**Draft a24, October 3, 2014**

**Ambiguity Attitudes, Framing, and Consistency[[1]](#footnote-1)**

Alex Voorhoeve, Philosophy, Logic and Scientific Method, LSE

Contact: a.e.voorhoeve [at] lse.ac.uk

Ken Binmore, Economics, UCL

Arnaldur Stefansson, Economics, Uppsala University

Lisa Stewart, Psychology, Australian National University

**Abstract**

We use probability-matching variations on Ellsberg’s single-urn experiment to assess both the sensitivity of ambiguity attitudes to framing and the consistency with which subjects display these attitudes within a single frame. Contrary to most other studies, we find very little change in ambiguity attitudes due to a switch from a gain to a loss frame; we also find that making ambiguity easier to recognize has little effect. Regarding consistency, we find that 28% of subjects are highly inconsistent choosers; roughly the same share are highly consistent. Ambiguity seeking is much more frequent among inconsistent choosers; consistent choosers are much more likely to be ambiguity neutral. The prevalence of ambiguity aversion is the same among inconsistent and consistent subjects. These findings support the conclusion in Charness et al (2013) and Stahl (2014) that consistent individuals are likely to be ambiguity neutral. However, unlike those studies, our findings suggest that it is ambiguity-seeking behavior, rather than ambiguity-averse behavior, that is the mark of an inconsistent chooser.

**Introduction**

It is widely thought that experiments on the Ellsberg paradox show that people’s decisions are sensitive to ambiguity. However, some recent experiments suggest that the level of sensitivity is much reduced when care is taken to assure subjects that there is no trickery on the part of the experimenters (Binmore et al. 2012; Charness et al. 2013; Stahl 2014). In this paper, we explore the robustness of this conclusion by altering the framing of one of Elleberg’s experiments in two ways that have been thought relevant to the expression of ambiguity-sensitive choices. Neither change has a significant effect. In particular, framing problems in terms of losses rather than gains does not lead to ambiguity-averse behavior being replaced by ambiguity-seeking behavior. Nor does emphasizing ambiguity lead to a stronger response to its presence.

We also explore the possibility emphasized by Stahl (2014) that cleaner results can be obtained by filtering out the input from “noisy” subjects. Our test is for consistency. How similar are a subject’s choices when she is asked essentially the same question multiple times? We find that the higher the level of consistency we require, the more ambiguity neutral and the less ambiguity seeking our subjects become.

This paper proceeds as follows. Section 1 surveys the literature. Section 2 summarizes our experiments. Section 3 reports the between-subject results on framing. Section 4 reports the results on within-subject consistency and ambiguity attitudes. Section 5 concludes. Additional data and analysis can be found on <http://personal.lse.ac.uk/voorhoev>/ellsberg .

**1. Three questions about ambiguity attitudes**

In Ellsberg’s (1961) famous single-urn experiment, a ball is drawn from an urn containing ten red balls and twenty other balls, of which it is known only that each is either black or white. Subjects are asked to choose between gamble *R*, in which a prize is won if and only if a red ball is drawn,and gamble *B*, in which this prize is won if and only if a black ball is drawn. They are also asked to choose between *B&W,* in which both white and black win, and *R&W*, in which both red and white win. Ellsberg predicted that most people will strictly prefer risky *R* to ambiguous *B* and risky *B&W* to ambiguous *R&W*. However, the first preference implies that the probability of drawing a red ball is greater than the probability of drawing a black ball, and the second preference implies the reverse. Such behavior is therefore inconsistent with standard expected utility theory.

Subjects who choose as Ellsberg hypothesized prefer risky to ambiguous gambles and are therefore said to be ambiguity averse. By contrast, subjects who choose in line with expected utility theory are said to be ambiguity neutral; indeed, if they employ the principle of insufficient reason, they will be indifferent between *R* and *B* and between *B&W* and *R&W.* Finally, subjects who prefer the ambiguous gamble in each case are said to be ambiguity seeking.

The prevalence of these ambiguity attitudes in this and related experimental designs has been extensively studied. Recently, there has been an upsurge of interest in the sensitivity of ambiguity attitudes to framing and in the consistency with which subjects display these attitudes within a given frame. We single out three sets of findings:

*(i) Gain/loss differences:* For moderate to high probability gains, the dominant finding is ambiguity aversion (Trautmann and van der Kuijlen 2014; Chew et al. 2013; and Wakker 2010). By contrast, for moderate to high probability losses, only a minority find ambiguity aversion.[[2]](#footnote-2) Some studies find ambiguity neutrality;[[3]](#footnote-3) others report a predominance of ambiguity seeking;[[4]](#footnote-4) some find a transition from approximate neutrality to ambiguity seeking as the probability of loss increases.[[5]](#footnote-5) A recent review concludes that “there is clear evidence for an effect of the outcome domain on ambiguity attitude” (Trautmann and van der Kuijlen 2014, p. 22; cf. Wakker 2010, p. 354).[[6]](#footnote-6)

*(ii) Difficulty in recognizing ambiguity.* Several studies of moderately likely gains suggest that some subjects find it difficult to grasp the nature of ambiguous alternatives. In previous experiments, we found a very modest degree of ambiguity aversion in a single-urn setup, in part because fewer subjects proved averse to ambiguity in the two-winning-color choice than in the one-winning-color choice (Binmore et al. 2012).We conjectured that this is due to the fact that some subjects find it difficult to grasp that in the two-winning-color variants, *B&W* is the unambiguous option, and *R&W* is the ambiguous option. Chew et al. (2013) support this conjecture by pointing out that ambiguity aversion is much less prevalent in experiments based on Ellsberg’s single-urn set-up than in experiments based on Ellsberg’s two-urn set-up, in which subjects must choose between drawing from a risky urn and drawing from an ambiguous urn. They also provide evidence that few subjects are able to discern ambiguity in more complex cases and those who do are generally ambiguity averse in a gain frame.

*(iii)* *Individual consistency of choice and ambiguity attitude.* Several recent studies for moderate-probability gains suggest that a substantial share of subjects make inconsistent or essentially random choices.[[7]](#footnote-7) This raises the question about the relationship between subjects’ consistency and their ambiguity attitude. This question has not been much studied, and the results are conflicting. Two recent studies suggest that ambiguity neutrality is predominant among consistent choosers. If one excludes subjects labeled “ambiguity-incoherent” from the sample in Charness et al. (2013), then 15% are ambiguity seeking, 75% are ambiguity neutral, and 10% are ambiguity averse. Similarly, if one excludes subjects classified as “random choosers” from Stahl (2014), then 74% are ambiguity neutral and 25% are ambiguity averse (his experiment does not allow for ambiguity seeking).[[8]](#footnote-8) By contrast, Chew et al. (2013) find a far higher rate of ambiguity aversion among “high comprehension” subjects (69%) than among “low comprehension” or “inattentive” subjects (48%).

Prompted by these findings, this paper addresses the following three questions:

(i) Is a substantial shift observable from ambiguity-averse towards ambiguity-neutral and/or ambiguity-seeking behavior when one changes from a gain to a loss frame?

(ii) Does clarifying which alternative is ambiguous increase the response to ambiguity, i.e., generate more ambiguity aversion for gains and more ambiguity neutrality and/or ambiguity seeking for losses?

(iii) How consistent are individual choices within a given frame, and how does consistency correlate with ambiguity attitude?

We address questions (i) and (ii) in a between-subject design. We address question (iii) by examining individual-level data from these experiments.

**2. Experiments**

Our experiments use decks of cards rather than urns and employ a titration to estimate the value *r1* of the probability *r* of drawing a red cardthat makes a particular subject indifferent between *R* and *B* and the value *r2* of drawing red that makes the same subject indifferent between *B&W* and *R&W*.[[9]](#footnote-9) The aim of this titration is to locate the values of *r1* and *r2* within eight subintervals of [0*,* ] using the scheme illustrated in Fig. 1.

Our titration locates an estimate of a subject’s (*r*1*, r*2)within one of 64 squares of a chessboard. Figure 2 shows the regions on this chessboard that we assign to the decision principles and ambiguity attitudes outlined below for the reasons given in Binmore et al. (2012).

*R B*

when *r* =?

*R B*

when *r* =?

*R B*

when *r* =?

*R B*

when *r* =?

*R B*

when *r* =?

yes

*R B*

when *r* =?

