Ambiguity Attitudes, Framing, and Consistency
Corresponding author: Alex Voorhoeve, Philosophy, Logic and Scientific Method, LSE, Houghton Street, London WC2A 2AE, United Kingdom
E-mail: a.e.voorhoeve [at] lse.ac.uk;
Tel: +44-7915150106; Fax: +44 (0)20 7831 9045
Ken Binmore, Economics, UCL,
Gower Street, London, WC1E 6BT, United Kingdom
Arnaldur Stefansson, Economics, Uppsala University
P.O. Box 513 SE-751 20 Uppsala, Sweden
Lisa Stewart, Research School of Psychology
The Australian National University,
Building 39, Science Road, Canberra, ACT 2601, Australia

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 attitudes depend on consistency: ambiguity seeking is much more frequent among inconsistent choosers; consistent choosers are much more likely to be ambiguity neutral.

Keywords: ambiguity attitudes, framing, consistency of choice, Ellsberg paradox,
JEL classification: C91, D03.

Acknowledgements

This paper was presented at the LSE, Uppsala University, and Warwick University. We thank those present as well as Richard Bradley, Casey Helgeson, Stefan Trautmann, and Peter Wakker for comments. We gratefully acknowledge support from the Suntory and Toyota International Centres for Economics and Related Disciplines, the British Arts and Humanities Research Council through grant AH/J006033/1 and 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.
Ambiguity Attitudes, Framing, and Consistency

Introduction

Experiments on the Ellsberg paradox generally show substantial aversion to ambiguity when people face choices involving moderate to high-probability gains. The literature also finds that ambiguity aversion is not robust to changes in framing, with subjects being more ambiguity loving when faced with losses (Trautmann and van der Kuijlen, 2014). It also suggests that some subjects find it hard to recognize ambiguity, so that emphasizing it generates a stronger response (Chew et al. 2013). In this paper, we examine whether these findings hold up by reframing one of Ellsberg’s experiments in terms of losses rather than gains and by emphasizing ambiguity. Neither change has a significant effect.

We also explore the hypothesis advanced by Charness et al. (2013) and Stahl (2014) that sensitivity to ambiguity is primarily due to “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 partial support for this hypothesis: the higher the level of consistency we require, the more ambiguity neutral and the less ambiguity seeking our subjects become. Ambiguity aversion, however, remains modest but constant across all levels of consistency.

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.

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 a risky gamble $R_r$, in which a prize is won if and only if a red ball is drawn, and an ambiguous gamble $B_a$, in which this prize is won if and only if a black ball is drawn. They are also asked to choose between risky $B&W_r$, in which both black and white win, and ambiguous $R&W_a$, in which both red and white win. Ellsberg predicted that most people will strictly prefer $R_r$ to $B_a$ and $B&W_r$ to $R&W_a$. 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 behaviour 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 $R_r$ and $B_a$, and between $B&W_r$ and $R&W_a$. 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.1 Some studies find ambiguity neutrality;2 others report a predominance of ambiguity seeking;3 some find a transition from approximate neutrality to ambiguity seeking as the probability of loss increases.4 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).5

(ii) Difficulty in recognizing ambiguity? In the space of moderate to high probability gains, 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 two-colour risky urn with a known probability of winning of 1/2 and drawing from a two-colour ambiguous urn.6

There are various hypotheses regarding the origin of this difference. One runs as follows. In the single-urn experiment, a subject reveals ambiguity aversion only if they both choose $R_r$ over $B_a$ when the known probability of the former is 1/3 and choose $B&W_r$ over $R&W_a$ when the known probability of the former is 2/3. This difference in the known probability of winning in these two bets may matter: some studies find that subjects who are ambiguity averse for moderate to high probability gains are ambiguity loving for low probability gains (Trautmann and van der Kuijlen 2014; Kocher et al. 2014). A subject who fails to register ambiguity aversion in the single-urn experiment (because she fails to choose $R_r$ over $B_a$ when the known probability, at 1/3, is perceived as low) might therefore display ambiguity aversion in the two-urn experiment (when the known probability, at 1/2, is perceived as moderate).

