which involved a single choice on the first trial.

The main analysis was conducted separately for each experienced probability condition (80% and 20%). The R-rates in each block were analyzed with a generalized linear mixed-effects model (assuming a binomial distribution with a logit link function) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) in R (R Core Team, 2014). The between-subjects conditions were the types of information presentation (E, DES or DEC). The D condition was excluded from the quantitative analysis because of the different nature of the paradigm, since it involved only one decision without feedback, which is not directly comparable to the repeated decisions of the other conditions (Mean R-rates: $D_{80\%}$: 67%; $D_{20\%}$: 20%). The within-subjects conditions were the blocks of 20 choices each. For each participant, the model contained a random intercept and a random slope for the effect of each block. This approach was used to capture the nested structure of the data. The random intercept accounts for differences in individual overall levels of risky choices and the random slope for differences in changes from block to block for each participant.

The main effect of information presentation condition on R-rates was significant for both probability levels (Figure 1). In the 80% probability condition, R-rates were lowest in the DEC condition (DEC: 70%; DES: 94%; E: 85%;$\chi^2(2) = 31.33$, $p < .001$). In the 20% condition, R-rates were highest in the DEC condition (DEC: 40%; DES: 24%; E: 21%; $\chi^2(2) = 17.50$, $p < .001$). The main effect of block was also significant for both probability levels. In the 80% probability condition, R-rates increased over time ($\chi^2(4) = 37.73$, $p < .001$). In the 20% condition, R-rates decreased over time ($\chi^2(4) = 37.94$, $p < .001$). The interactive effect of block and information condition was not significant in either probability condition ($\chi^2(8)$s$< 7.73$, $p$s$> .46$).

A post-hoc analysis for the last block of 20 trials was also conducted, for each of the two levels of probability separately, in order to compare the effects of the three different types of information presentation conditions (DEC, DES and E). It is of particular interest to look at the R-rates in the last block since by then participants can be expected to have stabilized in their preferred choice (Bechara et al., 1994; Ert & Erev, 2007). There was a marginally significant difference between the DES and the E conditions in the 80% condition (DES: 94%; E: 90%; $\chi^2(1) = 4.35$, $p = .04$, $\varphi_c = .07$), but no difference in the 20% condition (DES: 21%; E: 16%; $\chi^2(1) = 2.71$, $p = .10$, $\varphi_c = .06$), with small effect sizes in both cases. The presence of descriptive information congruent to experience influenced behavior weakly or not at all, consistent with the results observed by Lejarraga and Gonzalez (2011).

However, in the DEC conditions, which combined conflicting information from description and from experience, participants' behavior was shifted towards the choice predicted from the descriptive button labels, as observed in the relevant E and DES experimental conditions, and away from the choice observed in the other conditions. For example, if the $D_{20\%}$ description in the $D_{20\%}E_{80\%}C$ condition was influencing behavior, we would expect a shift away from the behavior observed in the $D_{80\%}E_{80\%}S$ and the $E_{80\%}$ conditions and towards what was observed in the $D_{20\%}E_{20\%}C$ and $E_{20\%}$ conditions, and vice-versa for the $D_{80\%}E_{20\%}C$ condition. This effect was observed. In the 80% condition, the DEC condition made the risky choice less attractive by describing a lower probability of rewards than experienced, and in the 20% it made it more attractive. Therefore we would expect a decrease in R-rates in the 80% condition and an increase in the 20% condition. In the 80% probability condition, R-rates in the DEC condition were significantly lower than in the other two conditions (DEC: 80%; DES: 94%; E: 90%; $\chi^2(2) = 45.98$, $p < .001$, $\varphi_c = .18$). A similar but mirrored effect was observed in the 20% condition, where R-rates in the DEC condition were significantly higher (DEC: 31%; DES: 21%; E: 16%; $\chi^2(2) = 28.09$, $p < .001$, $\varphi_c = .15$).

The conflicting descriptive information influenced behavior significantly, with strong effect sizes, and in the direction predicted by the misleading information provided. This behavior could be explained by participants taking into account the descriptive information and integrating it with the experiential information into their decision-making process. If participants were disregarding the descriptive information completely, the R-rates should not have differed between the comparable DEC, DES and E conditions for each probability level, since the experienced feedback in these three conditions was the same.

The behavior observed in the DEC conditions was also shifted towards what could be interpreted as a more random pattern, with R-rates closer to 50% than in the respective DES and E conditions. We propose that this shift is towards more predicted behavior as inferred by the conflicting descriptions, with participants being influenced by the content of the information available in the descriptions. However, other reasons might cause a similar shift towards random behavior, as a result of the increase in uncertainty in the DEC conditions. The conflicting information introduced further uncertainty into the task, and uncertainty can make participants believe less in their own experience and also explore more often, which could lead to more random-like behavior (Erev & Barron, 2005; Knox et al., 2012; Speekenbrink & Konstantinidis, 2015). In order to test if participants were being influenced by the content of the conflicting description, or simply behaving more randomly, we devised Experiment 2 in which the conflicting information should influence participants away from randomness.

## Experiment 2

While in Experiment 1 the conflicting information influenced participants towards what could potentially be interpreted as more random behavior, Experiment 2 was designed so that the conflicting information should influence participants behavior away from randomness. For example, in the 20% condition, when the conflicting information led to an increase in the risky choice rates (R-rates), the control R-rates (in the E and DES conditions, without the conflicting information) were initially below the random point of 50%; and above random for the 80% condition, when R-rates decreased. The new design used a paradigm in which an increase in the R-rate would be associated with an initial R-rate higher than 50% (Experiment 2a), and a decrease in R-rate associated with an initial R-rate lower than 50% (Experiment 2b). In this way, conflicting descriptions should move the observed behavior away from randomness.

### Method

**Design.** These experiments used a 3 × 2 between-subjects designs with three types of information presentation and two levels of probabilities in the risky option. Information was presented in one of the following: the experience-only (E) condition; the description-experience-same (DES) condition; or the description-experience-conflict (DEC) condition. The two levels of probabilities of the risky option were changed from Experiment 1: in Experiment 2a the probabilities used were 80% and 40%, and in Experiment 2b they were 40% and 20%. Each participant was presented with only one type of information presentation and only one level of probability.

**Participants.** Participants were recruited on-line using Amazon's Mechanical Turk service. Participation was restricted to individuals whose location was defined as in the United States. No participants were excluded from the analysis. Participants were paid a fixed amount of US$ 0.25 for participating and an additional bonus according to the outcomes of the choices they made during the experiment.

***Experiment 2a.*** 120 individuals participated (63 females; age: $M$ = 34 years, $SD$ = 11 years), 20 in each experimental condition. Average bonus paid was US$ 0.90 ($SD$ = US$ 0.06)[4].

***Experiment 2b.*** 120 individuals participated (43 females; age: $M$ = 31 years, $SD$ = 9 years), 20 in each experimental condition. Average bonus paid was US$ 0.55 ($SD$ = US$ 0.18)[5].

**Task.** The experimental paradigm was similar to that of Experiment 1, the only differences being the new values and probabilities for the risky option. The safe button still paid a sure outcome of two points 100% of the time in both experiments. In Experiment 2a, the risky choice paid rewards of six points either 80% or 40% of the time, according to the probability level condition. In Experiment 2b, the risky choice paid four points either 40% or 20% of the time. The new outcomes and button labels can be seen in Table 2. Points were converted to money at a rate of US$ 0.20/100 points in Experiment 2a and US$0.40/100 points in Experiment 2b. Participants completed the task in an average of 6.1 minutes (SD = 2.6 minutes).

Table 2
*Button labels according to condition in Experiment 2*

| Experimental conditions | Safe choice button label | Risky choice button label |
|---|---|---|
| Experiment 2a | | |
| $E_{80\%}$, $E_{40\%}$ | (blank) | (blank) |
| $D_{80\%}E_{80\%}S$, $D_{80\%}E_{40\%}C$ | 2 points with 100% probability. | 6 points with 80% probability; Zero otherwise. |
| $D_{40\%}E_{40\%}S$, $D_{40\%}E_{80\%}C$ | 2 points with 100% probability. | 6 points with 40% probability; Zero otherwise. |
| Experiment 2b | | |
| $E_{40\%}$, $E_{20\%}$ | (blank) | (blank) |
| $D_{40\%}E_{40\%}S$, $D_{40\%}E_{20\%}C$ | 2 points with 100% probability. | 4 points with 40% probability; Zero otherwise. |
| $D_{20\%}E_{20\%}S$, $D_{20\%}E_{40\%}C$ | 2 points with 100% probability. | 4 points with 20% probability; Zero otherwise. |

## Results and Discussion

As in Experiment 1, the main dependent variable was the average proportion of times individuals selected the risky choice (R-rate), in blocks of 20. The same analysis that was

[4] Bonus amount in Experiment 2a was not influenced by the information presentation condition (E: 0.90; DEC: 0.90; DES: 0.90;$F(2,114) = 0.079, p = .92$), but it was influenced by the probability level condition, with a significantly lower overall bonus in the 40% condition (80%: 0.92; 40%: 0.89; $F(1,114) = 7.538, p = .007$).