*R B*

when *r* =?

yes

yes

yes

yes

yes

no

no

no

no

no

no

no

yes

**Fig. 1** The tree shows the questions implicitly asked about the subjects’ preferences when the probability of red is *r* in order to locate *r1* in one of eight subintervals of [0*,* ½]. For example, someone who answers *BRR* is assigned a value of *r*1 satisfying . The same tree is used to locate *r*2, except that *B&W* replaces *B* and *R&W* replaces *R.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  | stp |  |  |  |  |  |  |  |  | waa |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  | saa |  |  |  | was |  |
|  |  |  |  | PIR |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  | waa |  |  |  | sas |  |  |
|  |  | stp |  |  |  |  |  |  |  |  |  |  | was |  |  |  |  |  |
| 0 |  |  |  |  |  |  |  |  |  | 0 |  |  |  |  |  |  |  |  |
|  |  |  |  | |  |  |  |  |  |  |  |  | |  |  |  |

**Fig. 2** Regions corresponding to ambiguity attitudes.

**Ambiguity neutrality.** We allow for two kinds of ambiguity-neutral behavior.

*Sure-thing principle* (stp): A subject who honors Savage’s (1954, p. 21) sure-thing principle has *r*1 = *r*2*.*

*Laplace’s Principle of Insufficient Reason* (PIR):This principle says that two events should be assigned the same subjective probability if no reason can be given for regarding one event as more likely than another. We then have *r*1 = *r*2 = . Since is an endpoint of two of our intervals (see Fig. 1), any value of (*r1, r2*)that lies in one of the four squares of our chessboard with a corner at is counted as being consistent with PIR.

**Ambiguity sensitivity.** We allow for two kinds of ambiguity sensitivity.

*Weak ambiguity sensitivity:* We say that *r*1 *< r*2 indicates weak ambiguity aversion (waa), because it implies that *R* is preferred to *B* and *B&W* to *R&W* when the probabilityof red being drawn lies between *r*1 and *r*2.We similarly say that *r*1 *> r*2 indicates weak ambiguity seeking (was).

*Strong ambiguity sensitivity:* We treat pairs (*r*1*, r*2)with *r*1 *<*  and *r*2 *>* as cases of strong ambiguity aversion (saa). We treat pairs (*r*1*, r*2)with *r*1 > and *r*2 <as cases of strong ambiguity seeking (sas).

**Common properties of the experiments.** Each subject was seated in front of a private computer screen. An on-screen introduction explained the structure of the experiment and the nature of the choices subjects would face. Subjects were told that they would choose between versions of *R* and *B* and between versions of *B&W* and *R&W* in ignorance of the mixture of black and white cards. Special care was taken in the introduction to illustrate the nature of this ignorance. Subjects were shown an illustrative deck of 6 red cards and 15 black or white cards, the latter marked on the screen with a “?” on the back. They were then told that the “?” cards could be any mixture of black and white cards, with three subsequent screens inviting them to move the mouse over the “?” cards, revealing three illustrative mixtures.

We sought to allay suspicion of skullduggery by having the decks from which the bets would be constructed visible at the start in closed boxes of cards and by making transparent the preparation of the gambles from these decks and the drawing of the winning card. After making two practice choices, subjects were invited to the front of the room to witness one of the practice bets being constructed and played. The experimenter opened a box of red cards and a box of black or white cards, and counted out the number of red and black or white cards in the first practice choice (respectively 6 and 15). These were placed in a card-shuffling machine to randomize the order. Finally, a subject exposed the card with the winning color. Subjects were informed that of the subsequent twenty-four choices they faced in the real experiment, two bets would be randomly selected by the computer to be played for money in this manner at the end of the experiment. We attach some importance to this feature of our “small word” framing (in Savage’s (1954) sense), insofar as it eliminates one possible source of uncontrolled ambiguity.

Each experiment consisted of four rounds. In Round 1, the subjects were first navigated through the tree of Fig. 1 to estimate the interval in which to locate the value *r*1 that makes a subject indifferent between *R* and *B*. Subjects were then navigated through a similar tree to estimate the value *r*2 that makes a subject indifferent between *B&W* and *R&W*. Subsequent rounds repeated this process with black replaced by blue, yellow, and green, thus yielding four estimates of (*r*1*, r*2) for each subject.

**Particular frames.** In Binmore et al. (2012), we reported the effects of variations in the presentation of alternatives in Versions 1-3 of this experiment. These experiments involved gains only, and left subjects to infer from the information provided on the number (and percentage share) of red, white, and black cards which option was ambiguous and which was merely risky. In our new experiments, we took as our starting point Version 3 and revised it to study the effects of switching to a loss frame and of making explicit any ambiguity present in an alternative.

Version 4 represented a loss frame in which ambiguity remained implicit. Upon entering, subjects were given twenty-five plastic £1 coins (£25 ≈ US$ 40). They were advised that every coin that they avoided losing at the end of the experiment would be worth £1 in real money. They were told that they could lose no more than twenty of these coins. For each choice, they were informed that they would keep their coins if they won and that they would lose a number of their coins otherwise.

Version 5 was identical to Version 4, except that throughout, subjects were told which alternative was ambiguous and how ambiguous it was. For example, in a choice between *B&W* and *R&W*, subjects were told that if they chose to bet on *B&W*, then their chance of winning was 67%, while if they chose to bet on *R&W*, then their chance of winning was anywhere from 33% to 100%.

Version 6 was identical to Version 3, except that throughout, subjects were told which alternative was ambiguous and how ambiguous it was. Appendix A provides sample screens for all versions. Table 1 summarizes these experiments.

|  |  |  |
| --- | --- | --- |
| **Framing** | **Gain** | **Loss** |
| **Implicit** (Subjects must infer ambiguous alternative) | Version 3 | Version 4 |
| **Explicit** (Subjects are told which alternative is ambiguous) | Version 6 | Version 5 |

**Table 1. Four experiments.**

Version 3 was performed at the ELSE lab at University College London; Versions 4, 5, and 6 at the Behavioural Science Lab at the London School of Economics. These labs recruit jointly and their demographic mix is similar.

**3. Results on framing**

As in Binmore et al. (2012), there was no evidence that subjects adjusted their choices during the experiment: the final round data is not significantly different from earlier rounds. We therefore aggregate the data across all rounds in each version of the experiment. The results for each version are given as percentages of the total number of observations in Figure 3. The results sorted by ambiguity attitudes appear in Table 2.

Before testing these hypotheses, a comment on the broad pattern of ambiguity attitudes displayed in our experiments may be helpful. Across all versions, roughly 30% are strongly ambiguity averse and roughly 45% are weakly ambiguity averse. Ambiguity neutrality is common, but more variable: the use of the principle of insufficient reason, for example, ranges from 24% to 38%. Ambiguity seeking is least common in all versions. Weak ambiguity seeking is quite constant around 21%, but strong ambiguity seeking shows more variability, varying from 7% to 16%. Only ambiguity neutrality is significantly more prevalent than would occur if subjects were to choose at random.

The hypotheses under consideration regarding framing are:

(i) For losses there will be less ambiguity aversion than for gains, and substantially more ambiguity-neutral or ambiguity-seeking behavior.

(ii) Making the ambiguous option clearer will magnify the response to ambiguity; it will generate more ambiguity-averse behavior for gains and more ambiguity-neutral and/or ambiguity-seeking behavior for losses.

We employ three complementary tests: (Section 3.1) a test at the level of each attitude taken separately; (Section 3.2) a version of the Kolmogorov-Smirnov test applied to the distribution across all 64 squares; and (Section 3.3) a model of the data.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Version 3** (gains; ambiguity inferred) | | | | | | | |  |  | **Version 4** (losses; ambiguity inferred) | | | | | | | |
|  | 80 subjects, percentages, 4 observations each | | | | | | | |  |  | 90 subjects, percentages, 4 observations each | | | | | | | |
|  | 1 | 1 | 2 | 2 |  | 1 | 1 |  |  | 1 | 2 | 1 |  | 1 | 1 | 2 |
|  |  | 1 | 2 | 3 | 1 |  | 1 |  |  |  |  | 1 | 1 | 1 |  | 1 | 1 |  |
|  |  |  | 3 | 5 | 1 | 2 | 1 | 1 |  |  |  |  | 3 | 6 | 6 | 4 |  |  |
|  |  | 1 | 4 | 6 | 6 | 1 | 1 |  |  |  |  |  | 2 | 11 | 5 | 3 | 1 |  |
|  | 1 | 7 | 17 | 3 | 1 |  |  |  |  | 1 | 6 | 16 | 3 | 1 | 1 | 1 |
|  |  | 1 | 7 | 5 | 1 |  |  |  |  |  |  | 1 | 5 | 3 |  | 1 |  |  |
|  |  | 2# | 2 | 1 |  |  |  |  |  |  |  | 3 | 2 | 2 |  |  |  |  |
| 0 | 1 |  | 1 |  |  |  |  |  |  | 0 |  | 1 |  |  |  |  | 1 |  |
|  |  |  |  | |  |  |  |  |  |  |  |  | |  |  |  |
|  | **Version 6** (gains; ambiguity explicit) | | | | | | | |  |  | **Version 5** (losses; ambiguity explicit) | | | | | | | |
|  | 81 subjects, percentages, 4 observations each | | | | | | | |  |  | 73 subjects, percentages, 4 observations each | | | | | | | |
|  |  |  | 1 | 1 |  |  |  |  |  | 1 | 2 | 2 | 2 | 1 |  | 1 |
|  |  |  | 1 | 1 |  | 1 | 1 |  |  |  |  | 1 |  | 2 | 2 |  |  |  |
|  |  | 1 | 5 | 6 | 3 | 7 |  |  |  |  |  | 1 | 5 | 7 | 5 | 6 | 1 |  |
|  |  |  | 2 | 14 | 10 | 2 |  |  |  |  |  |  | 3 | 7 | 3 | 2 |  |  |
| 1 | 2 | 4 | 12 | 2 | 2 | 1 |  |  |  |  | 6 | 11 | 3 | 4 |  |  |
|  |  |  | 4 | 3 | 2 | 2 |  |  |  |  |  | 1 | 4 | 2 | 1 | 4 | 1 |  |
|  |  | 2 | 1 |  | 1 | 1 |  |  |  |  |  |  | 1 | 1 |  | 1 |  |  |
| 0 |  |  |  |  |  |  |  |  |  | 0 |  |  |  | 1 |  |  |  | 1 |
|  |  |  |  | |  |  |  |  |  |  |  |  | |  |  |  |

