

# The Description/Experience Gap in the Case of Uncertainty

**Horacio Arlo-Costa**  
Carnegie Mellon University,  
Philosophy,  
hcosta@andrew.cmu.edu

**Cleotilde Gonzalez**  
Carnegie Mellon University,  
SDS<sup>†</sup>,  
coty@cmu.edu

**Varun Dutt**  
Carnegie Mellon University,  
EPP\*,  
varundutt@cmu.edu

**Jeffrey Helzner**  
Columbia University,  
Philosophy,  
jh2239@columbia.edu

## Abstract

We present empirical evidence indicating the existence of a description/experience gap for decisions under uncertainty. The nature of the gap is different than the one arising in the case of risk but both phenomena depend essentially on the use of limited sampling in experience. While subjects are ambiguity averse in description they are robustly ambiguity seeking in experience. A probabilistic explanation of this effect is provided as well as conjectures about the possibility of studying the effect with descriptive theories like Cumulative Prospect Theory.

**Keywords.** uncertainty, descriptive, normative, experience, description

## 1 Background

Traditionally, beliefs and desires are represented by subjective probabilities and utilities, respectively, and these subjective probabilities and utilities are combined in the calculation of expectations. This expected utility tradition is dominant within the decision sciences, extending from cases of decision making under risk, where objective probabilities are available to the decision maker, to cases of decision making under uncertainty, where information about objective probabilities is scarce [17].

The familiar axiomatizations of the expected utility hypothesis, from von Neumann-Morgenstern to Anscombe-Aumann to Savage (see [15] for an introductory presentation), are usually interpreted normatively, but they also serve as a diagnostic tool in that systematic deviations from their requirements are interpreted as pathology exhibited in human behavior and in need of explanation. The following example, one among a class of examples made famous by Allais [1], serves to illustrate the point: The subject is pre-

sented with two decision problems, each consisting of a pair of risky alternatives. In the first problem the subject is given a choice between a lottery  $A$  that pays \$4000 with probability 0.8 and \$0 with probability 0.2 and a lottery  $B$  that pays \$3000 with probability 1. In the second problem the subject is given a choice between a lottery  $C$  that pays \$4000 with probability 0.2 and \$0 with probability 0.8 and a lottery  $D$  that pays \$3000 with probability 0.25 and \$0 with probability 0.75. It has been observed that a significant number of subjects choose  $B$  in the first decision problem and  $C$  in the second decision problem. Assuming that such choices reveal strict preferences they are incompatible with the expected utility hypothesis: if  $B$  is strictly preferred to  $A$ , then the expected utility hypothesis requires a strict preference for the compound lottery that rewards  $B$  with probability 0.25 and \$0 with probability 0.75 over the compound lottery that rewards  $A$  with probability 0.25 and \$0 with probability 0.75 and, moreover, the expected utility maximizer must be indifferent between the first of these compound lotteries and  $D$  as well as between the second of these compound lotteries and  $C$ .

How are these observed deviations from expected utility theory to be interpreted? More generally, what is the significance of such deviations? According to one important class of interpretations such observed deviations are evidence that the normative theories at issue are not adequate when it comes to describing the decisions of human agents – but remain nonetheless valid normatively. According to another class of interpretations such deviations can be evidence that the theory being violated is inadequate as a normative theory of decision making – this is essentially the sort of interpretation that Ellsberg took in response to the violations that he made famous in connection with Savage’s theory [6]. For now we will focus on the first class of interpretations that was mentioned. Work on this class of has been dominated by two schools. Essential examples of the first of these schools can be found in the previously mentioned work of Simon [18]

---

\*Engineering and Public Policy.

† Social and Decision Science.

and Gigeren! zer [11]. A basic theme of such work is that deviations from a normative theory such as expected utility maximization are often just the result of computational limitations and that such deviations are not necessarily a sign of irrationality. Essential examples of the second of these schools is provided by the work of Kahneman and Tversky [14]. A basic theme of such work is that human decision processes, much like human senses, are subject to illusions and that these illusions lead to systematic deviations from expected utility and related norms. Work done in both these schools is potentially significant. For now we will focus on work done in the second of the two schools that were mentioned.

Let us consider what is perhaps the most well-known theory from this second school that attempts to address deviations from expected utility theory such as those associated with the Allais-type example mentioned previously. Roughly, *prospect theory* posits two phases of decision making. The first of these is an editing phase during which various operations (e.g., coding) are applied to the information that is available to the decision maker so that it can be arranged into an appropriate form. The second phase is concerned primarily with evaluation. The basic idea is that the various alternatives are assessed in terms of an index that is similar to expected utility but with “decision weights” replacing the probabilities and a “value function” replacing the utilities. The decision weights can be represented in terms of a weighting function  $\pi$  on the objective probabilities that are assumed to be accessible to the decision maker in the context of decision making under risk. According to Kahneman and Tversky, “decision weights measure the impact of events on the desirability of prospects, and not merely the perceived likelihood of these events.” [14]. Presumably, according to this view there are a significant number of cases where differences between  $\pi(p)$  and  $p$  indicate a pathology of systematic deviations from the expected utility hypothesis. Through an appeal to empirical and theoretical considerations, Kahneman and Tversky argue that these decision weights satisfy certain structural requirements, e.g., the overweighting of small probabilities. They also provide arguments, both theoretical and empirical, to show that the value function  $v$  of prospect theory, which is defined on “changes in wealth or welfare, rather than final states” satisfies certain structural requirements, e.g., concavity for gains.