[5] Bonus amount in Experiment 2b was not influenced by the information presentation condition (E: 0.54; DEC: 0.55; DES: 0.54;$F(2,114) = 0.732, p = .48$), but it was influenced by the probability level condition, with a significantly lower overall bonus in the 40% condition (40%: 0.37; 20%: 0.72; $F(1,114) = 1164, p < .001$).

used in Experiment 1 was conducted separately for Experiments 2a and 2b.

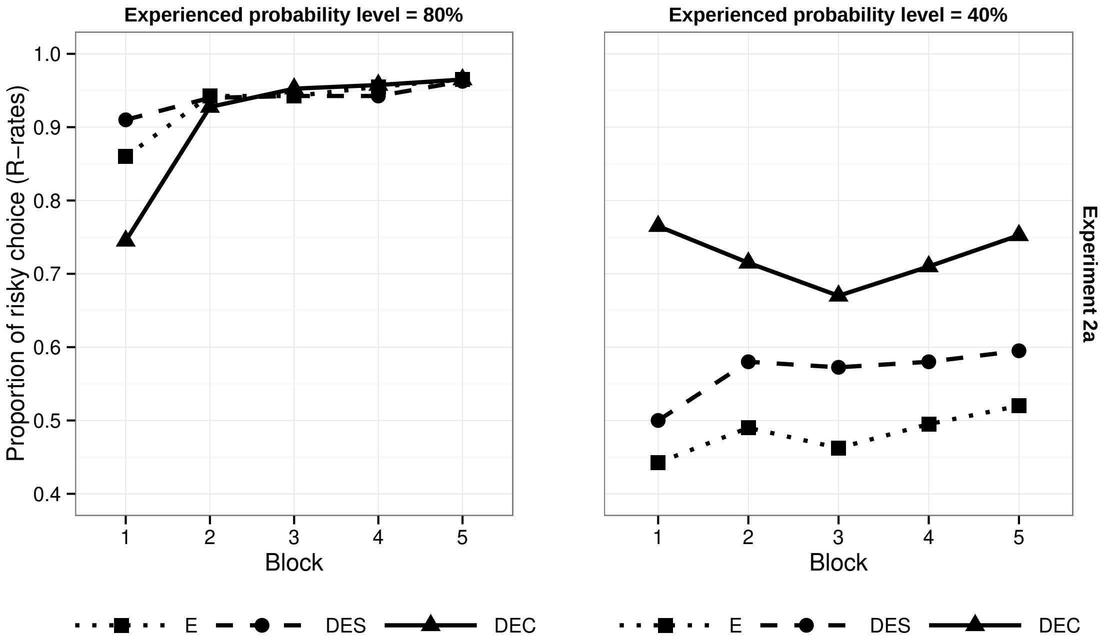


*Figure 2*. Evolution of the risky choice rate (R-rate) for each block of 20 trials for Experiment 2a. The left panel shows the results when the risky choice paid a reward of six points 80% of the time, while in the right panel the same reward was paid 40% of the time. The lines refer to the different descriptions presented to participants (E: experience only, no description; DES: description and experience same; DEC: description and experience conflicting).

**Experiment 2a.** The main effect of information presentation condition was significant for both probability levels. In the 80% probability level, R-rates were lower in the DEC condition ($\chi^2(2) = 12.29$, $p = .002$) but mostly influenced by the behavior in the first block (DEC: 75%; DES: 91%; E: 86%; $\chi^2(2) = 42.25$, $p < .001$). While in the 40% probability level, R-rates in the DEC condition were higher throughout the blocks, and away from randomness (DEC: 72%; DES: 57%; E: 48%; $\chi^2(2) = 23.20$, $p < .001$). The influence of conflicting descriptions was observed throughout all blocks in the 40% condition, but only in the first block in the 80% condition (Figure 2). The main effect of block was significant in the 80% probability condition, with R-rates increasing over time ($\chi^2(4) = 45.15$, $p < .001$), but not significant in the 40% condition ($\chi^2(4) = 7.57$, $p = .11$). The interactive effect of block and information condition was not significant in either probability condition ($\chi^2(8)$s$< 4.50$, $p$s$> .80$).

A post-hoc analysis for the last block was also conducted, for each of the two levels of probability, comparing the three types of information presentation conditions. In the 80% probability conditions, there were no significant differences between the R-rates according to condition (DEC: 97%; DES: 96%; E: 97%; $\chi^2(2) = 0.05, p = .98, \varphi_c = .01$). The availability of descriptive information, both congruent and conflicting, did not seem to move behavior away from that predicted by experience alone in the last block; its effect was restricted to the first block as mentioned before. In the 40% probability conditions, R-rates in the DEC condition were significantly higher than in the other two conditions (DEC: 75%; DES: 60%; E: 52%; $\chi^2(2) = 47.94$, $p < .001$, $\varphi_c = .20$). Conflicting descriptions influenced behavior significantly, in the direction predicted by the descriptive information provided. The descriptive information presented a probability of rewards for the risky option higher than experienced, so an increase in R-rates was expected. There was a marginally significant difference in R-rates between the DES and E conditions (DES: 60%; E: 52%; $\chi^2(1) = 4.56$, $p = .03$, $\varphi_c = .08$), albeit with a small effect size.

**Experiment 2b.** The main effect of information presentation condition was significant for the 40% probability levels ($\chi^2(2) = 13.33$, $p < .01$), with lower R-rates in the DEC condition, and away from randomness (DEC: 28%; DES: 37%; E: 40%). In contrast, the main effect of information presentation condition was not significant for the 20% probability levels ($\chi^2(2) = 1.66$, $p = .44$), with no significant differences in R-rates across conditions (DEC: 16%; DES: 18%; E: 16%). The influence of conflicting descriptions was observed in the 40% condition, but not in the 20% condition (Figure 3). The main effect of block was not significant in the 40% probability condition ($\chi^2(4) = 2.62$, $p = .62$), and significant in the 20% condition, with a reduction of R-rates over time ($\chi^2(4) = 42.08$, $p < .001$). The interactive effect of block and information condition was marginally significant in the 40% condition ($\chi^2(8) = 13.91$, $p = .08$) and significant in the 20% condition ($\chi^2(8) = 16.21$, $p = .04$). This effect was a result of the reduction in R-rates over time in the E conditions.

A post-hoc analysis for the last block was also conducted, for each of the two levels of probability, comparing the three types of information presentation conditions. In the 40% probability conditions, R-rates in the DEC condition were significantly lower than in the other two conditions (DEC: 26%; DES: 41%; E: 37%; $\chi^2(2) = 21.06$, $p < .001$, $\varphi_c = .13$). Conflicting descriptions influenced behavior significantly, in the direction predicted by the descriptive information provided. The descriptive information presented a probability lower than experienced, so a decrease in R-rates was expected. There were no significant differences between the DES and E conditions (DES: 41%; E: 37%; $\chi^2(1) = 1.52$, $p = .22$, $\varphi_c = .04$), therefore congruent descriptions did not significantly influence behavior. In the 20% probability conditions, there were no significant differences between the R-rates according to condition (DEC: 10%; DES: 15%; E: 11%; $\chi^2(2) = 5.64, p = .06, \varphi_c = .07$). The presence of different types of information did not seem to influence behavior strongly in this condition.

**Discussion.** The influence of conflicting description on behavior was observed in the 40% condition of Experiment 2a and the 40% condition of Experiment 2b. In these conditions, the conflicting description made the maximizing choice more attractive by increasing the difference in expected value between the risky and safe choices, and led to significant changes in the R-rates, away from random behavior, in the direction expected from the conflicting descriptions. In both of these conditions, the control R-rates (in the DES and E conditions) were close to 50%. In Experiment 2a, the descriptive probability of rewards of the risky option was higher than the experienced probability, and this led to an increase in R-rates. In Experiment 2b, the opposite occurred, with a lower descriptive probability leading to a reduction in the observed R-rates.