If this hypothesis were true, then in a single-urn experiment, one would expect to find a greater preference for the risky option in two-winning-colour choices such as between $B&W_r$ and $R&W_a$ than in the single-winning-colour choice between $R_r$ and $B_a$. 

---

1 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.
2 See Cohen et al. (1987); de la Resende and Wu (2010); Du and Budescu (2005); Einhorn and Hogarth (1986); Eisenberger and Weber (1995); Mangelsdorf and Weber (1994); Friedl et al. (2014); and Trautmann and Wakker (2014).
3 See Abdellaoui et al. (2005); Chakravarty and Roy (2009); and Baillon and Bleichrodt (forthcoming).
4 See Kahn and Sarin (1988); Viscusi and Chesson (1999); Di Mauro and Maffioletti (2004); and Vidier et al. (2012).
5 See Viscusi and Chesson (1999) and Bradley (2014) for proposed psychological explanations for such a shift.
6 Chew et al. (2013; Table 1) report that across 10 studies using the two-urn Ellsberg experiment, the average proportion of ambiguity-averse responses is 65.6%; across 7 studies using the single-urn experiment, the average proportion of ambiguity-aversion is 27.1%. Stahl (2014), which is not included in their review, equally finds lower ambiguity aversion in the one-turn set-up.
Contrary to this hypothesis, in previous work (Binmore et al. 2012, Figure 7), we found that the preference for the risky option is considerably stronger in the latter choice. We conjectured that this is due to the fact that some subjects find it difficult to grasp that in the two-winning-colour variants, $B & W_r$ is the risky option, and $R & W_a$ is the ambiguous option. Our conjecture gains support from Chew et al. (2013), who provide evidence that few subjects are able to discern ambiguity in more complex cases, but that those who do recognize ambiguity are generally ambiguity averse in a gain frame. The difficulty in recognizing ambiguity in the more complex, two-winning-colour choices is therefore an alternative candidate explanation for the observed disparity in ambiguity aversion between single-urn and two-urn experiments.

(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.\(^7\) 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 Charness et al. (2013), then 15% are ambiguity seeking, 75% are ambiguity neutral, and 10% are ambiguity averse. Similarly, if one excludes the 60% of subjects classified as “random choosers” from Stahl (2014), then 75% of the remaining subjects are ambiguity neutral and 25% are ambiguity averse.\(^5\) 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 aversion towards ambiguity neutrality and/or ambiguity seeking 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.

\(^7\) 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.

\(^5\) 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.
2. Experiments

Our experiments use decks of cards rather than urns and employ a titration to estimate the value $r_1$ of the probability $r$ of drawing a red card that makes a particular subject indifferent between $R_r$ and $B_a$ and the value $r_2$ of drawing red that makes the same subject indifferent between $B&Wr$ and $R&W_a$. The aim of this titration is to locate the values of $r_1$ and $r_2$ within eight subintervals of $[0, \frac{1}{2}]$ using the scheme illustrated in Figure 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).

---

**Fig. 1** The tree shows the questions implicitly asked about the subjects’ preferences when the probability of red is $r$ in order to locate $r_1$ in one of eight subintervals of $[0, \frac{1}{2}]$. For example, someone who answers $B, R, R_r$ is assigned a value of $r_1$ satisfying $\frac{1}{3} \leq r_1 \leq \frac{3}{8}$. The same tree is used to locate $r_2$, except that $B&Wr$ replaces $B_a$ and $R&W_a$ replaces $R_r$.