We now turn to an example that illustrates the manner in which prospect theory is tested in [13]. Recall from the previous discussion that prospect theory predicts the overweighting of small probabilities, i.e.,  $\pi(p) > p$  for small  $p$ . Kahneman and Tversky perform

the following experiment to test this prediction: Each subject in the study is asked to choose from a pair of alternatives. One of these alternatives is a lottery that pays \$5000 with probability 0.001 and pays \$0 with probability 0.999. The other alternative pays \$5 with certainty. Kahneman and Tversky [14] report that a majority of subjects have a strict preference for the first of the two alternatives just described. Consider a subject who demonstrates these preferences. Such preferences are representable in prospect theory just in case there are  $\pi$  and  $v$  such that  $\pi(.001)v(\$5000) + \pi(.999)v(\$0) > v(\$5)$ . Following Kahneman and Tversky we set  $v(\$0) = 0$  so that the previous inequality simplifies to  $\pi(.001)v(\$5000) > v(\$5)$ , which implies that

$$\pi(.001) > \frac{v(\$5)}{v(\$5000)}. \quad (1)$$

Finally, since  $v$  is assumed to be concave for gains it follows that

$$\frac{v(\$5)}{v(\$5000)} \geq .001. \quad (2)$$

Combining inequalities (1) and (2) yields  $\pi(.001) > .001$  as predicted.

In the experiment discussed in the previous paragraph subjects were presented with a menu of alternatives and a description of the relevant probabilities. Use of this sort of empirical methodology is widespread among work on the psychology of decision making. But to what extent does empirical support, such as that which was just discussed in connection with the overweighting of small probabilities, depend on this methodological choice? One might respond by maintaining that such a question presupposes that there are other plausible methodologies. An important example of an alternative methodology is the “sampling paradigm” that is used in more recent work such as [13]. For our purposes, the essential difference between this alternative methodology and the sort of approach that was taken in [14] is that in the former subjects get their information about the relevant probabilities through sampling rather than by reading a text description.

To illustrate the difference between the two approaches that were just mentioned, consider the following type of experiment from [13]: Divide the subjects into two groups. Subjects in the first group are given the previously discussed task from [13] in connection with the underweighting of small probabilities. That is, subjects in this first group are asked to choose between  $A$ , a lottery that pays \$5000 (\$0) with probability 0.001 (.999), and  $B$ , an alternative that pays \$5 with certainty. Subjects in the second group are asked to choose between pressing one of two buttons,  $A$  and  $B$ , on a computer screen. Although the

subjects in this group are never given such information,  $A$  is a chance setup that rewards either \$5000 or \$0 at the end of each trial and, furthermore, the objective probabilities (i.e., limiting frequencies) that are associated with  $A$  are .001 and .999 for \$5000 and \$0, respectively. Similarly, button  $B$  is a chance setup that rewards \$5 with probability 1. Finally, although subjects in the second group are not told the probabilities associated with  $A$  and  $B$ , they are permitted to sample both buttons as many times as they desire before making their decision between the two alternatives. It should be clear that the crucial distinction between the task that is given to the first group of subjects and the task that is given to the second is essentially the aforementioned distinction between first and second of the two empirical methodologies under consideration.

Let us assume that the subjects in the first group reveal preferences that are consistent with what Kahneman and Tversky observed in connection with the underweighting of small probabilities. Do we expect that the preferences that are revealed in the second group to essentially parallel those that are revealed in the first? Hertwig et al. have argued that we should not. Indeed, Hertwig et al., through experiments of the sort just mentioned, have shown that certain psychological effects – e.g., the overweighting of rare events – are not preserved when one changes from a description-based approach to an experience-based approach, and this lack of preservation is known as *experience-description gap* – as a matter of fact rare events are *underweighted* in experience. The gap is difficult to explain by appealing to theories like Prospect Theory.

Fox and Hadar have recently expressed criticisms in [10] about some of the claims presented in [13] concerning a possible experience-description gap. We will now consider two of the main theoretical criticisms that are discussed in [10]. First, Fox and Hadar do not believe that Hertwig et al. were sufficiently clear about what counts as experience-based decision making [EBDM]: “The generalization that EBDM differs from DBDM is difficult to evaluate because, surprisingly, no one has yet defined ‘experience-based decision making.’” [10]<sup>1</sup> Noting this lack of an adequate definition of experience-based decision making, Fox and Hadar offer what they take to be an adequate characterization of EBDM. The upshot of their analysis is that “[...] EBDM applies to any situation in which there is uncertainty and learning through sampling.” This point, which is significant, will be discussed later in this paper. For now, we turn to a

matter that is more directly related to the Fox and Hadar’s charge that EBDM had not been given an adequate definition.