However no significant influence of conflicting description (relative to the E and DES conditions) was observed in the 80% condition of Experiment 2a and the 20% condition of

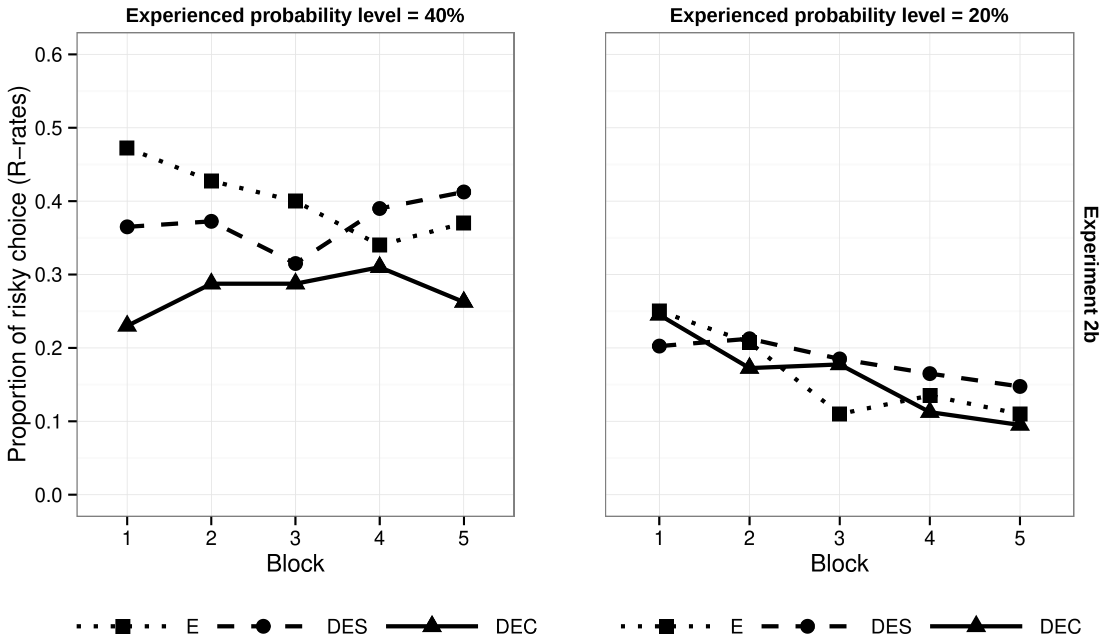


*Figure 3*. Evolution of the risky choice rate (R-rate) for each block of 20 trials for Experiment 2b. The left panel shows the results when the risky choice paid a reward of four points 40% of the time, while in the right panel the same reward was paid 20% of the time. The lines refer to the different descriptions presented to participants (E: experience only, no description; DES: description and experience same; DEC: description and experience conflicting).

Experiment 2b. Instead we observed a ceiling effect in Experiment 2a and a floor effect in Experiment 2b, with R-rates close to 100% in the former and close to 10% in the latter regardless of experimental condition. In comparison to Experiment 1, which provided the same probabilities of reward (five points with 80% and 20%), in Experiment 2a participants could earn more points (six points) and in Experiment 2b fewer points (four points). These changes led to an increase and decrease of R-rates respectively in comparison to Experiment 1, resulting in the ceiling and floor effects.

The opportunity cost of complying with warning labels seems to moderate behavior (Argo & Main, 2004) and also needs to be taken into account: In the 80% condition of Experiment 2a and the 20% condition of Experiment 2b, the expected value of one option was considerably higher than the other, and hence complying with the misleading description (which would lead to deviation from optimal behavior) was more costly. In addition, the influence of descriptive information might have been given less importance than that of experiential information, as previously suggested (Jessup et al., 2008; Lejarraga, 2010; Rogers et al., 2000; Shlomi, 2014). Descriptions, if discounted, would have less influence on behavior when the difference in expected values is higher, and this would explain the behavior observed. The discounting of descriptions will be further investigated in Experiment 3 and in the cognitive modeling section below.

## Experiment 3

In Experiment 3, we aimed to verify the boundaries of the influence of conflicting descriptive information. We propose that the influence of conflicting descriptions would be monotonically increasing at the center, but not at extreme levels of informational conflict. In extreme cases when information is highly implausible, it should be more readily discarded from the decision-making process. This could lead to a reduced marginal influence or even a contrasting effect at the extremes, similar to what has been found in research on anchoring (e.g., Chapman & Johnson, 1994), advice seeking (e.g., Yaniv, 2004), goal setting (e.g., Locke, 1982), and psychophysics (e.g., Brown, 1953). In order to check how a description's plausibility influences behavior, we created experimental conditions with varying levels of conflict: no conflict, plausible conflict and implausible conflict.

Plausibility was manipulated via the difference between the actual experienced and the verbally described frequencies of rewards. Across all conditions, the risky option returned rewards 50% of the time. In the two plausible conflict conditions, descriptions informed participants that rewards were paid either with a 25% or 75% probability, relatively close to the true experienced frequency of 50%, making the descriptions plausible explanations for the experience. In the two implausible conflict conditions, descriptions informed participants that rewards were paid either with a 1% or 99% probability, which were highly implausible given the experience. We expected participants to disregard the descriptions more easily in the implausible conflict conditions, therefore reducing the effect of their influence on their behavior.

### Method

**Design.** This experiment followed a between-subjects design with six experimental conditions, with manipulations of the descriptions that were provided to participants. Each participant was presented with only one type of description, assigned randomly from the following options: the experience-only (E) condition; the description-experience-same (DES) condition, in which the description matched the experience at 50% probability of receiving a reward; two *plausible* conflict ($\text{DEC}_p$) conditions, with descriptive probabilities of 25% and 75%; and two *implausible* conflict ($\text{DEC}_i$) conditions, with descriptive probabilities of 1% and 99%.

**Participants.** 240 participants (110 females; age: $M$ = 33 years, $SD$ = 11 years) were recruited on-line using Amazon's Mechanical Turk service. Participation was restricted to individuals whose location was defined as in the United States. There were 40 participants in each experimental condition. No participants were excluded from the analysis. Participants were paid a fixed amount of US$ 0.25 for participating and an additional bonus according to the outcomes of the choices they made during the experiment. Average bonus paid was US$ 0.80 ($SD$ = US$ 0.03)[6].

**Task.** The experimental paradigm was similar to that of Experiments 1 and 2, the only differences being the new values and probabilities for the risky option. In this experiment, the experience drew from the same underlying distributions across all conditions, with the risky option returning 4 points with 50% probability, and the safe option returning

[6]Because the expected value of the safe and risky options were matched, there was no significant difference in the bonus paid according to experimental condition ($F(1, 238) = 0.259, p = .611$).

2 points with 100% probability. The expected values of the risky and safe options were the same. The only between-subjects manipulation was the descriptive information. The new outcomes and button labels can be seen in Table 3. Points were converted to money at a rate of US$ 0.40/100 points. In addition, after participants had finished selecting between the two choices, the descriptions (if previously present) were hidden and a blank text box appeared under each button. Participants were asked to input their judgments for the actual experienced frequencies of rewards for each button, in the range of 0-100%, using the numbers in their keyboards. The removal of the descriptions from the screen was done to reduce any potential anchoring effect, and avoid participants from simply copying the descriptions as their answers. Participants completed the task in an average of 8.0 minutes (SD = 5.6 minutes).

Table 3
*Button labels according to condition in Experiment 3.*

| Description condition of risky option | Safe choice button label | Risky choice button label |
|---|---|---|
| Experience-only ($E_{50\%}$) | (blank) | (blank) |
| No conflict ($D_{50\%}E_{50\%}S$) | 2 points with 100% probability. | 4 points with 50% probability; Zero otherwise. |
| Plausible conflict 25% ($D_{25\%}E_{50\%}C_p$) | 2 points with 100% probability. | 4 points with 25% probability; Zero otherwise. |
| Plausible conflict 75% ($D_{75\%}E_{50\%}C_p$) | 2 points with 100% probability. | 4 points with 75% probability; Zero otherwise. |
| Implausible conflict 1% ($D_{1\%}E_{50\%}C_i$) | 2 points with 100% probability. | 4 points with 1% probability; Zero otherwise. |
| Implausible conflict 99% ($D_{99\%}E_{50\%}C_i$) | 2 points with 100% probability. | 4 points with 99% probability; Zero otherwise. |

**Results and Discussion**

As in Experiments 1 and 2, the main dependent variable was the average proportion of times individuals selected the risky choice (R-rate). The same analysis that was used in Experiments 1 and 2 was conducted for Experiment 3. The average judgments of the frequency of reward appearances were also analyzed using a one-way ANOVA by experimental condition.

The main effect of information presentation condition was significant. R-rates increased with higher description levels ($\chi^2(5) = 59.72$, $p < .001$). The main effect of block was not significant ($\chi^2(4) = 4.21$, $p = .38$), however the interaction of information and block was significant ($\chi^2(20) = 39.05$, $p = .007$). A post-hoc analysis for the last block was also conducted, comparing the R-rate in the DES condition against the five other types of information presentation conditions (Figure 4(a)). As before, presenting participants with congruent descriptions did not influence behavior in relation to no descriptions, in replication of the findings in Experiments 1 and 2 (DES: 48%; E: 50%; $\chi^2(1) = 1.00$, $p = .32$, $\varphi_c = .03$).