---

---

\[ A \text{ probability-matching approach is also used in MacCrimmon and Larsson (1979), Kahn and Sarin (1988), and Viscusi and Magat (1992). Dimmock et al. (2014) provide theoretical support for this method.} \]
**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 \( r_1 = r_2 = \frac{1}{3} \). Since \( \frac{1}{3} \) is an endpoint of two of our intervals (see Fig. 1), any value of \((r_1, r_2)\) that lies in one of the four squares of our chessboard with a corner at \((\frac{1}{3}, \frac{1}{3})\) 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 \( R_r \) is preferred to \( B_a \) and \( B&W_r \) to \( R&W_a \) when the probability of 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 \( r_1 < \frac{1}{3} \) and \( r_2 > \frac{1}{3} \) as cases of strong ambiguity aversion (saa). We treat pairs \((r_1, r_2)\) with \( r_1 > \frac{1}{3} \) and \( r_2 < \frac{1}{3} \) as cases of strong ambiguity seeking (sas).

**Common properties of the experiments.** The experiment took place in a classroom divided into a common area at the front and private cubicles. Each subject was seated in a cubicle in front of a 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 \( R_r \) and \( B_a \) and between versions of \( B&W_r \) and \( R&W_a \) 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 common area at 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 world” 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 \( R_r \) and \( B_a \). Subjects were then navigated through a similar tree to estimate the value \( r_2 \) that makes a subject indifferent between \( B&W_r \) and \( R&W_a \). 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 3g,i (the subscript denotes that it is a gain frame and ambiguity was implicit) and revised it to study the effects of switching to a loss frame and of making explicit any ambiguity present in an alternative.

Version 4l,i 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. Every screen that presented a “for real” choice informed them that they would keep their coins if they won and that they would lose a number of their coins otherwise.

Version 5l,e was identical to Version 4l,i, except that a greater effort was made throughout to make clear which alternative was risky and which alternative was ambiguous and how ambiguous it was. The effort started in the introduction. The text accompanying the two practice choices explained that betting on the risky option would give them a known chance of winning, whereas the ambiguous would give them an unknown chance in a range. It also gave the known probability of the former and the range of the latter. For example, the text accompanying the two-winning-colour practice choice read:
If you bet on BLACK & WHITE, your chance of winning is 71% (because the number of cards that are BLACK or WHITE is 15 out of a total of 21 cards).

If you bet on RED & WHITE, your chance of winning ranges from 29% to 100% (because the number of cards that are RED or WHITE can range from 6 to 21 out of a total of 21 cards).

Subsequently, in all real choices, the probability of winning (or range of probabilities of winning) was provided for each option. For example, in a choice between B&Wr and R&Wa, subjects were told that if they chose to bet on B&Wr, then their chance of winning was 67%, while if they chose to bet on R&Wa, then their chance of winning was anywhere from 33% to 100%.

Version 6g,e reverted to the gain frame but made explicit in the manner just outlined which alternative was ambiguous and how ambiguous it was. Appendix A provides sample screens for all versions.

Version 3g,i was performed at the ELSE lab at University College London; subsequent versions were performed at the Behavioural Science Lab at the London School of Economics. These labs recruit jointly and their demographic mix is similar (GIVE KEY DATA).

3. Results on framing

As in Binmore et al. (2012), there was no evidence that subjects adjusted their choices during the experiment. 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 1.

Across all versions, roughly 30% of subjects 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 at 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) When compared to gains, for losses there will be less ambiguity aversion and substantially more ambiguity-neutral or ambiguity-seeking behaviour.

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

We employ three complementary tests of these hypotheses: (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.
Fig. 3: Aggregate results by frame. 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.)
Table 1. Percentage shares of ambiguity attitudes in different versions.\(^{10}\) Entries in columns 3\(_{g,i}\), 4\(_{l,i}\), 5\(_{g,e}\), 6\(_{g,e}\), Gain, Loss; Impl.; and Expl. report fractions of the population in each version (or combination of versions) for each of the ambiguity attitudes we distinguish. The table also 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 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 hypotheses (i) and (ii) has several parts.

(a) We make a pairwise comparison of the prevalence of each ambiguity attitude in Versions 3 through 6 under the null hypothesis that observations from each subject are

\(^{10}\) 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.
independent and fall into a given ambiguity category with the same probability. We then
estimate the probability (p-value) of obtaining the observed absolute difference or more in the
prevalence of each attitude between each pair of versions of the experiment 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 is
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. Our
null hypothesis is that the prevalence of an attitude is independent of this change in frame.