**Fig. 3: Aggregate results, framing**. Shaded squares indicate a particularly high concentration of responses.

# This entry was omitted in error in Binmore et al. (2012, p. 226). (The analysis in that paper was done with the correct data.)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **0** |  | **3** | **4** | **5** | **6** | **Sign.**  **Diff. a** |  | **Gain**  **(3&6)** | **Loss**  **(4&5)** | | **Sign. Diff. b** |  | **Impl.**  **(3&4)** | **Expl.**  **(5&6)** | **Sign.**  **Diff. c** |
|  | Pop | 100 |  | 100 | 100 | 100 | 100 |  |  | 100 | | 100 |  |  | 100 | 100 |  |
|  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |
| **Averse** | saa | 25 | 28 | 27 | 31 | 31 |  |  | 30 | | 29 | 0.838 |  | 28 | 31 | 0.227 |
| waa | 44 | 43 | 44 | 49 | 44 |  |  | 43 | | 46 | 0.330 |  | 43 | 46 | 0.327 |
|  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |
| **Neutral** | stp | 13 | 37 | 37 | 26 | 36 |  |  | 36 | | 32 | 0.093\* |  | 37 | 31 | 0.025\*\* |
| PIR | 6 | 32 | 36 | 24 | 38 | 0.092\* (4&5) 0.055\*  (5&6) |  | 35 | | 30 | 0.070\* |  | 34 | 31 | 0.258 |
|  |  |  |  |  |  |  |  |  |  | |  |  |  |  |  |  |
| **Seeking** | sas | 25 | 7 | 8 | 16 | 12 | 0.084\*  (3&5) |  | 9 | | 12 | 0.201 |  | 8 | 14 | 0.000\*\*\* |
| was | 44 | 20 | 19 | 25 | 21 |  |  | 20 | | 22 | 0.443 |  | 20 | 23 | 0.161 |

= deviation towards the attitude compared to the null hypothesis (column 0) that subjects choose at random is significant at the 1% level using one sided z-tests of goodness-of-fit without continuity correction. (No attitude is significant at only the 10% or 5% level.)

\* = difference is significant at the 10% level.

\*\* = difference is significant at the 5% level.

\*\*\* = difference is significant at the 1% level.

**Table 2. Percentage shares of ambiguity attitudes in different versions.[[10]](#footnote-10)** The table reports the results of several tests. **Shaded** cells indicate that there is a deviation towards the ambiguity attitude compared to the null hypothesis (column 0) that subjects choose at random. **Sign. Diff. a** reports a pairwise comparison of the prevalence of the ambiguity attitude in question in Version *i* with Version *j*, for all *i*, *j*, under the null hypothesis that the prevalence of this attitude is independent of the version. We report only those *p*-values for comparisons for which we can reject the null hypothesis at the 10% level (or lower). **Sign. Diff. b** reports the *p-*values for the null hypothesis that the prevalence of the ambiguity attitude in question is independent of whether choices take place in a gain or in a loss frame. **Sign. Diff. c** reports the *p-*values for the null hypothesis that the prevalence of the ambiguity attitude in question is independent of whether ambiguity is implicit or explicit.

**3.1 Does prevalence of attitudes differ significantly across versions?**

Our first test of hypothesis (i) has several parts.

(a) We do a pairwise comparison of the prevalence of each ambiguity attitude in Version *i* with Version *j*, for all *i*, *j*, under the null hypothesis that that the averages of the four observations from each subject are independent, and fall into a given ambiguity category with the same probability for each of  Versions 3-6. The difference of the percentages of these averages in any two versions is then approximately normal with zero mean and a variance that depends on the number of subjects in each version, which allows us to calculate the probability (*p*-value) of obtaining the observed absolute difference or more in the prevalence of the relevant attitude between Version *i* and Version *j* if the null hypothesis were true.

(b) For each ambiguity attitude, we compare its prevalence in the aggregate of the gain versions with its prevalence the aggregate of the loss versions. Our null hypothesis that the prevalence of an attitude is independent of gain or loss framing.

(c) We compare the prevalence of each attitude in the aggregate of the “ambiguity implicit” versions with its prevalence the aggregate of the “ambiguity explicit” versions under the null hypothesis that the prevalence of an attitude is independent of this change in frame.

Table 2 reports our results (for a description of our test and all values of this test, see: our online appendix on <http://personal.lse.ac.uk/voorhoev> ).

**Gains versus losses:** The mere switch from gains to losses while leaving subjects to infer the ambiguous option in every choice (Version 3 to Version 4) makes no significant difference for any ambiguity attitude. The mere switch from gains to losses while making the ambiguous option explicit (Version 6 to Version 5) equally has no significant effect, except a decrease in one type of ambiguity-neutral behavior, PIR, which is significant only at the 10% level. Comparing the aggregate of both gain versions with the aggregate of both loss versions, we find a similar effect: a modest decrease in ambiguity-neutral behavior (stp and PIR), which is significant only at the 10% level. In sum, on this test, there is no evidence within our small-world framing for the common finding that changing from gains to losses leads to a reduction in ambiguity aversion and an increase in ambiguity-neutral and/or ambiguity-seeking behavior.

**Making ambiguity easier to recognize:** Making the ambiguous option explicit in every choice has no significant effect in a gain frame (Version 3 versus Version 6). In a loss frame (Version 4 versus Version 5), there is some increase in the predicted direction of ambiguity seeking, but it is not significant. Indeed, in this frame only a decrease in the use of the PIR is significant (at the 10% level). This test therefore finds no support for hypothesis (ii).

We also compared the aggregate of versions that leave ambiguity implicit (Versions 3 & 4) with the aggregate of versions that make it explicit (Versions 5 & 6). Each of these groupings combines a gain and a loss frame. Since making ambiguity explicit was meant to magnify the difference in ambiguity attitudes between gains and losses, our hypotheses make no prediction regarding this comparison. The “explicit” aggregate has a modestly lower share of behavior in accordance with the sure-thing principle (significant at the 5% level) and a higher share of strong ambiguity seeking (significant at the 1% level). These effects are primarily driven by Version 5. We therefore interpret this result as suggesting that combining our two reframings leads to the greatest differences with the standard framing of the Ellsberg single-urn experiment.

Overall, neither of our two hypotheses regarding framing gains support from a direct comparison of the prevalence of each ambiguity attitude taken separately.

**3.2 Kolmogorov-Smirnov test**

Next, we compare the distributions of each version across our 64-square grid using the Kolmogorov-Smirnov (K-S) test. This provides a criterion for deciding whether two samples are generated by the same probability distribution. It is important that the K–S test is non-parametric, because its use shows that some of our data is not normally distributed, which rules out various alternative approaches. With one-dimensional data, the K–S statistic is obtained by computing the cumulative distribution functions of the two samples to be compared. Its value is the maximum of the absolute difference between them. Low values indicate that the evidence is not good enough to reject the null hypothesis that the two samples are from the same distribution.

Lopes et al. (2007) review the problem of applying the K–S test with multidimensional data. The problem arises because the manner in which the data points are ordered then becomes significant. Their very severe recommendation is to maximize over all possible orderings of the data points. Such a procedure might be appropriate when the data is unstructured, but we exploit the underlying structure of our problem by applying the one-dimensional K–S test separately to the sums of the columns, the sums of the rows, and the sums of both types of diagonal in each of the 8×8 matrices of Fig. 2. Table 3 reports the largest of these four K-S statistics. We reject the hypothesis that two versions *i* and *j* come from the same distribution if this exceeds the relevant threshold (which depends on the sample sizes). The results are as follows.

**Gains versus losses:** The mere switch from gains to losses while leaving subjects to infer the ambiguous option in every choice (Version 3 to Version 4) makes no significant difference. The mere switch from gains to losses while making the ambiguous option explicit (Version 6 to Version 5) equally has no significant effect. Nor does the aggregate of the gain versions differ significantly from the aggregate of the loss versions.

**Making ambiguity easier to recognize:** Making the ambiguous option explicit in every choice has no significant effect in a gain frame (Version 3 versus Version 6) or a loss frame (Version 4 versus Version 5). Nor does the K-S test reveal any significant difference between the “ambiguity implicit” and “ambiguity explicit” groupings.