In the plausible conflict conditions of 25% and 75%, we observed an associated change

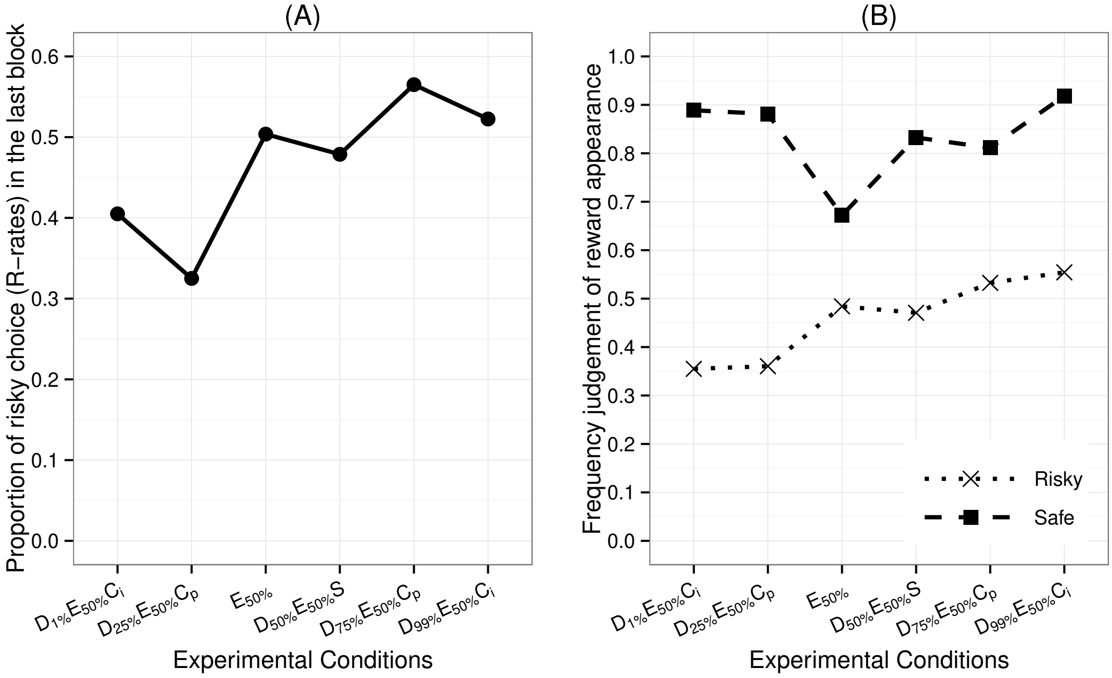


*Figure 4.* Results from Experiment 3. Left panel (A): Proportion of risky choice (R-rates) in the last block of 20 trials. Right panel (B): Frequency judgments for the appearance of rewards for the safe and risky choices. Across all conditions, participants experienced a reward of 4 points 50% of the time for the risky option and 2 points 100% of the time for the safe option. Experimental conditions refer to the descriptions presented to participants for the risky option ($E_{50\%}$: experience-only, no description, $D_{50\%}E_{50\%}S$: same description (50%), $D_{25\%}E_{50\%}C_p$: plausible conflicting description (25%), $D_{75\%}E_{50\%}C_p$: plausible conflicting description (75%), $D_{1\%}E_{50\%}C_i$: implausible conflicting description (1%), $D_{99\%}E_{50\%}C_i$: implausible conflicting description (99%)).

in the R-rates as predicted by the directionality of the description. In the case of lower probability of rewards in the risky option (25%) there was a reduction in the R-rate in comparison to the DES condition ($DEC_p$: 33%; DES: 48%; $\chi^2(1) = 39.34$, $p < .001$, $\varphi_c = .16$). In the case when participants were presented with a higher probability of reward (75%) we observed an increase in the R-rate ($DEC_p$: 57%; DES: 48%; $\chi^2(1) = 11.93$, $p < .001$, $\varphi_c = .09$). This suggests a monotonic influence of descriptions in the central conditions.

In the implausible conflict conditions of 1% and 99%, we observed a reversal in the influence of the conflicting descriptions. In the case of the 1% description, the R-rate was significantly higher than in the plausible 25% condition ($DEC_i$: 41%; $DEC_p$: 33%; $\chi^2(1) = 11.05$, $p < .001$, $\varphi_c = .08$) and still significantly lower than in the DES condition ($DEC_i$: 41%; DES: 48%; $\chi^2(1) = 8.82$, $p < .01$, $\varphi_c = .07$). In the case of the 99% description, the R-rate was lower than in the plausible condition, but only marginally significant ($DEC_i$: 52%; $DEC_p$: 57%; $\chi^2(1) = 2.91$, $p = .09$, $\varphi_c = .04$), and also only marginally significantly higher than in the DES condition ($DEC_i$: 52%; DES: 48%; $\chi^2(1) = 3.06$, $p = .08$, $\varphi_c = .04$).

The frequency judgment of reward appearances was also analyzed, using a one-way

ANOVA by experimental condition (Figure 4(b)). All participants experienced frequencies of rewards of 100% for the safe choice and 50% for the risky choice, which would have been their unbiased correct answers. For the safe choice, participants' judgments were not different across the individual description conditions ($M = 87\%; F(4, 195) = 1.587, p = .179$); in the E condition however, their judgments were significantly lower (E: 67%; $F(5, 234) = 5.851, p < .001$). There was also a significant difference in the judgments for the risky choice ($F(5, 234) = 8.952, p < .001$). This effect was analyzed with five polynomial contrasts: The linear, quadratic and cubic contrasts were all significant ($p$s<.001); while the two remaining higher order contrasts were not significant ($p$s>.52). This would indicate a sigmoid-shaped monotonically increasing judgment in relation to the description: participants presented with descriptions of higher probabilities of rewards responded with higher frequency judgments of the observed rewards, with diminishing sensitivities at the extremes (see Figure 4(b)). The individual frequency judgments for the risky choice were also significantly correlated to the individual R-rates ($r = .50, n = 238, p < .001$).

The lack of influence of congruent descriptions and the significant influence of conflicting descriptions on behavior were again observed in Experiment 3, replicating the results found in Experiments 1 and 2. In addition, plausible conflicting descriptions influenced R-rates in a monotonic way, with high described probabilities of rewards increasing R-rates, and vice-versa. However, in the case of implausible conflicts, a contrast effect was observed. A more extreme and more implausible described probability had a weaker effect on behavior than a less extreme but plausible one. If the description is highly unlikely to be a true representation of the experience, participants give it lower weight in their decision-making process. These differences in decision weights will be specified with a cognitive model in the next section.

## Cognitive Modeling Analysis

To further test if participants integrate the descriptive information into their decision-making processes, a set of cognitive computational models was fitted to the experimental data. If the descriptive information influenced human behavior, then a model that includes representations of both description and experience should fit better than a model that relies on experience alone. We therefore compared *experience-only* against *description-experience* models. Within the description-experience models, we tested two different approaches: a *fixed-weight* approach, in which the weights given to description are fixed over time and over conditions, and a *Bayesian-updating* approach, in which the weights given to description change over time according to the plausibility of the evidence observed in contrast with the description.

### The Models

The aim of fitting a cognitive model to the data was to assess and formalize how the two sources of information, descriptive and experiential, are combined. We did not aim to have an extensive comparison between different decision-making models in decision making paradigms (such as the one in Yechiam & Busemeyer, 2005).

We fitted three models to the behavioral data. They all share the same basic structure, which is defined by the final expected value $FEV_j(t)$ of each choice $j$ available to participants at time $t$:

$$FEV_j(t) = \xi_j(t) \cdot D_j + [1 - \xi_j(t)] \cdot E_j(t).$$

We propose that the two sources of information are combined via $\xi_j(t)$, a parameter which determines the weight given to description at each point in time for each option. A representation of the descriptive information is included via $D_j$, which is the expected value calculated from the descriptive information available to participants, using cumulative prospect theory (CPT; Tversky & Kahneman, 1992). The experience is represented by $E_j(t)$, which is the expected value calculated from the experiential information received by participants in the form of feedback up to trial $t$, based on a delta-rule reinforcement learning model.