Table 1 reports our results (for a description of our test and all values of this test, see
our Web Appendix on http://personal.lse.ac.uk/voorhoev/ellsberg.htm).

Gains versus losses: The mere switch from gains to losses while leaving subjects to infer the
ambiguous option in every choice (Version 3g,i to Version 4i,l) makes no significant
difference for any ambiguity attitude. The mere switch from gains to losses while making the
ambiguous option explicit (Version 6g,e to Version 5l,e) 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 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 effect in a gain frame (Version 3g,i versus Version 6g,e). In a loss frame
(Version 4i,l versus Version 5l,e), 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 with the
aggregate of versions that make it explicit. 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 5l,e. 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 receives 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 test 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 2 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 ambiguity implicit (Version 3g,i to Version 4l,i) makes no significant difference. The mere switch from gains to losses while making ambiguity explicit (Version 6g,e to Version 5l,e) 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 3g,i versus Version 6g,e) or a loss frame (Version 4l,i versus Version 5l,e). Nor does the K-S test reveal any significant difference between the “implicit” and “explicit” groupings.

In sum, the K-S test does not permit us to say with confidence of any version that it generates 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.
Table 2. Significantly different distributions? For each comparison between versions, we compute four K-S statistics comparing, respectively: (i) the distribution of $r_1$; (ii) the distribution of $r_2$; (iii) the distribution of the NE-SW diagonal sums in our 8×8 chessboards (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 roughly 0.19; for the aggregates, it is roughly 0.14.)

3.3 Modelling the data

Our third test involves fitting the 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 $\mu$ and standard deviation $\sigma$. 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 behaviour, 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: $\mu$, $\sigma$, $a$, $d$. 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, $K_i$ and $K_{ij}$, using the observed data on the main diagonal for $K_i$, and the sums of data points along parallels to the main diagonal for $K_{ij}$. Low values of $K_i$ indicate that the observed data points on the main diagonal of our data matrix are consistent with our model. Low values of $K_{ij}$ indicate that deviations from ambiguity neutrality are consistent with the model. Table 3 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 be confidently 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 seeking changes little in response to changes in framing. It indicates that ambiguity aversion is more prevalent than ambiguity seeking. Our model therefore provides no evidence for a substantial shift in the balance of ambiguity aversion and ambiguity seeking between the various versions, or between aggregated gain versions on the one side and aggregated loss versions on the other. (Our model is available on the aforementioned website.)

| Model parameters | \( 3_{g,i} \) | \( 4_{l,i} \) | \( 5_{e} \) | \( 6_{e} \) | Gain | Loss |
|------------------|-------------|-------------|-------------|-------------|------|------|
| \( 1-a \) (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 |
| \( \mu \) (mean of the distribution along the SW-NE diagonal) | 0.31 | 0.32 | 0.33 | 0.33 | 0.32 | 0.32 |
| \( \sigma \) (standard deviation) | 0.035 | 0.045 | 0.060 | 0.056 | 0.045 | 0.051 |

| Fit | \( K_I \) | \( K_{II} \) |
|-----|-------------|-------------|
| Gain | 0.06 | 0.08 | 0.10 | 0.05 | 0.05 | 0.09 |
| Loss | 0.05 | 0.03 | 0.07 | 0.02 | 0.02 | 0.04 |

**Table 3. The best-fitting models.** In the top part of the table, the columns give the parameters of the model that best fits the data for the respective versions (or aggregation of versions). The bottom part gives our two K-S statistics for these instantiations of the model (lower numbers indicate a better fit). All best-fitting models pass our K-S tests. For individual versions, the lower limit for the 10% confidence level is roughly 0.13; for combined versions, it is roughly 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 behaviour and substantially more ambiguity-neutral and/or ambiguity-seeking behaviour for losses than for gains.*
In our experiments, the prevalence of ambiguity-averse behaviour 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 behaviour. 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, we find no significant shift towards ambiguity neutrality and/or ambiguity seeking, so that our experiments also refute hypothesis (ii) for losses. We note that the failure of this hypothesis in the gain frame means that we lack an explanation of the finding, reported in Section 1, that ambiguity aversion appears less prevalent in single-urn experiments.