We think that the analysis of EBDM given in [10] is not well-suited to a study of the experience-description gap as understood in [13]. In particular, the analysis that is supplied in [10] does not say anything about what it means to be an experience-based counterpart to a given description-based task, something which is crucial to the interpretation of the work that is reported in [13]. In light of this, Arló-Costa and Helzner [2] proposed the following analysis of this counterpart relation that is essential if one is to examine how well a given psychological effect travels across experience-description gap:

- In a *decision from description* the subject is presented with a specification of the type of chance mechanism.
- In a *decision from experience* the subject is not presented with such a specification but rather is allowed to observe the behavior of a chance mechanism that has the specified type.

In [2], Arló-Costa and Helzner suggest that this analysis might be useful in examining the extent to which an experience-description gap exists for certain psychological effects associated with decision making under uncertainty. The argument that was given in [2] on behalf of this suggestion is that, while classical descriptions of uncertainty – e.g., the Ellsberg urn – have no experiential counterparts, since the relevant uncertainties in such cases are epistemic, one can specify mechanisms that, at least psychologically, approximate descriptions of uncertainty and, moreover, have an experiential counterpart in the sense of Arló-Costa and Helzner’s analysis of the counterpart relation. In the next section we will examine recent experimental work concerning the way effects associated with these approximations of uncertainty can vary as one moves from EBDM to DBDM. It is worth noting that the present article may be seen as building on the approach considered in [21] and [5]. In a more recent paper Yoram Halevy [12] makes a forceful case for establishing a strong correlation between ambiguity neutrality and the reduction of compound objective lotteries. Halevy concludes that his results suggest that failure to reduce compound (objective) lotteries is the underlying factor of the Ellsberg paradox. We do not want to make such a strong claim but we rely on the idea that a chance setup like  $B^*$ , as described in what follows, can be treated as an operational approximation of uncertainty.

<sup>1</sup>DBDM of course refers to description-based decision making.

## 2 Experimental Work

### Example 1 (Ellsberg’s two-color problem [6])

Consider the following two cases:

**Urn A** contains exactly 100 balls. 50 of these balls are solid black and the remaining 50 are solid white.

**Urn B** contains exactly 100 balls. Each of these balls is either solid black or solid white, although the ratio of black balls to white balls is unknown.

Consider now the following questions: How much would you be willing to pay for a ticket that pays \$25 (\$0) if the next random selection from Urn A results in black (white) ball? Repeat then the same question for Urn B.

It is well known that subjects tend to offer higher maximum buying prices for urn A than for urn B. This seems to be so even in non-comparative cases (see [4] and [3]) contrary to the so-called *comparative ignorance hypothesis* formulated in [9]. On the other hand, consider the following description of a chance setup:

**B\***: First, select an integer between 0 and 100 at random, and let  $n$  be the result of this selection. Second, make a random selection from an urn consisting of exactly 100 balls, where  $n$  of these balls are solid black and  $100 - n$  are solid white.

In a previous ISIPTA paper Arló-Costa and Helzner reported experimental results indicating that maximum buying prices for  $B^*$  are intermediate with respect to the ones for  $A$  and  $B$ . This confirms previous results reported in [21] and [5]. In a more recent paper Yoram Halevy [12] makes a forceful case for establishing a strong correlation between ambiguity neutrality and the reduction of compound objective lotteries (that would lead to treat urns  $A$  and  $B^*$  equally). He therefore concludes that his results suggest that failure to reduce compound (objective) lotteries is the underlying factor of the Ellsberg paradox. We do not want to claim something as strong as that but we rely on the idea that  $B^*$  can be treated as an operational approximation of urn  $B$ . The interest of this move is that  $B^*$  is easily implementable in experience while it is notoriously difficult to find an experiential counterpart of  $B$ . The main experimental finding reported below is that while in description subjects are averse to ambiguity (they prefer  $C$  over  $B^*$  and  $B$ ) in experience this effect is reversed and subjects are ambiguity seeking (they prefer  $B^*$  over  $C - B$  has no experiential counterpart). This shows that the description-experience gap also appears (in a different form) for decisions under uncertainty.

## 3 Method: First Experiment

One hundred and nineteen students at Carnegie Mellon University (Pittsburgh, USA) were presented with the three decision problems presented below. Maximum buying prices for these games were requested. The options  $C$ ,  $B^*$  and  $B$  described above were implemented in the following way:

**C**: A fair chance setup with possible outcomes  $\{1, 2, \dots, 99, 100\}$  has been constructed. If the outcome on the next run of this setup is less than or equal to 50, then you win \$25. Otherwise, you get \$0.