### Description ($D_j$)

The subjective expected value of the descriptive information for choice $j$, $D_j$, was fixed over time and calculated as the CPT value based on the descriptions provided to participants in the button labels. According to Tversky and Kahneman (1992), the CPT value is calculated using the curvature parameter for values and weighting parameter for probabilities, $\nu$ and $\omega$ respectively,

$$D_j = \sum_m W(p_{jm}) V_{jm}^{\nu},$$

where $p_{jm}$ are the probabilities and $V_{jm}$ are the potential values for each outcome $m$ of option $j$; $\nu$ is the free parameter that determines the curvature of the value function ($0 \leq \nu \leq 1$), with lower values reducing the distance between extreme values of rewards; and $W(\cdot)$ is the probability weighting function. $W(\cdot)$ is defined as:

$$W(p) = \frac{p^{\omega}}{(p^{\omega} + (1-p)^{\omega})^{\frac{1}{\omega}}},$$

where $\omega$ is the free parameter ($0 \leq \omega \leq 5$) that determines the curvature of the probability weighting function. Values of $\omega$ below 1 lead to overweighting of rare events, while values above 1 lead to underweighting of rare events.

### Experience ($E_j(t)$)

A simple reinforcement learning model was used to fit the experimental choice data. This model has extensively been used before in research on repeated decisions from experience (Erev & Barron, 2005; Lejarraga & Gonzalez, 2011; Yechiam & Busemeyer, 2005; Yechiam & Rakow, 2012).

Firstly, observed outcomes are evaluated by a prospect-theory type of utility function (Kahneman & Tversky, 1979). The utility function $v_j(t)$ of option $j$ is defined as:

$$v_j(t) = [payoff_j(t)]^{\nu},$$

where $payoff_j(t)$ is equal to the payoff, in points, at each trial $t$ for each option $j$, and $\nu$ is same parameter that determines the curvature of the utility function for the description.

Secondly, expectancies for the value of rewards for each option are formed via a learning rule, which integrates the experienced feedback after each trial. The learning rule used was a delta rule, which uses a learning rate that determines how much the new information gathered via feedback, in the form of prediction error, influences the updating of the expectancies at each trial (Sutton & Barto, 1998; Yechiam & Busemeyer, 2005). Feedback observed is integrated after each trial, to arrive at the experienced expectancy $E_j(t)$ for option $j$ at time $t$:

$$E_j(t) = E_j(t-1) + \phi \cdot \{\delta_j(t) + \gamma \cdot [1 - \delta_j(t)]\} \cdot [v_j(t) - E_j(t-1)],$$

where $\phi$ is the free learning rate parameter ($0 \leq \phi \leq 1$), which is a weight given to new information observed, with lower values resulting in slower learning. The free parameter $\gamma$, ($0 \leq \gamma \leq 1$), denotes the weight associated with the feedback of the foregone option, such that when $\gamma$=1 the foregone and observed payoffs are weighted the same, and when $\gamma$=0 foregone payoffs are disregarded. The variable $\delta_j(t)$ is a dummy variable, which is equal to one if option $j$ was chosen on trial $t$, and zero otherwise.

## The weight given to description ($\xi_j(t)$)

Three different approaches were used to combine description and experience by manipulating the parameter $\xi_j(t)$, which is the weight given to description: experience-only, fixed-weight and Bayesian-updating. Each of these approaches was fitted to the behavioral data individually, and the fit results were then compared.

**Experience-only model.** In the experience-only model, no representation of description was included, with $\xi_j(t)$ fixed to zero across all trials and options. Therefore, the final expectancy $FEV_j(t)$ was defined to be equal to the expectancy derived from experience alone, via the reinforcement learning model, $E_j(t)$. This model assumes that descriptions do not influence the decision making process.

**Fixed-weight model.** In the fixed-weight model, the weight given to description, $\xi_j(t)$, was set as a free parameter, $\xi$, constant across all trials, options and conditions ($0 \leq \xi \leq 1$). This model assumes that description influenced the individual choices, but with a fixed weight that did not depend on the experimental condition and did not change over time.

**Bayesian-updating model.** Based on our behavioral results, we observed that descriptions influenced the decision-making processes in different ways in each experimental condition. More plausible descriptions seemed to have a stronger influence on decisions than less plausible ones. We propose a Bayesian-updating model in which the weight given to description, $\xi_j(t)$, equals the subjective probability that the description is true on that trial, given the evidence observed thus far. In this model, the weights given to description will differ for each option and change over time, according to the experimental conditions.

The Bayesian model assumes that, at each trial, either the description is true, denoted as $\mathcal{D}_j(t)$, or the task is in a different state, denoted as $\mathcal{E}_j(t)$, where the probabilities of rewards are not as described but instead are as experienced. From trial to trial, the state of the task can change, such that if the description was true on the previous trial, it no longer is true on the next. The task of a participant is twofold: to determine whether the task is in state $\mathcal{D}_j(t)$ or $\mathcal{E}_j(t)$, and to estimate the relevant probabilities of the outcomes when the task is in state $\mathcal{E}_j(t)$. The probabilities in state $\mathcal{D}_j(t)$ do not need to be estimated as they come from the description itself. Initially it makes sense to rely on the description, as there is no information to estimate the probabilities of winning in the other state. Over time however, it is possible to learn that the true probabilities of the outcomes are different than those described, in which case the weight given to description should diminish. With this model, less plausible conditions lead to lower weights given to description than more plausible conditions, thus making this approach more adaptive to the non-monotonicity observed in the behavioral results than using a fixed weight (Figure 5).

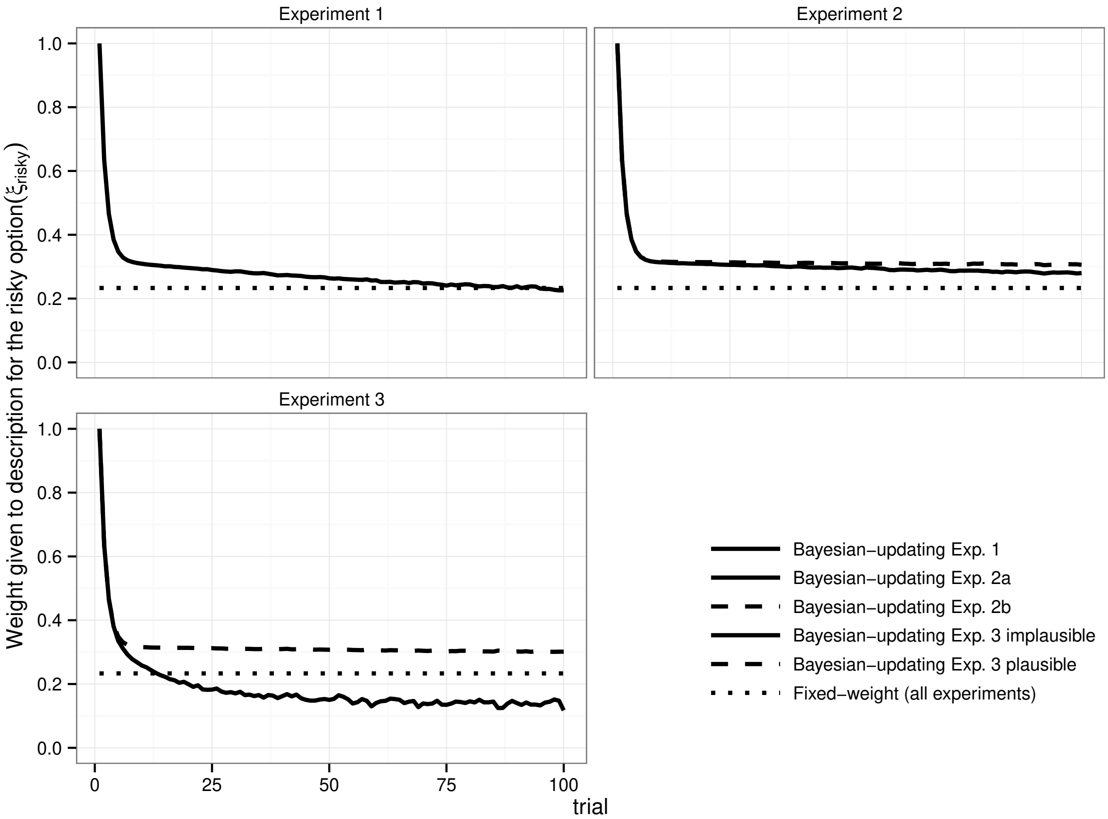


*Figure 5*. Average weights given to description for the risky option ($\xi_{risky}$) in the two different description-experience models, based on the best fit parameters, for the conflicting (DEC) conditions only. The Bayesian-updating approach shows how the weight changes over time according to the accumulation of evidence, and the plausibility of the description, as measured by the difference between descriptive and experiential information.