4. Results on consistency

We now turn to the following questions: How consistent are individual choices within a given frame? Do ambiguity attitudes correlate with consistency? To address these questions, we look at individual-level data. As mentioned, each experiment yields four estimates for a subject’s \((r_1, r_2)\). We measure how inconsistent a subject \(n\) is by finding the minimal number \(M_n\) 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 \((r_1, r_2)\). If \(M_n \leq 12\), one can think of the representative pair as the maximum likelihood estimator of an agent who has a “true pair” \((r_1, r_2)\) but who makes mistakes with probability \(p < \frac{1}{2}\) when answering titration questions. More precisely, we define the “distance” from square \((k, l)\) to square \((i, j)\) by
\[
d = |k - i| + |l - j|.
\]
Given our four data points \((i_1, j_1), (i_2, j_2), (i_3, j_3), (i_4, j_4)\) for individual \(n\):
\[
m(k, l) = \sum_{s=1}^{4} [|k - i_s| + |l - j_s|]
\]
for each \((k, l)\). The measure \(M_n\) of consistency for each individual is then
\[
M_n = \min m(k, l)
\]
taken over all \((k, l)\).

Table 4 reports the percentage share of subjects at each level of consistency for several versions of our experiment. In order to assess the relationship between our measure of consistency 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 different at the 10% significance level from any other version. This excludes only Version 1.

11 There will often be more than one representative pair \((k, l)\) that achieves the minimum \(M_n\). Whenever it matters, we take the average of these values.
2. For the remaining versions, 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 2g,i, 3g,is and 6g,e) to “all loss versions” (4l,i and 5l,e). No version or combination of versions is excluded by this test (see our Online Appendix). The upshot is that this procedure permits us to aggregate Versions 2-6 of our experiment. However, since we saw in Section 3.1 that Version 5l,e 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 5l,e from our data.

Next, we classify subjects into three groups: high consistency ($0 \leq M_n \leq 2$); medium consistency ($3 \leq M_n \leq 5$); and low consistency ($6 \leq M_n$). 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 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).\footnote{See Charness et al. (2013), Duersch et al. (2013), Dimmock et al. (2014), and Stahl (2014).}

| $M_n$ | 2g,i | 3g,i | 4l,i | 5l,e | 6g,e | Aggregate |
|-------|------|------|------|------|------|-----------|
| 0     | 9    | 3    | 16   | 6    | 11   | 9         |
| 1     | 4    | 15   | 9    | 6    | 20   | 11        |
| 2     | 12   | 9    | 9    | 8    | 5    | 9         |
| 3     | 17   | 15   | 19   | 10   | 11   | 15        |
| 4     | 13   | 16   | 18   | 26   | 10   | 16        |
| 5     | 21   | 11   | 9    | 8    | 17   | 13        |
| 6     | 15   | 8    | 8    | 11   | 7    | 10        |
| 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 $M_n$ | 4    | 4    | 3    | 4    | 4    | 4         |
| $N$   | 76   | 80   | 90   | 73   | 81   | 400       |

Table 4. The percentage share of subjects for each level of consistency, for each version. $M_n$ is our measure of consistency; it is the aggregate “distance” in our $8 \times 8$ 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. The entries in columns labelled 2g,i through 6g,e give the percentage of the population for each version for each level of consistency; the next column reports the percentages in the aggregate of these versions.

So defined, how does consistency relate to ambiguity attitude? Table 5 gives our results. Three things are apparent. First, the prevalence of ambiguity aversion is nearly identical in all three groups. Second, ambiguity-neutral behaviour 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) the Kolmogorov-Smirnov test for differences between the distributions for each consistency group over our 64-square grid; 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                   |

Table 5. Ambiguity attitude by consistency group. The first three columns show the fractions of the population from the Low, Medium and High consistency groups for each of our six ambiguity attitudes. The fourth column is the probability (p-value) of seeing the observed difference or more between the High and Low fractions on the null hypothesis that consistency does not influence ambiguity attitudes.