**B\***: Two fair chance setups, I and II, have been constructed. Setup I has possible outcomes  $\{0, 1, \dots, 99, 100\}$ . Setup II has possible outcomes  $\{1, \dots, 99, 100\}$ . The game is played by first running setup I and then running setup II. If the outcome of the run of setup II is less than or equal to the outcome from the run of setup I, then you win \$25. Otherwise, you get \$0.

**B**: An integer  $n$  has been selected from the set  $\{0, 1, \dots, 99, 100\}$ . Nothing is known about the mechanism by which  $n$  has been selected. A fair chance setup with possible outcomes  $\{1, \dots, 99, 100\}$  has been constructed. If the outcome on the next run of this setup is less than or equal to  $n$ , then you win \$25. Otherwise, you get \$0.

### 3.1 The Description Condition

Fifty eight students from the pool of one hundred and nineteen students mentioned above faced the description condition for the first experiment. We have two types of trials. In the first type we consider gains. The subjects face a computer window with two rectangles containing the text used above to describe the options  $C$  and  $B^*$ . The subjects in this condition are asked the following question: **Which one out of the two games will you choose to play?** They then have three possible options for a response:

**Left Button**: You were indifferent between the two alternatives. (A)

**Middle Button**: You had a strict preference for one of the alternatives. (B)

**Right Button**: Neither (A) nor (B) reflect my attitudes

The second type of trials involved losses. For example the loss version of the  $C$ -option is: ‘A fair

chance setup with possible outcomes  $\{1, \dots, 99, 100\}$  has been constructed. If the outcome on the next run of this setup is less than or equal to 50, then you lose 25. Otherwise, you get 0.'

### 3.2 The Experience Condition

Sixty one students from the pool of one hundred and nineteen students mentioned above faced the experience condition for the first experiment. In the experience condition the subjects faced two rectangles containing the labels  $C$  and  $V$ . They can sample these options by clicking on them. Clicking the option  $C$  triggers a random selection from  $\{1, 2, \dots, 99, 100\}$ . If the outcome is less than or equal to 50, then the subject is told that he won \$25. Otherwise, he is told that he got \$0. So, this option corresponds to option  $C$  in description. Clicking the option  $V$  yields an output obtained by triggering the double sampling procedure  $B^*$  presented in description. For example, if one clicks on  $V$  a number is selected at random in the set  $\{0, 1, \dots, 99, 100\}$  and then another from  $\{1, 2, \dots, 99, 100\}$ . If the outcome of the run of second selection is less than or equal to the outcome from the first selection, then the subject is told that he won \$25. Otherwise, he gets \$0. So, button  $V$  is the experiential counterpart of option  $B^*$ . The subjects can sample as much as they want, and then they make a final selection of the  $C$  or  $V$  button. Sampling is used as follows: after the subject presses  $V$ , for example, she has the option of sampling again *the same game* selected at the second stage. The sampling option is also available for the  $C$  button. Of course, in this case one continues to sample the unique game implemented for this button (another random number will be generated from  $\{1, 2, \dots, 99, 100\}$  and if the outcome is less than or equal to 50, then the subject is told that he won \$25. Otherwise, he is told that he got \$0.

As in the description condition the subjects have three buttons at their disposal to respond to. Clicking the middle button, as before, reveals a strict preference for one of the two options. A gain and a loss version of  $C$  and  $B^*$  were implemented.

## 4 Results: First Experiment

At the end of the experiment subjects were asked to provide maximum buying prices for options  $C$ ,  $B$  and  $B^*$  as presented in the description condition. This was done for the subjects in the experience condition and for the subjects in the description condition. So, it makes sense to pool subjects from both conditions in order to compute results. Confirming previous results presented in ISIPTA (see [4] and [3]) there

is a significant difference between maximum buying prices for options  $C$  and  $B$  even when the subjects do not compare vague and clear options. And confirming results reported in [2] option  $B^*$  appears as an intermediate option between options  $C$  and  $B$ . A few remarks are in order before presenting the results from this first experiment: First, although we recognize that doing multiple comparisons to test for independent hypotheses might inflate the Type 1 error, it is important to note that we only compare experience ! with description (a single hypothesis) in different ways (for gain, for loss, and for pooled gain and loss). Thus, we do not consider multiple hypotheses and do not need the Bonferroni correction of  $\alpha = \frac{.05}{3} = 0.02$  (to reduce the inflation in Type 1 error). Second, it is important to note that we are not using a normal approximation on small samples. We tested for normality of the data and we found the data to be non-normal in both experience and description conditions of experiments 1 and 2 using separate Shapiro-Wilk tests. For example, the data were non-normal in both the experience and description conditions (experience:  $D(124) = .64, p < .001$ ; description:  $D(122) = .63, p < .001$ ). Again, the data was non-normal in both experience and description conditions in experiment 2. Therefore, we used non-parametric Mann-Whitney tests to evaluate significant differences between experience and description conditions. In fact, a binomial distribution assumption, as the one anonymous referee suggested, might also not be a correct assumption. Therefore, the safest thing for us to do was to report non-parametric statistics, as we do in the current paper. The  $Z$ -score that we report as part of the statistics belongs to this Mann-Whitney non-parametric test.