In sum, the K-S test does not permit us to say with confidence of any version that it generates significantly different behavior from any other version. Just as the analysis of Section 3.1, it therefore offers no support for either of our two hypotheses.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 4 | 5 | 6 |  |  | Loss  (4&5) | Implicit  (3&4) |
| 3 | 0.09 | 0.17 | 0.15 |  | Gain  (3&6) | 0.08 |  |
| 4 |  | 0.11 | 0.06 |  | Explicit |  | 0.11 |
| 5 |  |  | 0.09 |  |  |  |  |

**Table 3. Significantly different distributions?** For each comparison between versions, we compute four K-S statistics comparing, respectively: (i) the distribution of *r1*; (ii) the distribution of *r2*; (iii) the distribution of the NE-SW diagonal sums (an indicator of ambiguity attitudes, since NE is ambiguity aversion and SW is ambiguity seeking); and (iv) the distribution of the SE-NW diagonal sums. We report the largest of these. None of these are sufficiently large to confidently reject the hypothesis that the distributions are drawn from the same underlying distribution. (The borderline for rejecting the hypothesis that two distributions are drawn from the same distribution at the 10% level varies with population size. For comparisons of individual versions it is ≈ 0.19. For the groupings, it is ≈ 0.14.)

**3.3 Modeling the data**

Our third test involves fitting the simple econometric model from Binmore et al. (2012) to each version. The model assumes that subjects basically follow the sure-thing principle, but sometimes diverge from it when answering the questions in our titration. More precisely, the model assumes that the baseline preferences of all subjects satisfy *r*1 = *r*2, and that *r*1 is normally distributed with mean *μ* and standard deviation *σ*. When an answer consistent with the baseline preference is in the direction of ambiguity-averse behavior, the model assumes that subjects respond in line with this preference with probability *a*, where *a <* 1. When an answer consistent with the baseline preference is in the direction of ambiguity-seeking behavior, we assume that subjects respond in line with the baseline preference with probability *d*, where *d <* 1. We therefore have a model with four parameters: *μ*, *σ*, *a*, *d*. For *a* > *d*, subjects are more likely to diverge from ambiguity neutrality in an ambiguity-averse direction; for *a* < *d*, subjects are more likely to diverge in an ambiguity-seeking direction. We find the best-fitting set of parameters for each version separately, and for the aggregate of the gain versions and the aggregate of the loss versions. If each instantiation of the model fits tolerably well, a test of the effects of framing is this: Does the best-fitting model for version *i* have parameters that are substantially different from the best-fitting model for version *j*?

To address this question, we compute two K–S statistics, *KI*  and *KII*, using the observed data on the main diagonal for *KI*, and the sums of data points along parallels to the main diagonal for *KII*. Low values of *KI* indicate that the observed data points on the main diagonal of our data matrix are consistent with our model. Low values of *KII* indicate that deviations from ambiguity neutrality are consistent with the model. Table 4 lists the results of a hill-climbing exercise in parameter space.

Several things emerge from this analysis. First, the best-fitting instantiations of this model fit well. These instantiations of the model therefore cannot confidently be rejected as representations of the data. (The parameter range within which this result can be sustained for each version is small.) Second, our indicator *a – d* for the relative prevalence of ambiguity aversion compared to ambiguity seekingchanges little in response to changes in framing. It indicates that ambiguity aversion is more prevalent than ambiguity seeking throughout, although there is a slight shift towards ambiguity seeking in moving between the gain and loss frame (*a – d* = 0.12 for the aggregate of the gain-frame experiments and *a – d* = 0.10 for the loss-frame experiments). Our model therefore provides no evidence for a substantial shift in the balance of ambiguity aversion or ambiguity seeking between the various versions, or between aggregated gain versions on the one side and aggregated loss versions on the other.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model parameters** | **3** | | **4** | | **5** | | **6** | |  | | **Gain**  **(3&6)** | | **Loss**  **(4&5)** | |
| 1-*a* (likelihood of diverging from stp in an ambiguity-seeking direction in any choice) | 0.09 | | 0.11 | | 0.10 | | 0.12 | |  | | 0.11 | | 0.11 | |
| 1-*d* (likelihood of diverging from stp in an ambiguity-averse direction in any choice) | 0.22 | | 0.20 | | 0.20 | | 0.21 | |  | | 0.23 | | 0.21 | |
| *a – d* (indicator of the prevalence of ambiguity aversion over seeking) | 0.13 | | 0.09 | | 0.10 | | 0.09 | |  | | 0.12 | | 0.10 | |
| *μ* (mean of the distribution along the SW-NE diagonal) | 0.31 | | 0.32 | | 0.33 | | 0.33 | |  | | 0.32 | | 0.32 | |
| *σ* (standard deviation) | 0.035 | | 0.045 | | 0.060 | | 0.056 | |  | | 0.045 | | 0.051 | |
| **Fit** |  |  | |  | |  | |  | |  | |  | |
| *KI* | 0.06 | | 0.08 | | 0.10 | | 0.05 | |  | | 0.05 | | 0.09 | |
| *KII* | 0.05 | | 0.03 | | 0.06 | | 0.02 | |  | | 0.02 | | 0.04 | |

**Table 4. The best-fitting models**. Each of these instantiations of the model passes our K-S tests for the version(s) to which it is tailored. For *KI* and *KII,* lower numbers indicate a better fit; for individual versions, the lower limit for the (stringent) 10% confidence level is ≈ 0.13; for combined versions, it is ≈ 0.10 because of a larger population size.

**3.4 Conclusions with regard to framing**

Our three complementary tests of our hypotheses with regard to framing yield similar conclusions.

*Hypothesis (i) There will be substantially less ambiguity-averse behavior and substantially more ambiguity-neutral and/or ambiguity-seeking behavior for losses than for gains.*

In our experiments, the prevalence of ambiguity-averse behavior does not depend on framing. Comparing the aggregates of our gain and loss frames, the only significant finding (at the 10% level) is an unpredicted *decrease* in ambiguity-neutral behavior. We conclude that our experiments refute hypothesis (i).

*Hypothesis (ii): Making the ambiguous option clearer will generate more ambiguity aversion for gains and more ambiguity neutrality and/or seeking for losses.*

In a gain frame, clarifying ambiguity has no effect. This refutes hypothesis (ii) for gains. In a loss frame, while there is a modest increase in ambiguity seeking, this is not significant. Moreover, there is also a (marginally significant) shift *away* from ambiguity-neutral behavior. Therefore, we find no significant shift towards ambiguity neutrality and/or ambiguity seeking, so that our experiments also refute hypothesis (ii) for losses.

**4. Results on consistency**

We now turn to the question how consistent individual choices are within a given frame and how consistency correlates with ambiguity attitude. To address this question, we look at individual-level data. As mentioned, each experiment yields four estimates for a subject’s (*r1*, *r2*). We measure how inconsistent a subject *n* is by finding the minimal number *Mn* of mistakes in the titrations of the experiment that a subject would have to make from some “representative pair” to end up with his or her four values of (*r1*, *r2*). If *Mn* ≤ 12, one can think of the representative pair as the maximum likelihood estimator of an agent who has a “true pair” (*r1*, *r2*) but who makes mistakes with probability *p* < when answering titration questions. More precisely, we define the “distance” from square (*k*, *l*) to square (*i*, *j*) by Given our four data points (*i1*, *j1*), (*i2*, *j2*), (*i3*, *j3*), (*i4*, *j4*) for individual *n*:

for each (*k*, *l*). The measure *Mn* of consistency for each individual is then

taken over all (*k*, *l*).[[11]](#footnote-11)

Table 5 yields this measure for several versions of our experiment. In order to assess the relationship between our measure of inconsistency and ambiguity attitudes, it will be useful to aggregate across different versions of our experiment where our data permits. We therefore employ the K-S test outlined in Section 3.2 for all six versions of our experiment (Versions 1-3 reported in Binmore et al. 2012 and Versions 4-6 reported here) as follows:

1. We exclude a version from our aggregated data if the largest of the four K-S statistics indicates that it is significantly different at the 10% level from any other version. This excludes Version 1. No other version is excluded by this criterion.

2. For the versions that pass test (1), we assess whether it is permissible to add various salient combinations of versions to other such combinations. E.g. we compare “all gain versions” (Version 2, 3, and 6) to “all loss versions” (4 and 5). No version or combination of versions is excluded by this test (see Appendix B).

The upshot is that this procedure permits us to aggregate Version 2 through 6 of our experiment. However, since we saw in Section 3.1 that Version 5 is most different from our other versions, we confirmed that our results regarding the relationship between consistency and ambiguity attitude are robust to removing Version 5 from our data.