We set the weight given to the description as the Bayesian predicted probability that the description is true on that trial, $\xi_j(t) = p(\mathcal{D}_j(t)|k_j(1:t-1))$. This probability is based on all the observations made up to the previous trial $t-1$, denoted here as $k_j(1:t-1)$, which are all the observed outcomes of option $j$, whether rewards were obtained or not, from trial 1 to trial $t-1$. We will simplify this notation by using a subscript to denote the information used to calculate the probability, with $p_{t-1}(\mathcal{D}_j(t)) = p(\mathcal{D}_j(t)|k_j(1:t-1))$, and analogously using $p_t$ to include all the information from trial 1 to trial $t$, $k_j(1:t)$. We also assume that the state of the task can change over trials according to the transition probabilities $\kappa_{\mathcal{D}}$ and $\kappa_{\mathcal{E}}$ which are free parameters (range: 0-1). $\kappa_{\mathcal{E}}$ is the probability that state $\mathcal{D}_j$ is true in trial $t$ if state $\mathcal{E}_j$ was true in trial $t-1$:

$$p(\mathcal{D}_j(t)|\mathcal{E}_j(t-1)) = \kappa_{\mathcal{E}}$$

and conversely for $\kappa_{\mathcal{D}}$:

$$p(\mathcal{E}_j(t)|\mathcal{D}_j(t-1)) = \kappa_{\mathcal{D}}.$$

The effect of the transition probabilities is to change the prior distributions at each trial. The prior probability that the description is true on trial $t+1$ for option $j$ is then:

$$p_t(\mathcal{D}_j(t+1)) = p_t(\mathcal{D}_j(t)) \cdot (1-\kappa_{\mathcal{D}}) + p_t(\mathcal{E}_j(t)) \cdot \kappa_{\mathcal{E}}.$$

Since one of the two states has to be true at any point, the two probabilities $p_{t-1}(\mathcal{D}_j(t))$ and $p_{t-1}(\mathcal{E}_j(t))$ are complimentary and the prior probability that state $\mathcal{E}_j$ is true is simply $p_{t-1}(\mathcal{E}_j(t)) = 1 - p_{t-1}(\mathcal{D}_j(t))$. The prior $p_{t-1}(\mathcal{D}_j(t))$ and its compliment $p_{t-1}(\mathcal{E}_j(t))$ can be used to calculate the Bayesian posterior:

$$p_t(\mathcal{D}_j(t)) = \frac{p(k_j(t)|\mathcal{D}_j(t))p_{t-1}(\mathcal{D}_j(t))}{p(k_j(t)|\mathcal{D}_j(t))p_{t-1}(\mathcal{D}_j(t)) + p(k_j(t)|\mathcal{E}_j(t))p_{t-1}(\mathcal{E}_j(t))}.$$

We set the initial prior, $p_0(\mathcal{D}_j(1))$ at the first trial, to be equal to one, since at that point there was no information experienced so far, and participants had to rely solely on the descriptive information provided to base their decisions.

According to the descriptive information $\mathcal{D}_j(t)$, the probability of a win $k_j(t)$ observed in $t$ if the description holds on trial $t$ is

$$p(k_j(t)|\mathcal{D}_j(t)) = pd_j^{k_j(t)}(1-pd_j)^{(1-k_j(t))},$$

where $pd_j$ is the probability for a obtaining a reward for option $j$ as provided by the description. For example, in Experiment 1, $pd_j$ could be either 0.2 or 0.8 for the risky options, depending on the experimental conditions, and 1.0 for the safe option.

The relevant probabilities for state $\mathcal{E}_j(t)$ have to be learned from experience. Assuming that people start with a Beta prior over these probabilities, the posterior distributions over these probabilities are also Beta distributions, and the likelihood of outcome $k_j(t)$ if the task is in state $\mathcal{E}_j(t)$ is

$$p(k_j(t)|\mathcal{E}_j(t)) = \frac{\mathrm{B}(\alpha_j(t-1) + k_j(t), \beta_j(t-1) + 1 - k_j(t))}{\mathrm{B}(\alpha_j(t-1), \beta_j(t-1))},$$

where B is the Beta function, and $\alpha_j(t)$ and $\beta_j(t)$ are its parameters, which are updated as follows, with new experiential evidence $k_j(t)$ weighed by the probabilities given to $\mathcal{E}_j(t)$, for $t \geq 1$:

$$\alpha_j(t+1) = \alpha_j(t) + p_t(\mathcal{E}_j(t)) \cdot k_j(t);$$
$$\beta_j(t+1) = \beta_j(t) + p_t(\mathcal{E}_j(t)) \cdot (1-k_j(t)).$$

For $t = 1$ the values of $\alpha_j(1)$ and $\beta_j(1)$ are defined by the initial expected value of the Beta distribution which is set to the probability provided in the description, $\alpha_j(1)/(\alpha_j(1) + \beta_j(1)) = pd_j$. We constrained the total weight of the initial Beta prior as a free parameter $\mathcal{S} = \alpha_j(1) + \beta_j(1)$ for all options $j$ ($1 \leq \mathcal{S} \leq 500$[7]). Higher values of $\mathcal{S}$ led to a slower accumulation of new evidence towards $p(\mathcal{E}_j(t))$.

## Choice rule

After description and experience were integrated into the final expectancy calculation, $FEV_j(t)$, the choice rule used was a time-independent soft-max rule (Yechiam & Busemeyer, 2005) that combined the $FEV_j$ across all options at each trial to determine the probability of choosing option $j$ among all options $J$:

$$Pr[Choice(t+1) = j] = \frac{e^{\theta \cdot FEV_j(t)}}{\sum_J e^{\theta \cdot FEV_J(t)}},$$

where $\theta$ is the choice sensitivity free parameter, $(0 \leq \theta \leq 1)$. If $\theta = 0$, the model randomly guesses between the expectancies regardless of their values, while higher values of $\theta$ will lead to more deterministic maximization behavior.

## Model fitting

Data sets containing 100 simulated participants were generated for each of the 4 experiments $\times$ 6 conditions[8], with the same methodology used to generate actual data sets for the experiments. A total of 2,400 modeled simulated participants were confronted with 608 observed human participants. All simulated participants across all experiments and underlying experimental conditions shared the same set of free parameters. The best fit parameters were found by minimizing the log-likelihood between the average observed proportions of risky choice and the average model-predicted risky choice for each of the individual conditions separately, with each condition receiving the same weight (Erev & Barron, 2005). Because of the different number of parameters between the models, the Bayesian Information Criterion (BIC), which penalizes for additional parameters, was calculated to compare the models, $BIC = LL + f \cdot ln(N)$, where $f$ is the number of free parameters and $N$ is the number of fitted observations for each evaluation. Lower BIC values represent better fitting models.

## Model Evaluation and Results

Three models were evaluated: the experience-only model, which did not account for the influence of descriptive information, with four free parameters; the fixed-weight model, which assumed a single fixed weight across all trials and all conditions, with six free parameters; and the Bayesian-updating model, with its reducing weight given to description over time according to the feedback received, with eight free parameters.

[7]We also fitted a model in which the experience was initially set to follow a Beta(1,1) distribution, $\alpha_j(1) = \beta_j(1) = 1$, with an expected value of 0.5. This model also outperformed the fixed-weight model but by a smaller margin.

[8]The description-only (D) condition from Experiment 1 was excluded from the cognitive modeling, as it included a single-shot decision without experience.

The best fit parameters were relatively consistent across the different models (Table 4). In the best-performing Bayesian-updating model, the modeled influence of description started at 1 and converged towards a stable level ranging between 0.15 and 0.30 according to the plausibility of the descriptions (Figure 5). It seems that even after many trials and despite the large amount of evidence gathered via feedback, participants were still taking the descriptive information into consideration, albeit at a discounted level; experience never gained participants' full attention. This could be explained by the constant presence of the descriptive information on the buttons, which might have continuously reinforced its influence. Barron et al. (2008) found a similar lingering influence of descriptive information even when descriptions were only presented briefly. In comparison, the fixed-weight model predicted a constant weight given to description of $\xi = 0.23$ throughout all trials and conditions.