The following results were obtained.

Condition	Alternative	Max. Buying Price
Exp. + Desc.	B	4.93 (n = 119)
Exp. + Desc.	$B^*$	6.36 (n = 119)
Exp. + Desc.	C	7.04 (n = 119)

Table 1

Alternative  $C$  is significantly greater than alternative  $B^*$ , with  $T=622, Z=-2.329, p < .05$ , Effect Size = -0.15. Alternative  $C$  is significantly greater than Alternative  $B$ , with  $T=527, Z=-4.751, p < .001$ , Effect Size = -0.31. Alternative  $B^*$  is significantly greater than Alternative  $B$ , with  $T=411, Z=-3.716, p < .001$ , Effect Size = -0.24. So, as we explained above, alternative  $B^*$  can be used as an operational proxy of condition  $B$  in experience.

#### 4.1 The Description condition: Results

The main hypothesis here is that subjects will be ambiguity averse (will significantly prefer  $C$  to  $B^*$ ). The numbers in the tables are the number of subjects clicking the corresponding buttons for each alternative. The boldfaced results for the pooled population show the magnitude of the effect.

Alternative	Left Button	Middle B.	Right B.
$B^*$	10	11	6
$C$	11	18	5

Table 2. Gain Trials (Description)

Alternative	Left Button	Middle B.	Right B.
$B^*$	8	14	5
$C$	7	20	7

Table 3. Loss Trials (Description)

Alternative	Left Button	Middle B.	Right B.
$B^*$	18	<b>25</b>	11
$C$	18	<b>38</b>	12

Table 4. Gain and Loss Trials (Description)

#### 4.2 The Experience condition: Results

Here the hypothesis is the subjects will be ambiguity seekers. The hypothesis is confirmed by the following results.

Alternative	Left Button	Middle B.	Right B.
$B^*$	10	17	4
$C$	14	11	6

Table 5. Gain Trials (Experience)

Alternative	Left Button	Middle B.	Right B.
$B^*$	16	15	2
$C$	10	12	7

Table 6. Loss Trials (Experience)

Alternative	Left Button	Middle B.	Right B.
$B^*$	26	<b>32</b>	6
$C$	24	<b>23</b>	13

Table 7. Gain and Loss Trials (Experience)

One index that seems interesting is based on computing the proportion of subjects who expressed a strict preference for  $C$ , both in Description and Experience. These are the subjects who clicked the Middle Button and  $C$  in experience or the subjects who clicked the Middle Button in description expressing a strict

preference for the ‘clear’ 50-50 lottery. We will refer indistinctly to these subjects as ‘Middle &  $C$ ’ subjects. The analysis reveals the following:

**Gain Trials:** Proportion Middle &  $C$  buttons (Description) ( $18/29 = .62$ ) = Proportion Middle &  $C$  buttons (Experience) ( $11/28 = .4$ ), with  $U=314$ ,  $Z=-1.705$ ,  $p = .09$ , Effect Size =  $-0.23$ .

**Loss Trials:** Proportion Middle &  $C$  buttons (Description) ( $20/34 = .58$ ) = Proportion Middle &  $C$  buttons (Experience) ( $12/27 = .4$ ), with  $U=393$ ,  $Z=-1.108$ ,  $p = .27$ , Effect Size =  $-0.14$ .

**Gain and Loss Trials:** Proportion Middle &  $C$  buttons (Description) ( $38/63 = .6$ ) > Proportion Middle &  $C$  buttons (Experience) ( $23/55 = .4$ ), with  $U=1412$ ,  $Z=-1.998$ ,  $p < .05$ , Effect Size =  $-0.18$ .

It is very interesting to notice that the proportions of Middle &  $C$  subjects in Description remains almost constant for both gain and loss trials (minimum and maximum values are, respectively, .58 and .62). By the same token the proportion of Middle &  $C$  subjects in Experience remains exactly constant with a value of .4. The constancy of the proportions across conditions is clearly depicted in Figure 1 below. Nevertheless the proportion of Middle &  $C$  subjects in Description is not significantly different from the proportion of Middle &  $C$  subjects in Experience for both the gain and loss trials taken separately. But when the two types of trials are pooled the proportion of Middle &  $C$  subjects in Description is indeed significantly greater than the proportion of Middle &  $C$  subjects in Experience.

We believe that the reason why significant results are not obtained for gain or loss trials separately is because we do not have enough subjects in these type of trials taken separately. But it seems that pooling the data for these two type of trials makes sense given that the proportions remain constant across the different types. In other words, the effect seems to have the same polarity in both types of trials.

The analysis of the pooled data reveals that subjects were more ambiguity-averse in the description condition than in the experience condition. To put this in other terms, if we compute the ratio  $R_D$  of the number of Middle &  $C$  subjects in Description divided by the number of Middle &  $B^*$ , and we compute the corresponding ratio  $R_E$  in Experience, we have that  $R_D = \frac{1}{R_E}$ .