Next, we classify subjects into three groups: high consistency (0 ≤ *Mn* ≤ 2); medium consistency (3 ≤ *Mn* ≤ 5); and low consistency (6 ≤ *Mn*). We employ this classification because it both brings together subjects who are plausibly similar in terms of consistency of choice and creates groupings that are sufficiently similar in size to ensure statistical power. On this classification, 29% of our subjects are highly consistent; 44% are moderately consistent; and 28% have a low level of consistency. Our findings are broadly in line with the incipient literature on consistency of choice and ambiguity, which reports that between 20% and 60% of subjects are inconsistent (using different measures).[[12]](#footnote-12)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Version** | | | | |  | **Consistency** | |
| ***Mn*** | **2** | **3** | **4** | **5** | **6** | **Aggregate** |  | |
| 0 | 9 | 3 | 16 | 6 | 11 | 9 |  | High: 29%  (*N* = 113) |
| 1 | 4 | 15 | 9 | 6 | 20 | 11 |
| 2 | 12 | 9 | 9 | 8 | 5 | 9 |
| 3 | 17 | 15 | 19 | 10 | 11 | 15 |  | Moderate: 44%  (*N* = 177) |
| 4 | 13 | 16 | 18 | 26 | 10 | 16 |
| 5 | 21 | 11 | 9 | 8 | 17 | 13 |
| 6 | 15 | 8 | 8 | 11 | 7 | 10 |  | Low: 28%  (*N =* 110) |
| 7 | 4 | 6 | 1 | 6 | 6 | 4 |
| 8 | 3 | 8 | 2 | 8 | 4 | 5 |
| 9 | 0 | 4 | 3 | 4 | 4 | 3 |
| ≥10 | 3 | 6 | 7 | 8 | 5 | 6 |
| *Median* | 4 | 4 | 3 | 4 | 4 | 4 |  | |
| *N* | 76 | 80 | 90 | 73 | 81 | 400 |  | |

**Table 5. Consistency.** *Mn* is the aggregate “distance” in our 8 8 grid from a subject’s four answers to an answer that best represents them. “0” occurs when a subject answers all questions identically four times.

So defined, how does consistency relate to ambiguity attitude? Table 6 gives our results. Three things are immediately apparent. First, the prevalence of ambiguity aversion is nearly identical in all three groups. Second, ambiguity-neutral behavior increases as subjects become more consistent: it is roughly twice as prevalent among highly consistent subjects as it is among inconsistent subjects, with moderately consistent subjects in the middle. Third, ambiguity seeking decreases as subjects become more consistent: it is three to four times more prevalent among inconsistent subjects than among highly consistent subjects, with moderately consistent subjects again in the middle.

To establish whether these differences are significant, we employ tests similar to those of Section 3: (Section 4.1) a test for the prevalence of each ambiguity attitude; (Section 4.2) a test at the level of the distributions for each consistency group as a whole; and (Section 4.3) a test involving the best-fitting model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Low** | **Medium** | **High** | **Significant**  **Difference?** |
| **Averse** | saa | 28 | 27 | 33 | 0.724 |
| waa | 43 | 44 | 45 | 0.967 |
|  |  |  |  |  |  |
| **Neutral** | stp | 24 | 33 | 45 | 0.026 |
| PIR | 21 | 33 | 48 | 0.002 |
|  |  |  |  |  |  |
| **Seeking** | sas | 19 | 10 | 4 | 0.007 |
| was | 33 | 22 | 11 | 0.004 |

= difference between groups is significant at the 5% level.

= difference between groups is significant at the 1% level.

**Table 6. Ambiguity attitude by consistency group.** Our null hypothesis that the proportion of an ambiguity attitude is independent of the consistency group. Our *p-*value in the final column is the chance of seeing the observed or greater difference between the proportions of a given attitude between the group with the lowest prevalence and the group with the highest prevalence if the null hypothesis were true, conditional on… .

**4.1 Is the prevalence of any attitude significantly different?**

In order to test whether the proportions of ambiguity attitudes differ significantly between consistency groups, our null hypothesis will be that the proportions are the same. (For a description of the test, see our Web Appendix.) The final column in Table 6 shows our results.

**4.2 Kolmogorov-Smirnov test**

To assess whether the distributions are different, we employ the K-S test as outlined in Section 3.2. Table 7 reports the results. While the K-S test does not allow us to conclude that medium and high consistency individuals choose differently, we can be very confident that the low consistency group displays a pattern of choices that differs from the pattern in the high consistency group. We can also be moderately confident that the low consistency group differs from the medium consistency group. This is in line with the results reported in Table 6.

|  |  |  |
| --- | --- | --- |
|  | Med | High |
| Low | 0.16 | 0.23 |
| Med |  | 0.12 |

= no more than 10% chance of wrongly rejecting the hypothesis that the distributions are drawn from the same underlying distribution.

= no more than 1% chance of wrongly rejecting the hypothesis that the distributions are drawn from the same underlying distribution.

**Table 7. Significantly different distributions?** For each comparison, we compute four K-S statistics comparing, respectively: (i) the distribution of *r1*; (ii) the distribution of *r2*; (iii) the distribution of the NE-SW diagonal sums; and (iv) the distribution of the SE-NW diagonal sums. We report the largest of these. We can be moderately confident that Low and Medium are different distributions; we can be very confident that Low and High are different distributions.

**4.3 Modeling the data**

Finally, we find the best-fitting parameter values for the model outlined in Section 3.3. Table 8 reports these parameters. Each of these models fits the data well, according to our K-S tests; moreover, the parameter range within which this result can be maintained is small. Our key finding is that inconsistent individuals are most likely to diverge from the sure-thing principle because ambiguity seeking is much more common among them. In addition, medium-consistency and high-consistency choosers are quite similar. This analysis therefore confirms the results of the previous two tests.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model parameters** | **Low** | **Medium** | **High** |
| 1-*a* (likelihood of diverging from stp in an ambiguity-seeking direction in any choice) | 0.24 | 0.12 | 0.08 |
| 1-*d* (likelihood of diverging from stp in an ambiguity-averse direction in any choice) | 0.28 | 0.22 | 0.22 |
| *a – d* (indicator of the prevalence of ambiguity aversion over seeking) | 0.04 | 0.10 | 0.14 |
| *µ* (mean of the distribution along the SW-NE diagonal) | 0.35 | 0.32 | 0.32 |
| *σ* (standard deviation) | 0.076 | 0.033 | 0.032 |

**Table 8. Best-fitting models**. Each of these instantiations of the model passes our K-S tests for the version(s) to which it is tailored at the strict 10% level (none of them can be confidently rejected).

**4.4 Conclusion regarding consistency**

While the prevalence of ambiguity aversion is the same among inconsistent and consistent subjects, other ambiguity attitudes depend on consistency. Ambiguity seeking is much more frequent among inconsistent choosers; consistent choosers are much more likely to be ambiguity neutral. Our results are therefore in tension with the findings in Chew et al. (2013) that ambiguity aversion is more prevalent among “sophisticated, competent choosers.” But they support the conclusions of Charness et al (2013) and Stahl (2014) that consistent individuals are likely to be ambiguity neutral. However, contrary to Stahl (2014), we find that the prevalence of ambiguity aversion does not depend on consistency; contrary to Charness at al. (2013), we find that a substantial share of consistent choosers are ambiguity averse. Rather than ambiguity aversion, we find that it is ambiguity seeking which disappears as choosers become consistent. At least in the relatively small world of our experiment, indifference to ambiguity is a sign of a subject who has made up her mind, while ambiguity seeking indicates irresolution.

**5. Concluding discussion**

We assessed both the sensitivity of ambiguity attitudes to framing and the consistency with which subjects display these attitudes within a single frame. Our key findings are:

1. Contrary to the predominant finding in the literature, a switch from a gain to a loss frame does not generate much of an effect on the distribution of ambiguity attitudes.
2. Making the ambiguous alternatives easier to recognize also has little effect.
3. The prevalence of ambiguity neutrality strongly increases with consistency, while ambiguity seeking virtually disappears as individuals become more consistent.
4. Across all our experiments, subjects’ behavior is explained quite well by a model in which subjects are basically ambiguity neutral, but occasionally diverge from ambiguity neutrality by responding to questions in an ambiguity-sensitive manner, with a somewhat stronger tendency to diverge in an ambiguity-averse direction than in an ambiguity-seeking direction. More consistent individuals diverge less frequently from ambiguity neutrality.

It may be that Savage’s (1954, p.16) insistence that Bayesian decision theory could only reasonably be expected to apply in a small world is relevant to these results. Savage is not very precise about what he would count as a small world, but there seems little doubt that he would have regarded the world in which Ellsberg originally set his paradox as small. Our finding that subjects are broadly ambiguity neutral can therefore be regarded as offering support for the claim that orthodox theories of rationality can sometimes have predictive power when experiments are framed as Savage would have recommended. Further support is provided in this paper by the additional finding that varying the framing in some ways[[13]](#footnote-13) that should be irrelevant to a rational subject does indeed leave the behavior largely unaffected. The same goes for the finding that eliminating subjects who choose inconsistently sharpens our results. We are anxious not to be thought to making wider claims for ambiguity neutrality. In particular, we see no necessary conflict between our results and those of papers that find widespread evidence of ambiguity sensitivity in experiments framed in terms more typical of

the economic problems that arise in macroeconomics and finance.

**References**

Abdellaoui, M., Vossmann, F. & Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science*, *51*, 1384–1399.

Baillon, A. & Bleichrodt, H. (forthcoming). Testing ambiguity models through the measurement of probabilities for gains and losses. *American Economic Journal: Micro*

Binmore, K. (2009). *Rational decisions*. Princeton: Princeton University Press.