Table 4
*Best fit parameters of the three cognitive models. n.a. = not applicable. Note: the weight given to description ($\xi$) in the experience-only model was fixed to zero.*

| Parameter | Bayesian updating | Fixed weight | Experience only |
|---|---|---|---|
| $\nu$ (curvature of value function) | 0.91 | 0.93 | 0.92 |
| $\omega$ (curvature of probability function) | 1.02 | 1.35 | n.a. |
| $\phi$ (learning rate) | 0.08 | 0.08 | 0.09 |
| $\gamma$ (foregone's weight) | 0.99 | 0.96 | 1.00 |
| $\theta$ (choice sensitivity) | 0.83 | 0.80 | 0.74 |
| $\xi$ (description's weight) | n.a. | 0.23 | *zero (fixed)* |
| $\kappa_{\mathcal{D}}$ (switch rate for description) | 0.17 | n.a. | n.a. |
| $\kappa_{\mathcal{E}}$ (switch rate for experience) | 0.36 | n.a. | n.a. |
| $\mathcal{S}$ (Initial Beta prior sum for experience) | 454 | n.a. | n.a. |
| Total number of free parameters | 8 | 6 | 4 |

The results of the model fitting analysis were in line with the behavioral results (Table 5). The mean Bayesian Information Criteria (BIC) values for the two description-experience models were substantially smaller than those for the experience-only model (range: 16-17% smaller overall; lower BIC values represent better fitting models). Therefore, models that included the descriptive information provided a better fit for the observed behavior than a model that did not include the influence of description (Figure 6). The Bayesian-updating model, which allows for the plausibility of the description as a source of information to influence the weights given to description and experience, was the best fitting model overall, with a 17% lower mean BIC value than the experience-only model, and 2% lower than the fixed-weight model. Among individual conditions, the highest reduction in mean BIC values was obtained in the DEC conditions, with the Bayesian-updating model 30% lower than experience-only and 4% lower than fixed-weight models, showing the influence of the conflicting information on participants' choices.

The reduction in mean BIC values across all the DES conditions was lower at only 6% when comparing the Bayesian-updating with the experience-only model. This finding is similar to that found in Lejarraga and Gonzalez (2011), where a model that did not include descriptions fit the behavioral data relatively well, likely due to the equivalent information provided by both sources, and where the addition of description did not improve the model

Table 5
*Mean BIC values for the experience-only and description-experience cognitive models fitted on the three groups of experimental conditions: description-experience-conflict (DEC), description-experience-same (DES) and experience-only (E). Values in brackets are the differences in relation to the base model at the top of each column. Lower BIC values represent better fitting models.*

| Model | Overall | DEC | DES | E |
|---|---|---|---|---|
| Experience-only (Base) | 503 | 627 | 417 | 412 |
| Fixed-weight | 425 (-16%) | 454 (-28%) | 400 (-4%) | 410 (-1%) |
| Bayesian updating | 417 (-17%) | 437 (-30%) | 394 (-6%) | 410 (-1%) |

fit substantially. While we did observe a small improvement with the description-experience model, most of the reduction in BIC values came from the first few trials. Comparing the mean BIC values of the DES conditions between the Bayesian-updating and the experience-only models, there was a 39% reduction in the first five trials, a 9% reduction in the next 15 trials, and only a small 0.1% reduction in the last 80 trials. While the experiential information had to be learned over many trials, the descriptive information was available from the beginning of the task. Thus, if participants were only relying on experience, they should have chosen randomly in the first few trials until enough information was learned to steer their decisions, while if participants used descriptions, they could rely on that information to direct their earlier choices. The reduction in BIC values observed in the DES condition comes from this earlier availability of descriptive information which helps explain participants' choices in the initial trials. According to the behavioral and modeling results, participants chose in accordance with the descriptions available, showing the influence of the congruent information as well on their choices.

The largest improvement in model fit among the two description-experience models was in the DEC conditions of Experiment 3: The mean BIC values of the Bayesian-updating model were 8% lower than those in the fixed-weight model. In this experiment, we manipulated for the plausibility of the descriptive information. While a fixed-weight model had a single $\xi$ parameter for all conditions, the Bayesian-updating model was able to adapt to the different levels of plausibility of the information given (Figure 5), and therefore did a better job to explain behavior than the fixed-weight model. This is further evidence that, when the descriptions are implausible, participants give them lower weights in their decision processes, and vice versa, as predicted by the best fit Bayesian-updating model.

## General Discussion

Our aim was to shed additional light on how individuals combine information acquired by description and by experience when making decisions. The current experiments have shown that the choice behavior of individuals exposed to a combination of description and experience is influenced by both sources of information. In support of this, a cognitive model that included both the descriptive and experiential information fitted the observed behavior better than a model that relied on experience alone. However, both observed behavior and fitted cognitive models were significantly influenced by descriptions only when they provided novel information and not when the description and experience transmitted the same information. In addition, the influence of experience dominated the influence of

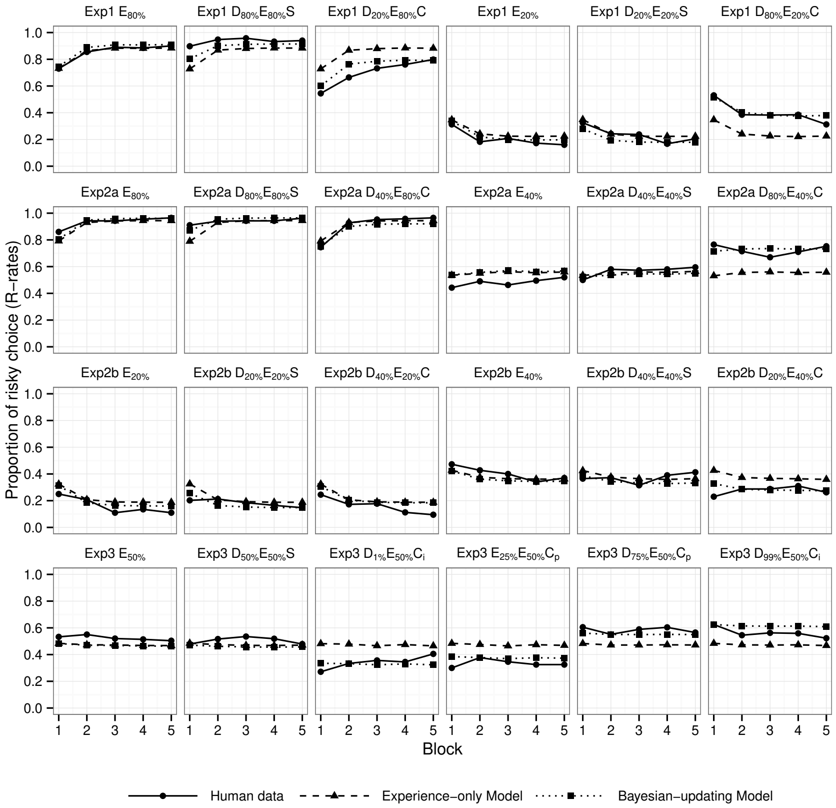


*Figure 6*. Comparison of observed human data, and two cognitive models, for each of the 24 experimental conditions. The description-experience model shown is the Bayesian-updating model, which is closer to human behavior than the experience-only model, in particular in the conflicting conditions. Each row shows data for a separate experiment: from top to bottom, Experiments 1, 2a, 2b and 3. The subscripts after each D and E indicate the probability of rewards for the risky choice in that condition for the description and experience respectively, while S stands for same description, C for conflicting description, $C_p$ for plausible conflict and $C_i$ for implausible conflict.

description, likely driven by the accumulation of experiential information over time. This caused the influence of descriptive information to be increasingly discounted by individuals; however, descriptions were still taken into consideration even after many trials. We also observed that the plausibility of the description has an effect on behavior: only plausible descriptions influenced behavior monotonically, with a reversal of the effect in the case of highly implausible descriptions.

Previous research has shown that information provided by experience can overwhelm descriptions, making experience the preferred source of information (Jessup et al., 2008). When Lejarraga and Gonzalez (2011) exposed participants to a combination of descriptive and experiential information, they observed that choice behavior could be explained by experience alone, as if the descriptive information was neglected. Based on the findings from the current experiment, an alternative explanation can be considered. Because in Lejarraga and Gonzalez's experiment both the description and the experience carried the same information, the description might not have been actively disregarded. Instead, it is possible that because it did not add any relevant information that could not be inferred from feedback, it did not lead to any observable differences in behavior. In contrast, Barron et al. (2008) used a paradigm with partial descriptions, which alerted participants to the presence of a rare negative event. Since most of the participants in their study did not experience this rare event, the description provided novel information, which in turn influenced behavior. The addition of descriptions that carried the same information as experience to the paradigm was therefore not enough to shift behavior - especially under the assumption that it is ultimately the choice mechanism paradigm (e.g., one-choice/single-outcome versus repeated-choice/multiple-outcomes; the presence of feedback), not the type of information presentation, that generates the behavioral differences in the description-experience gap (Camilleri & Newell, 2013a; Jessup et al., 2008).

Instead of individuals neglecting specific sources of information, participants might integrate experience with prior beliefs about the outcomes of their choices (Rakow & Newell, 2010), such that different weights are given to each source of information, depending on their relevance. Prior beliefs could be in the form of descriptions, or come from memory in pure experience-only conditions. Our results indicate that participants were influenced by descriptive information, albeit at a discounted rate in comparison to experience. One factor behind the apparent discounting of descriptive information might be the accumulation of experiential evidence in the form of feedback after each trial. Alternatively, the discounting could be explained by the higher costs associated with processing descriptive information (Lejarraga, 2010). When presented with both descriptive and experiential information, individuals have to deal with competing sources of control (Fantino & Navarro, 2012). In addition, there is strong research evidence supporting the evolutionary adaptation of human-decision making to dynamic environments, which would justify the preference for experiential information because it reacts more quickly to changes in the environment (for a review, see Hertwig & Pleskac, 2010).