We collected additional experience data in a second experiment in order to see whether we can observe significant effects not only for the pooled population

but also for gains and losses. The results are reported in the next section.

## 5 Method: Second Experiment

Thirty students at Carnegie Mellon University (Pittsburgh, USA) participated in a second experiment. They faced an experiential version of the first part of the experiment. Of course in this case we can only implement  $C$  and  $B^*$ .

## 6 Results: Second Experiment

Since we already had data for description we selected at random thirty subjects from the first experiment. First we report the maximum buying prices for the conditions  $C$ ,  $B^*$  and  $B$  in description for the selected subjects of the first experiment and for the experiential version of  $B^*$  and  $C$  in experience. In this experiential version the subjects face the  $V$  and  $C$  buttons used in experience. There is a preparatory phase where they can see results from each button and then maximum buying prices for each alternative are requested.

Condition	Alternative	Max. Buying Price
Description	$C$	5.38 (n = 30; SD = 3.61)
Description	$B^*$	4.71 (n = 30; SD = 3.81)
Description	$B$	3.57 (n = 30; SD = 3.39)
Experience	$C$	6.25 (n = 30; SD = 4.05)
Experience	$B^*$	5.75 (n = 30; SD = 4.75)

Table 8

Although in description the maximum buying prices for  $B^*$  also occupy an intermediate position between prices for  $B$  and  $C$ , the difference between  $C$  and  $B^*$  is not statistically significant ( $T=118$ ,  $Z=-1.401$ ,  $p = .16$ , Effect Size =  $-0.18$ ). But this seems to be due to the fact that the effect verified in the larger population of the first experiment is not verified in this arbitrarily selected sub-population.

There is a clear effect verified in the first experiment and in a previous paper [2] according to which in description the mean maximum buying prices for  $C$  are higher than the mean maximum buying prices for  $B^*$ . Moreover the values for  $B^*$  appear as intermediate between  $C$  and  $B$ . This effect seems to disappear or suffer a complete inversion in experience. This is partly verified by considering mean maximum buying prices. In fact, mean maximum buying prices for  $B^*$  and  $C$  cannot be distinguished statistically in experience ( $T=172$ ,  $Z=-1.037$ ,  $p = .30$ , Effect Size =

$-0.13$ ). A more clear reversal is verified in the following experiments for gains and losses. This manipulation repeats the design used in the first experiment. Rather than providing buying prices the subjects express preferences for the different buttons displayed in their screens.

### 6.1 The Experience condition: Results

The following results were observed for gain and loss trials for experience:

Alternative	Left Button	Middle B.	Right B.
$B^*$	4	11	3
$C$	5	4	3

Table 9. Gain Trials (Experience)

Now here we can see the first clear reversal for experience of the pattern  $B^* < C$  for description. In fact we verify here that  $B^* > C$  is indeed statistically significant in spite of the relatively small size of the population ( $p < .001$ , Effect Size =  $-0.47$ ).

Alternative	Left Button	Middle B.	Right B.
$B^*$	6	12	5
$C$	3	2	2

Table 10. Loss Trials (Experience)

It is interesting to see that the pattern gets repeated here *also for losses* and with a similar ratio. We do have as above  $B^* > C$  is indeed statistically significant ( $p < .01$ , Effect Size =  $-0.71$ ).

Alternative	Left Button	Middle B.	Right B.
$B^*$	10	<b>23</b>	8
$C$	8	<b>6</b>	5

Table 11. Gain and Loss Trials (Experience)

And, of course, we do have the same effect verified for the pooled population.  $B^* > C$  is indeed statistically significant ( $p < .001$ , Effect Size =  $-0.58$ ). Moreover now we have an even nicer result paralleling the one obtained in the first experiment:

**Gain Trials:** Proportion Middle &  $C$  buttons (Description) (**17/24**) > Proportion Middle &  $C$  buttons (Experience) (**4/11**), with  $U=101$ ,  $Z=-2.657$ ,  $p < .01$ , Effect Size =  $-0.45$ .

**Loss Trials:** Proportion Middle &  $C$  buttons (Description) (**14/26**) = Proportion Middle &  $C$  buttons (Experience) (**2/14**), with  $U=110$ ,  $Z=-2.405$ ,  $p < .05$ , Effect Size =  $-0.38$ .

**Gain and Loss Trials:** Proportion Middle & C buttons (Description) ( $31/50 = .62$ ) > Proportion Middle & C buttons (Experience) ( $6/29$ ), with  $U=426$ ,  $Z=-3.524$ ,  $p < .001$ , Effect Size =  $-0.40$ .

So, the second experiment verifies the effect (the fact that subjects are ambiguity seeking in experience) seen in the first experiment for the pooled population. In addition the effect also holds for gains and losses separately and it holds with similar intensity in both conditions.