Binmore, K., Stewart, L. & Voorhoeve, A. (2012). How much ambiguity aversion? Finding indifferences between Ellsberg’s risky and ambiguous bets. *Journal of Risk and Uncertainty*, *45*, 215–238.

Bradley, R. (2014). Ellsberg’s Paradox and the Value of Chances. Working paper, Department of Philosophy, Logic and Scientific Method, LSE.

Budescu, D., Kuhn, K., Kramer, K. & Johnson, T. (2002). Modeling certainty equivalents for imprecise gambles. *Organizational Behavior and Human Decision Processes* *88*, 748–768.

Chakravarty, S. & Roy, J. (2009). Recursive expected utility and the separation of attitudes towards risk and ambiguity: An experimental study. *Theory and Decision*, *66*, 199–228.

Charness, G., Karni, E. & Levin, D. (2013). Ambiguity attitudes and social interactions: An experimental investigation. *Journal of Risk and Uncertainty* 46, 1–25.

Chew. S.H., Ratchford, M. & Sagi, J. (2013). You need to recognize ambiguity to avoid it. Working paper, NUS. (Accessed September 2014).

Cohen, M., Jaffray, J.-Y. & Said, T. (1987). Experimental comparisons of individual behavior under risk and under uncertainty for gains and for losses. *Organizational Behavior and Human Decision Processes*, *39*, 1– 22.

De La Resende, J. & Wu, G. (2010). Competence effects for choices involving gains and losses. *Journal of Risk and Uncertainty*, *40*, 109–132.

Di Mauro, C. & Maffioletti, A. (1996). An experimental investigation of the impact of ambiguity on the valuation of self-insurance and self-protection. *Journal of Risk and Uncertainty*, *13*, 53–71.

Dimmock, S., Kouwenberg, R. & Wakker, P. (2014). Ambiguity attitudes in a large representative sample. Working paper on people.few.eur.nl/wakker/pdf/dwk\_amb\_finance.pdf (Accessed September 2014).

Du, N. & Budescu, D. (2005). The effects of imprecise probabilities and outcomes in evaluating investment options. *Management Science* 51, 1791–1803.

Duersch, P., Roemer, D. & Roth, B. (2013). Intertemporal stability of ambiguity preferences. Discussion Paper Series No. 548, Heidelberg University.

Einhorn, H. & Hogarth, R. (1986). Decision making under ambiguity. *Journal of Business*, *59*, S225–S250.

Eisenberger, R. & Weber, M. (1995). Willingness-to-pay and willingness-to-accept for risky and ambiguous lotteries. *Journal of Risk and Uncertainty*, *10*: 223–233.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643– 669.

# Friedl, A., Lima de Miranda, K. & Schmidt, U. (2014). Insurance demand and social comparison: An experimental analysis. *Journal of Risk and Uncertainty*, *48*, 97-109.

Ho, J., Keller, L., & Keltyka, P. (2002). Effects of outcome and probabilistic ambiguity on managerial choices. *Journal of Risk and Uncertainty*, *24*, 47–74.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica,75*, 503-536.

Inukai, K. & Takahasi, T. (2009). Decision under ambiguity: effects of sign and magnitude. *International Journal of Neuroscience*, *119*, 1170–1178.

Kahn, B. & Sarin, R. (1988). Modeling ambiguity in decisions under uncertainty. *Journal of Consumer Research*, *15*, 265–272.

Keren, G. & Gerritsen, L. (1999). On the robustness and possible accounts of ambiguity aversion. *Acta Psychologica*, *103*, 149–172.

Lopes, R., Reid, I. & Hobbes, P. (2007). The Two-Dimensional Kolmogorov–Smirnov Test. XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research, Amsterdam. (Published by Proceedings of Science).

MacCrimmon, K. & Larsson, S. (1979). Utility theory: Axioms versus ‘paradoxes’. In M. Allais, & O. Hagen (Eds.), *Expected utility and the Allais paradox* (pp. 333–409). Dordrecht: Reidel.

Mangelsdorff, L. & Weber, M. (1994). Testing Choquet expected utility. *Journal of Economic Behavior and Organization*, *25*, 437–457.

Raiffa, H. (1961). Risk, ambiguity and the Savage axioms; A comment. *Quarterly Journal of Economics*, *75*, 690-694.

Savage, L. (1954). *The foundations of statistics*. New York: Wiley.

Stahl, D. (2014) Heterogeneity of ambiguity preferences. *The Review of Economics and Statistics*, *96*: 609–617.

Trautmann, S. & Van der Kuijlen, G. (forthcoming). Ambiguity attitudes. In G. Keren & G. Wu *Blackwell Handbook of Judgment and Decision-Making*. Oxford: Blackwell.

Trautmann, S. & Wakker, P. (2014). Making the Anscombe-Aumann approach to ambiguity suited for descriptive applications. Working paper on <http://people.few.eur.nl/wakker/newps.htm> (Accessed September 2014).

Vidier, F., Martinsson, P. & Medhin, H. (2012). Stake effects on ambiguity attitudes for gains and losses. Working paper, WZB Berlin from <http://www.ferdinandvieider.com/stake_effects_ambiguity.pdf> (Accessed September 2014).

Viscusi, W.K. & Chesson, H. (1999). Hopes and fears: The conflicting effects on risk ambiguity. *Theory and Decision*, *47*, 153–178.

Viscusi, W.K. & Magat, W. (1992). Bayesian decisions with ambiguous belief aversion. *Journal of Risk and Uncertainty*, *5*, 371–378.

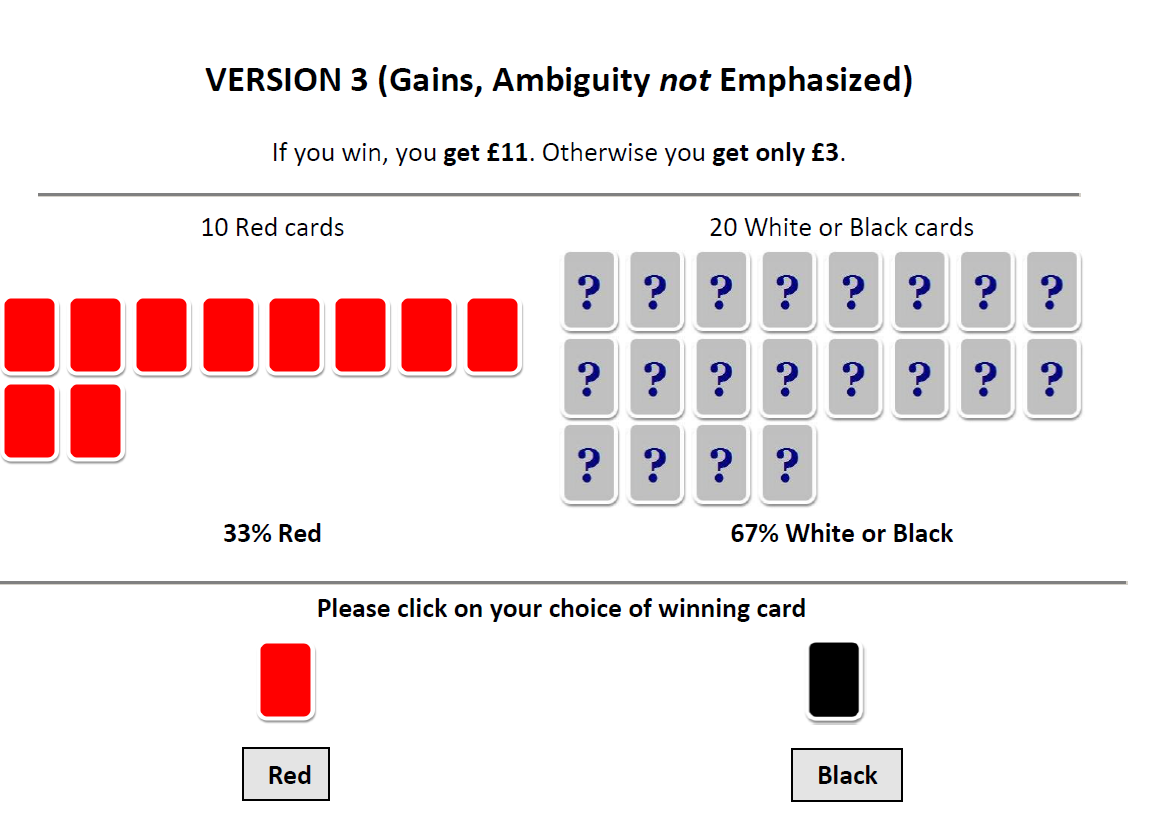
Wakker, P. (2010). *Prospect theory for risk and ambiguity*. Cambridge: Cambridge University Press.