Even though experiential information was dominant, the discounted influence of descriptive information still remained after many trials. The influence of experience appeared to grow steeply in the first few trials, but quickly reached an asymptotic level where it remained for the remainder of the task. Even after many trials, participants still behaved overall as if description received around a quarter of their decision weight, although this proportion was influenced by the plausibility of the information, with implausible descriptions receiving lower weights. This finding has implications for research of warning labels: these are commonly disregarded by individuals, or have limited impact on behavior (e.g., Argo & Main, 2004; Wagenaar, Hudson, & Reason, 1990). This phenomenon could be driven by the heavy discounting of descriptive information, as suggested by Rogers et al. (2000), and shown here in our modeling results. This implies that where a strong dominant choice is prescribed by experience, the reduced salience of descriptive information might not be enough to sway behavior, as observed in Experiment 2. To counteract this apparent discounting effect, designers of warning labels might be tempted to exaggerate risks and appeal to emotions and personal experience in order to increase compliance. However highly exaggerated descriptions might become implausible and could have the opposite effect on behavior, as observed in Experiment 3.

Further exploration of description-plus-experience paradigms, looking at how descriptive and experiential sources of information are integrated, is crucial for understanding decision-making. In particular, future research should focus on what determines the weights given to description and experience, for example, the trustworthiness of different sources of information, and the importance of descriptions in more complex decision spaces. In some cases, individuals might discount the experience, and give higher weights to descriptions. Decision-making biases originally found in descriptive paradigms are not always replicated in experiential paradigms, and in some instances reversals are observed (e.g., Barron & Erev, 2003; Fantino & Navarro, 2012; Hau et al., 2010). These decision-making biases have been widely explored in applications such as behavioral interventions, social marketing and governmental policy-making. Perhaps the reason why some of these attempted manipulations are not successful is because the research behind them is based on descriptive paradigms and applied in experiential settings. More relevant research would be based on real-life experiential situations and would allow for the influence of descriptions and other prior beliefs by giving them appropriate weightings.

## References

Argo, J. J., & Main, K. J. (2004). Meta-analyses of the effectiveness of warning labels. *Journal of Public Policy & Marketing*, *23*, 193–208. doi: 10.1509/jppm.23.2.193.51400

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 warnings' impact: Sex, drugs, and rock-n-roll. *Organizational Behavior and Human Decision Processes*, *106*, 125–142. doi: 10.1016/j.obhdp.2007.11.002

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4.* R package version 1.1-6. Retrieved from `http://cran.r-project.org/package=lme4`

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15. doi: 10.1016/0010-0277(94)90018-3

Brown, D. R. (1953). Stimulus-similarity and the anchoring of subjective scales. *The American Journal of Psychology*, *66*, 199–214. doi: 10.2307/1418726

Camilleri, A. R., & Newell, B. R. (2009). Within-subject preference reversals in description- and experience-based choice. In N. Taatgen, J. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (pp. 449–454). Austin, TX: Cognitive Science Society.

Camilleri, A. R., & Newell, B. R. (2011). Description- and experience-based choice: Does equivalent information equal equivalent choice? *Acta Psychologica*, *136*, 276–284. doi: 10.1016/j.actpsy.2010.11.007

Camilleri, A. R., & Newell, B. R. (2013a). The long and short of it: Closing the description-experience "gap" by taking the long-run view. *Cognition*, *126*, 54–71. doi: 10.1016/j.cognition.2012.09.001

Camilleri, A. R., & Newell, B. R. (2013b). Mind the gap? Description, experience, and the continuum of uncertainty in risky choice. In N. Srinivasan & P. Chandrasekhar (Eds.), *Progress in brain research* (pp. 55–72). Oxford, UK: Elsevier. doi: 10.1016/B978-0-444-62604-2.00004-6

Carson, J., & Mannering, F. (2001). The effect of ice warning signs on ice-accident frequencies and severities. *Accident Analysis & Prevention*, *33*, 99–109. doi: 10.1016/S0001-4575(00)00020-8

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making*, *7*, 223–242. doi: doi: 10.1002/bdm.3960070402

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674. doi: 10.1126/science.2648573

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

Ert, E., & Erev, I. (2007). Replicated alternatives and the role of confusion, chasing, and regret in decisions from experience. *Journal of Behavioral Decision Making*, *20*, 305–322. doi: 10.1002/bdm.556

Fantino, E., & Navarro, A. (2012). Description-experience gaps: Assessments in other choice paradigms. *Journal of Behavioral Decision Making*, *25*, 303–314. doi: 10.1002/bdm.737

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.

Glasgow, R. E., Green, L. W., Klesges, L. M., Abrams, D. B., Fisher, E. B., Goldstein, M. G., ... Orleans, C. T. (2006). External validity: We need to do more. *Annals of Behavioral Medicine*, *31*, 105–108. doi: 10.1207/s15324796abm3102_1

Glöckner, A., Fiedler, S., Hochman, G., Ayal, S., & Hilbig, B. E. (2012). Processing differences between descriptions and experience: A comparative analysis using eye-tracking and physiological measures. *Frontiers in Psychology*, *3*, 173. doi: 10.3389/fpsyg.2012.00173

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

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., & 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., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*, 225–237. doi: 10.1016/j.cognition.2009.12.009

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: An analysis of decision under risk. *Econometrica*, *47*, 263–292. doi: 10.2307/1914185

Kamal, A. H., & Peppercorn, J. M. (2013). The generalizability paradox within palliative care clinical trials. *Annals of Palliative Medicine*, *2*, 101–104. doi: 10.3978/j.issn.2224-

5820.2013.02.06

Knox, W. B., Otto, A. R., Stone, P., & Love, B. C. (2012). The nature of belief-directed exploratory choice in human decision-making. *Frontiers in Psychology*, *2*, 398. doi: 10.3389/fpsyg.2011.00398

Lejarraga, T. (2010). When experience is better than description: Time delays and complexity. *Journal of Behavioral Decision Making*, *23*, 100–116. doi: 10.1002/bdm.666

Lejarraga, T., & Gonzalez, C. (2011). Effects of feedback and complexity on repeated decisions from description. *Organizational Behavior and Human Decision Processes*, *116*, 286–295. doi: 10.1016/j.obhdp.2011.05.001

Locke, E. A. (1982). Relation of goal level to performance with a short work period and multiple goal levels. *Journal of Applied Psychology*, *67*, 512–514. doi: 10.1037/0021-9010.67.4.512

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic Bulletin & Review*, *14*, 1133–1139. doi: 10.3758/BF03193102

R Core Team. (2014). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from `http://www.r-project.org/`

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

Rakow, T., & Miler, K. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with nonstationary payoffs. *Memory & Cognition*, *37*, 985–1000. doi: 10.3758/MC.37.7.985

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14. doi: 10.1002/bdm.681

Rakow, T., Vincent, C., Bull, K., & Harvey, N. (2005). Assessing the likelihood of an important clinical outcome: new insights from a comparison of clinical and actuarial judgment. *Medical Decision Making*, *25*, 262–282. doi: 10.1177/0272989X05276849

Rogers, W. A., Lamson, N., & Rousseau, G. K. (2000). Warning research: An integrative perspective. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *42*, 102–139. doi: 10.1518/001872000779656624

Shlomi, Y. (2014). Subjective integration of probabilistic information from experience and description. *Judgment and Decision Making*, *9*, 491–499.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in Cognitive Science*, *7*, 351–367. doi: 10.1111/tops.12145

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

Tversky, A., & Kahneman, D. (1992). Advances in Prospect-Theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. doi: Doi 10.1007/Bf00122574

Wagenaar, W. A., Hudson, P. T. W., & Reason, J. T. (1990). Cognitive failures and accidents. *Applied Cognitive Psychology*, *4*, 273–294. doi: 10.1002/ acp.2350040405

Weber, E. U. (2006). Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climatic Change*, *77*, 103–120. doi: 10.1007/s10584-006-9060-3

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes*, *93*, 1–13. doi: 10.1016/j.obhdp.2003.08.002

Yechiam, E., & Busemeyer, J. R. (2005). Comparison of basic assumptions embedded in learning models for experience-based decision making. *Psychonomic Bulletin & Review*, *12*, 387–402. doi: 10.3758/BF03193783

Yechiam, E., & Rakow, T. (2012). The effect of foregone outcomes on choices from experience. *Experimental Psychology*, *59*, 55–67. doi: 10.1027/1618-3169/a000126