## 7 Discussion

It is tempting to reason as follows: a play in the chance set up  $B^*$  is equivalent to a play on chance set up  $C$ . The line of reasoning is roughly as follows: The random selection in the first stage of  $B^*$  entails that, for each integer  $i$ , where  $0 \leq i \leq 100$ , there is a probability of  $\frac{1}{101}$  that the urn sampled in the second stage consist of  $i$  black balls and  $100 - i$  white balls. Moreover, according to this line of reasoning the random selection in the second stage entails that if  $i$  is selected in the first stage, then the probability of selecting a black ball in the second stage is  $\frac{i}{100}$ . This line of reasoning then continues by combining the first and second stage probabilities to conclude that the probability of getting a black ball on a trial of  $B^*$  is  $\frac{1}{101} (\sum_{i=0}^{100} \frac{i}{100}) = \frac{1}{2}$ , as in the case of chance set up  $C$ . First note that if this line of reasoning were correct then the results presented in this paper would be rather surprising. Perhaps  $B^*$  and  $C$  can be distinguished in description (due to cognitive limitations of the players), but the two chance set ups should not be distinguishable in experience according to such an argument. However, arguments of the given sort are mistaken as they fail to account for the interaction between the subject's choices and the frequencies that are observed. For example, consider the following set up which is basically equivalent to the one we implemented. Suppose that you have a sequence of 101 possible urns with black and white balls. Each urn contains 100 balls in total but the proportion of white and black changes in each case. The subject could generate output from  $B^*$  by employing a strategy where according to which she samples from the current urn until she sees a white ball and, upon seeing a white ball, advances to the next urn in the arrangement. This strategy should eventually stabilize on the all black urn so that the observed frequencies converge to those associated with the all black urn.

### 7.1 Prospect Theory and its capacity to model the gap experience-description

We explained above that Fox and Hadar offered in [10] an ingenious explanation of the gap experience-description by appealing to a version of Prospect Theory applicable to decisions under uncertainty rather than risk. This version of Prospect Theory is presented in detail in the recent (and excellent) book by Peter Wakker on Prospect Theory [16] (the theory was presented first in [20]). The central idea of the theory is to use event-decision weights rather than probability-based decision weights. In fact if  $P$  is the probability used for risk we can define a function  $W$  on events by applying decision weights to  $P$ . So we have that  $W(E) = w(P(E))$ . Since  $w$  can be non-linear,  $W$  need not be additive. The corresponding function has the properties of a capacity.

The idea that Fox and Hadar considered in the aforementioned paper is to apply the decision weight  $w$  to *judged probabilities* rather than the objective probabilities of the lotteries considered in the case of risk. This ingenious move fits the data reasonably well. So, one can claim that decisions from experience are essentially decisions under uncertainty and one can appeal to Cumulative Prospect Theory to analyze the data. The event-decision weights are calculated by appealing to judged rather than risky probabilities. These judged probabilities are estimated in terms of observed frequencies through sampling.

Is it possible to do something similar in the case we are studying? Perhaps there is a possible strategy one can use to test the predictive power of Cumulative Prospect Theory in this setting. To see the point it is important to stress that we do agree with Hadar and Fox about the fact that decisions from experience are cases of decisions under uncertainty. Let's first see how this applies to our experiment. For each particular play of  $V$  the subject can do some sampling and obtain a judged probability in the sense of Hadar and Fox. So, it seems that one does not have any alternative except representing the subjects playing  $V$  as entertaining a *set* of judged probabilities. The only thing that the subjects know is that there is a chance set-up that is producing a set of probabilities that he can estimate by repeated sampling. But he knows nothing about the nature of the chance set up that produces this set of probabilities. In particular the chance set up that is producing the set of probabilities need not obey the law of large numbers as in the experimental set up used by Stecher et al. [19]. In this case even the computation of winning frequencies would lead to erroneous estimation of the chances of



the chance set up that generates the probabilities.<sup>2</sup>

So, when one implements  $B_S^*$  in experience by fixing a sampling rule various types of indeterminacy arise. First the implemented game does not have a single objective long run frequency associated to it. Subjects can employ different playing strategies that are associated with different long run frequencies. Second, under the point of view of the subject who plays his probabilities are indeterminate also. He only knows that there is a chance set up that produces sets of probabilities that he can eventually estimate. A sophisticated player can learn that the set of probabilities associated with  $C$  are produced by a chance set up of objective probability 0.5.<sup>3</sup> And if the subject has a fixed playing strategy he can perhaps learn the objective probability corresponding to this strategy. But most players will not use a fixed strategy and in this case it seems that there is no learnable objective chance associated to the chance set up that produces the set of probabilities associated with  $B_S^*$ .

So, the probabilities of the subjects remain undetermined. This is so even if we guarantee that the winning frequencies of the two chance set ups converge to the same number (for example by guaranteeing fair sampling procedures for plays of  $B^*$ ), in the short run agents cannot but remain uncertain about the probabilities of the two chance set ups in experience.