**Appendix A: Sample interfaces.**

**Version 3 (gains; subjects must infer ambiguous option)**

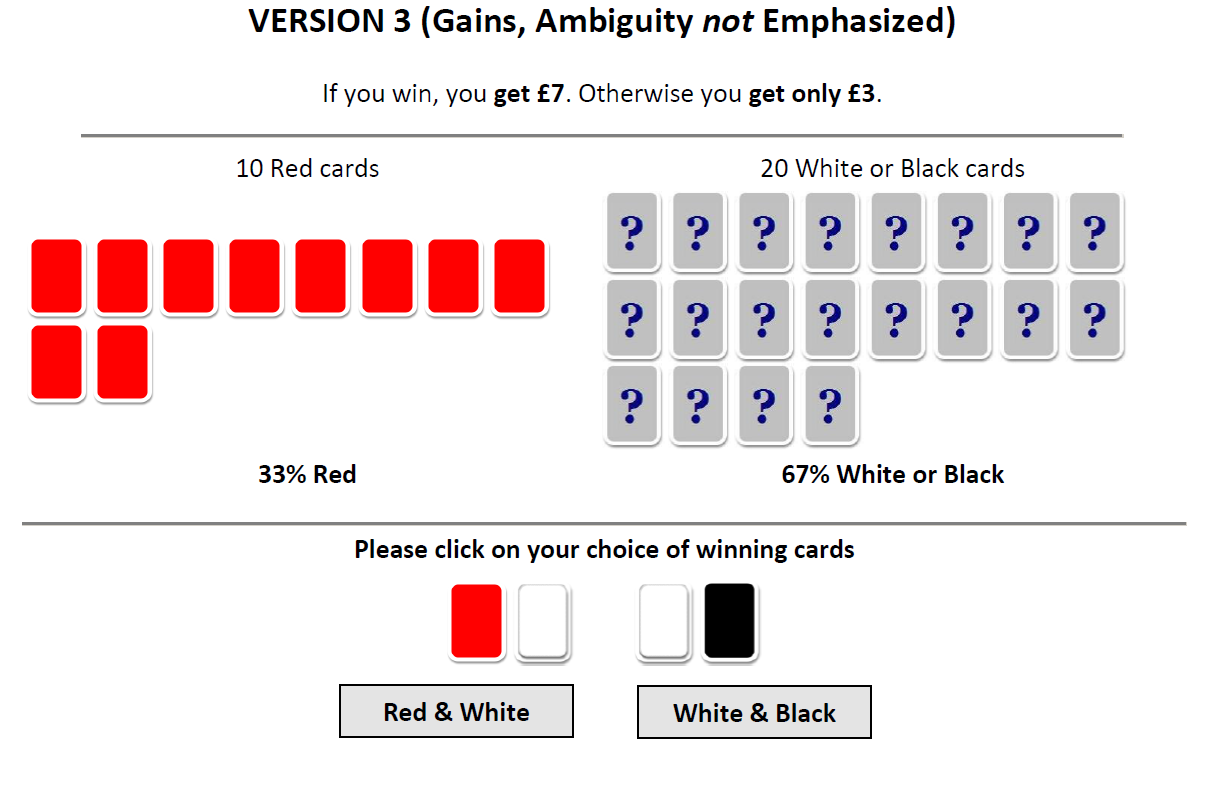
***R versus B screen****:*

If you win, you **get £11**; otherwise, you get only **£3.**



***B&W versus R&W screen****:*

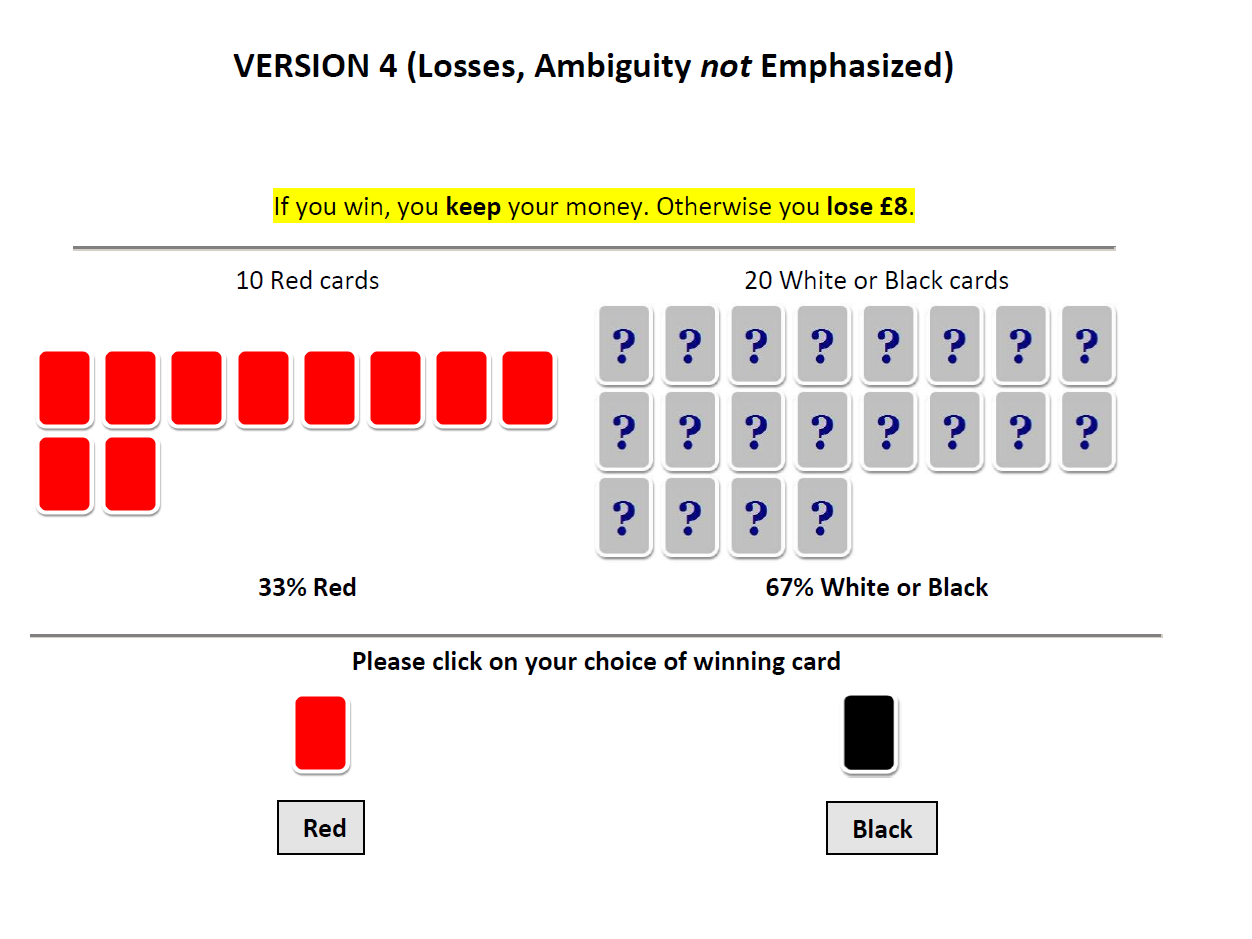
If you win, you **get £7;** otherwise, you get only **£3.**



**Version 4 (losses; subjects must infer ambiguous alternative)**

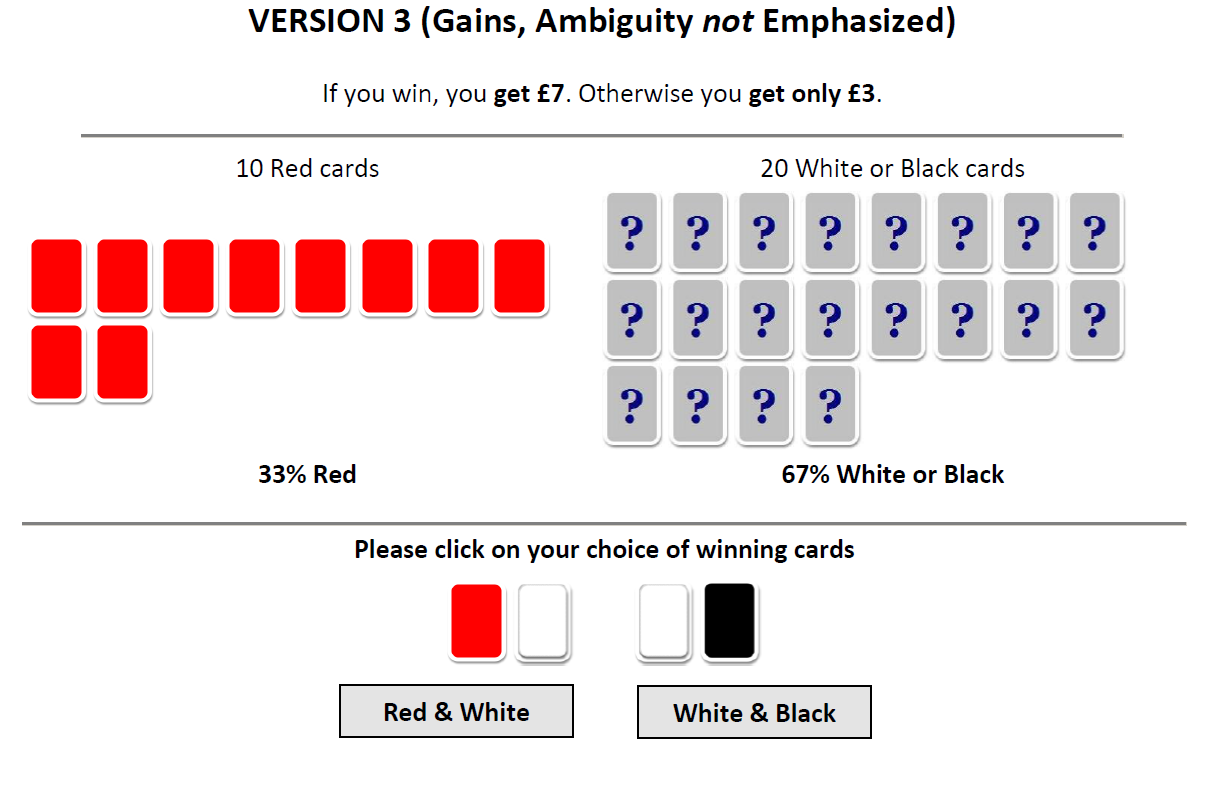
***R versus B screen****:*

If you win, you **keep** your money**;** otherwise you **lose £8.**



***B&W versus R&W screen:***

If you win, you **keep** your money; otherwise you **lose £7.**



**Version 5 (losses; subjects are told which alternative is ambiguous)**

***R versus B screen****:*

If you win, you **keep** your money. Otherwise you **lose £8**.