Is there a way of connecting this set of priors with event-decision weights? Here is a possible way of doing it. Call the set of priors  $\mathbf{C}$ . Then for each event  $E$  we have the interval  $I_E = \{P(E) : P \in \mathbf{C}\}$ . Now, one can define an event-decision weight  $W$  as  $W = \inf(I_E)$  or  $W = \sup(I_E)$ . Is it possible to approximate out empirical results via this procedure? We propose a careful investigation of this issue for future work.<sup>4</sup>

We can point out here that a theory like Cumulative Prospect Theory will tend to predict asymmetries regarding ambiguity for gains and losses. The majority of the existing evidence seems to indicate, for example, that subjects are ambiguity seeking for losses while they are ambiguity averse for gains (see the evidence and references in [16]). This patterns does not

seem to arise in our experiment. At least in the second experiment it is clear that subjects are equally ambiguity seeking for gains and losses.

The presentation of Cumulative Prospect Theory in [16] makes clear that the main idea of extending Prospect Theory to uncertainty is to avoid the representation of uncertainty via multiple priors. Wakker is quite explicit about rejecting this strategy which he sees as problematic for various reasons that have to do with measurement and elicitation. But it seems that there are experimental situations of the sort we presented in this article where the use of multiple priors seems unavoidable. In spite of this aversion to use multiple priors there might be ways of finding a connection with the way in which Cumulative Prospect Theory represents uncertainty. If this were possible a second step would consist in testing the predictive power of the extended version of Prospect Theory to uncertainty. We also propose to tackle this issue in future work.

## References

- [1] M. Allais. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Americaine. *Econometrica* 21: 503-546, 1953.
- [2] H. Arló-Costa and J. Helzner. Iterated Random Selection as Intermediate Between Risk and Uncertainty. *Proceedings of the Sixth International Symposium on Imprecise Probabilities and their applications*, eds. T. Augustin, F. P. A. Coolen, S. Moral and M. C. M. Troffaes, 2009.
- [3] H. Arló-Costa and J. Helzner. On the Explanatory Value of Indeterminate Probabilities. *Proceedings of the Fifth International Symposium on Imprecise Probabilities and Their Applications*, eds. G. De Cooman, J. Vejnarova and M. Zafalon, 2007.
- [4] H. Arló-Costa and J. Helzner. Comparative Ignorance and the Ellsberg Phenomenon. *Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*, eds. F. Cozman, R. Nau and T. Seidenfeld, 2005
- [5] C.C. Chow and R. Sarin. Known, Unknown and Unknowable Uncertainties. *Theory and Decision*, 52, 127-138, 2002.
- [6] D. Ellsberg. Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, Vol. 75, No. 4, 643-669, 1961.
- [7] D. Ellsberg. *Risk, Ambiguity and Decision*. Garland Publishing, 2001.

---

<sup>2</sup>The main idea in the aforementioned paper is to use a chance set up that cannot be learned in experience. Assuming that fair sampling procedures could be used to play  $B^*$  then its objective chance could be learned by observing enough data. But in the short run the agents cannot but remain uncertain about the determinate or indeterminate chances associated with the chance set up producing the given set of probabilities.

<sup>3</sup>Notice that the subject does not even know that  $C$  is based on single-sampling. For each time they play  $C$  again they do not know whether they are sampling the *same* risky lottery. So, they have to entertain as well a set of probabilities for  $C$ .

<sup>4</sup>See [7] and [8] for earlier discussions along these lines.

- [8] H.J. Einhorn and R.M. Hogarth. Ambiguity and uncertainty in probabilistic inference. *Psychological Review*, 92, 433-461, 1985.
- [9] C.R. Fox and A. Tversky. Ambiguity Aversion and Comparative Ignorance. *The Quarterly Journal of Economics*, Vol. 110, No. 3, 585-603, 1991.
- [10] C. R. Fox and L. Hadar. Decisions from experience = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber and Erev (2004). *Judgment and Decision Making*, 1, 159-161, 2006.
- [11] G. Gigerenzer. *Adaptive thinking : Rationality in the real world*. New York: Oxford University Press, 2000.
- [12] Y. Halevy. Ellsberg revisited. An experimental study. *Econometrica* 75(2): 503-536, 2007.
- [13] R. Hertwig, G. Barron, E.U. Weber and I. Erev. Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, Vol. 15, No. 8, 534-539, 2003.
- [14] D. Kahneman and A. Tversky. Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, Vol. 47, 263-291, 1979.
- [15] Kreps, D. M.: *Notes on the Theory of Choice*, Westview Press, 1988.
- [16] P. Wakker. *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press, 2010.
- [17] L. J. Savage. *The Foundations of Statistics*. Dover Publications; 2 Revised edition, 1972.
- [18] H. A. Simon, H. A. Invariants of human behavior. *Annual Review of Psychology*, 41, 1-19, 1990.
- [19] J. Stecher, T. Shields and J. Dickhaut. Generating Ambiguity in the Laboratory. *Management Science*, Forthcoming.
- [20] A. Tversky and D. Kahneman. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5:297-323, 1992.
- [21] F.J. Yates and L.G. Zukowski. Characterization of Ambiguity in Decision Making. *Behavioral Science*, 21, 19-25, 1976.