|  |
| --- |
| 10 Red cards |
| http://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.png |
| 20 White or Black cards |
| http://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpg http://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpg |
| **Please click on your choice of winning card** |

|  |  |  |
| --- | --- | --- |
| 33% Red |  | From 0% to 67% Black |
| http://alkami.org/ells/ells04/cardRed.png |  | http://alkami.org/ells/ells04/cardBlack.png |

***B&W versus R&W screen:***

If you win, you **keep** your money. Otherwise you **lose £4**.

|  |
| --- |
| 10 Red cards |
| http://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardRed.png |
| 20 White or Black cards |
| http://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpg http://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpghttp://alkami.org/ells/ells04/cardWhich.jpg |
| **Please click on your choice of winning cards** |

|  |  |  |
| --- | --- | --- |
| From 33% to 100% Red or White |  | 67% White or Black |
| http://alkami.org/ells/ells04/cardRed.pnghttp://alkami.org/ells/ells04/cardWhite.png |  | http://alkami.org/ells/ells04/cardWhite.pnghttp://alkami.org/ells/ells04/cardBlack.png |

**Version 6 (gains; subjects are told which alternative is ambiguous)**

***R versus B screen:***

If you win, you **get £11**. Otherwise you **get only £3**.

|  |
| --- |
| **10 Red cards**  http://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.png |
| **20 White or Black cards**  http://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpg http://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpg |
| **Please click on your choice of winning card** |

|  |  |  |
| --- | --- | --- |
| **33% Red** http://alkami.org/ells/ells05/cardRed.png |  | **From 0% to 67% Black** http://alkami.org/ells/ells05/cardBlack.png |

***B&W versus R&W screen:***

If you win, you **get £7**. Otherwise you **get only £3**.

|  |
| --- |
| **10 Red cards**  http://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardRed.png |
| **20 White or Black cards**  http://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpg http://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpghttp://alkami.org/ells/ells05/cardWhich.jpg |
| **Please click on your choice of winning cards** |

|  |  |  |
| --- | --- | --- |
| **From 33% to 100% Red or White** http://alkami.org/ells/ells05/cardRed.pnghttp://alkami.org/ells/ells05/cardWhite.png |  | **67% White or Black** http://alkami.org/ells/ells05/cardWhite.pnghttp://alkami.org/ells/ells05/cardBlack.png |

**Appendix B Tests for aggregating versions in Section 4**

Version 2 is explained in detail in Binmore et al. (2012). Figure B1 reports the results for Version 2.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Version 2** (gains; ambiguity inferred) | | | | | | | | | | |
|  | 76 subjects, percentages, 4 observations each | | | | | | | | | | |
|  | 1 |  | 2 | 1 | | 1 | |  | |  | |
|  |  |  | 2 | 1 | 1 | |  | | 1 | |  | |
|  |  |  | 3 | 5 | 3 | | 2 | | 1 | |  | |
|  |  |  | 4 | 9 | 7 | | 1 | |  | |  | |
|  |  | 7 | 16 | 7 | | 1 | |  | |  | |
|  |  | 1 | 7 | 3 | 2 | | 1 | |  | |  | |
|  |  |  | 4 | 3 |  | |  | |  | |  | |
| 0 |  |  |  |  | 1 | |  | |  | | 1\* | |
|  |  |  |  |  | |  | |  | |  | | | |

**Fig. B1. Results for Version 2.** Shaded squares indicate an especially high concentration of responses.

\* This cell entry was erroneously omitted in Binmore et al (2012, p. 226). (The analysis in that paper was done with the correct data.)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2 | 3 | 4 | 5 | 6 |  |  | 6 | 6&3 |
| 1 | 0.27 | 0.24 | Unnecessary, Version 1 excluded. | | |  | 2&3 | 0.14 |  |
| 2 |  | 0.07 | 0.11 | 0.16 | 0.13 |  | 2 |  | 0.04 |
| 3 |  |  | 0.09 | 0.17 | 0.15 |  |  |  |  |
| 4 |  |  |  | 0.11 | 0.06 |  |  | Losses  (4&5) | Explicit (5&6) |
| 5 |  |  |  |  | 0.09 |  | Gains  (2, 3&6) | 0.08 |  |
| 2&3&6 |  |  |  | 0.12 |  |  | Implicit  (2,3 & 4) |  | 0.12 |
| 2&3&4&6 |  |  |  | 0.11 |  |  |  |  |  |

**Table B1. Kolmogorov-Smirnov test.** For each comparison between (combinations of) versions, we compute four K-S statistics comparing, respectively: (i) the distribution of *r1*; the distribution of *r2*; (iii) the distribution of the NE-SW diagonal sums; and (iv) the distribution of the SE-NW diagonal sums. We report the largest of these. We focus on the salient groupings and on the individual version that, besides Version 1, was most different from others (Version 5). Version 1 fails these tests; all other combinations investigated pass our K-S tests.

= no more than 5% chance of wrongly rejecting the hypothesis that the distributions are drawn from the same underlying distribution.

1. This paper was presented at the LSE. We thank those present as well as Richard Bradley, Casey Helgeson, and Peter Wakker for comments. We gratefully acknowledge support from STICERD, from the British Arts and Humanities Research Council through grant AH/J006033/1 and from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant 295449. We thank LSE’s Behavioural Science Lab for the use of their facilities. [↑](#footnote-ref-1)
2. See Keren and Gerritsen (1999) and Inukai and Takahasi (2009). Di Mauro and Maffioletti (1996) generally find ambiguity aversion for losses, but most of their findings are not significant. [↑](#footnote-ref-2)
3. See Cohen et al (1987); de la Resende and Wu (2010); Du and Budescu (2005); Einhorn and Hogarth (1986); Eisenberger and Weber (1995); Mangelsdorff and Weber (1994); Friedl et al. (2014); and Trautmann and Wakker (2014). [↑](#footnote-ref-3)
4. See Abdellaoui et al. (2005); Chakravarty and Roy (2009); and Baillon and Bleichrodt (forthcoming). [↑](#footnote-ref-4)
5. See Kahn and Sarin (1988); Viscusi and Chesson (1999); Di Mauro and Maffioletti (2004); and Vidier et al. (2012). [↑](#footnote-ref-5)
6. See Viscusi and Chesson (1999) and Bradley (2014) for proposed psychological explanations for such a shift. [↑](#footnote-ref-6)
7. For example, Stahl (2014), finds that 60% of subjects choose close to randomly; Charness et al. (2013) find that 20% of subjects choose incoherently; Duersch et al. (2013) find that, after a short interval, 29% choose differently in repetitions of the same question; and Dimmock et al. (2014) find that up to 45% do not choose according to their previously stated preferences. [↑](#footnote-ref-7)
8. If one views adherence to the compound lottery axiom as a requirement of rationality, then the results in Halevy (2007) show that ambiguity aversion is absent among those who are rational in this sense. [↑](#footnote-ref-8)
9. A probability-matching approach is also used in MacCrimmon and Larsson (1979), Kahn and Sarin (1988), and Viscusi and Magat (1992). Dimmock et al. (2014) provide theoretical support for this method. [↑](#footnote-ref-9)
10. We omit the ‘lax’ criteria for ambiguity-sensitive attitudes and of the sure-thing principle used in Binmore et al. (2012) because the consistency criterion of Section 4 now handles any mistakes. [↑](#footnote-ref-10)
11. There will often be more than one representative pair (*k*, *l*) that achieves the minimum *Mn*. Whenever it matters, we take the average of these values. [↑](#footnote-ref-11)
12. See Charness et al. (2013), Duersch et al. (2013), Dimmock et al. (2014), and Stahl (2014). [↑](#footnote-ref-12)
13. Some, but not all. In particular, Version 1 of our experiment is excluded from our tests because describing “black or white wins” as “not-red wins” seems to make a significant difference to subjects. [↑](#footnote-ref-13)