4.1 Is the prevalence of any attitude significantly different?

The fact that the fraction of observations satisfying PIR increases markedly with our measure of consistency leads us to compute the probability (p-value) of seeing the observed difference or more between the High and Low fractions on the null hypothesis that consistency does not affect ambiguity attitudes. In particular, the deviations in the High and Low fractions from the Medium fraction are assumed independent. We condition this probability on the observed ordering of the fractions because this ordering was instrumental in our choosing the test (see our Online Appendix). We use the same test for the other ambiguity attitudes (although the fractions are only approximately monotone). The final column in Table 5 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 6 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.

Table 6. Significantly different distributions? For each comparison, we compute four K-S statistics comparing, respectively: (i) the distribution of $r_1$; (ii) the distribution of $r_2$; (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 Modelling the data

Finally, we find the best-fitting parameter values for the model outlined in Section 3.3. Table 7 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 |
| $\mu$ (mean of the distribution along the SW-NE diagonal) | 0.35 | 0.32 | 0.32 |
| $\sigma$ (standard deviation) | 0.076 | 0.033 | 0.032 |

Table 7. 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

Our results contrast interestingly with other studies. Our finding that the prevalence of ambiguity aversion does not depend on consistency is contrary to Chew et al. (2013), who find that ambiguity aversion is more prevalent among “sophisticated, competent choosers”, and to Stahl (2014), who finds that ambiguity aversion is less prevalent among consistent choosers. Our finding that a substantial share of consistent choosers are ambiguity averse also contrasts with Charness et al. (2013), who find very little ambiguity aversion among consistent choosers.

On the other hand, our conclusion that consistent choosers are much more likely to be ambiguity neutral is entirely in line with Stahl (2014) and Charness et al. (2013). Our finding that ambiguity-seeking subjects are likely to be inconsistent also meshes well with the latter’s report that ambiguity-seekers can be readily persuaded by ambiguity-neutral subjects to abandon their love of ambiguity. (Ambiguity seekers are apparently fickle.)

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 ambiguity easier to recognize also has little effect.
3. As choosers become more consistent, ambiguity neutrality becomes more prevalent and ambiguity seeking disappears. In our experiments, indifference to ambiguity is therefore a sign of a subject who has made up her mind, while ambiguity seeking indicates irresolution.
4. Across all our experiments, subjects’ behaviour 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.

Why do we find less ambiguity sensitivity than some studies? Our conjecture is that Savage’s (1954, p.16) insistence that Bayesian decision theory should only be expected to apply in a small world is relevant here (Binmore 2009). If we accept this restriction on the range of application of Savage’s theory, then it is not surprising that experiments framed in terms of large world problems—like those of finance or macroeconomics—should result in subjects failing to maximize subjective expected utility. In particular, we see no conflict between our results and papers that find widespread evidence of ambiguity sensitivity in experiments that allow subjects a more open interpretation of their laboratory tasks.

However, if the Ellsberg experiment is conceived of as a small-world problem, it should generate rational behaviour that is ambiguity neutral (Raiffa 1961). Indeed, recent
papers that, like ours, find relatively high levels of ambiguity neutrality and low levels of ambiguity aversion all take care to minimize the ambiguity that stems from subjects’ unawareness of the way that gambles are constructed or the experimenter’s intentions (Charness et al. 2013; Stahl 2014). Our findings that subjects are broadly ambiguity neutral (and become more so the more consistent they are) and are unaffected by some variations in framing\textsuperscript{13} are therefore at least to some extent in line with the orthodox theory of rationality.

References

Abdellaoui, M., Vossman, 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.

\textsuperscript{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.
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.

Kocher, M., Lahno, A. and Trautmann, S. (2014). Ambiguity aversion and ambiguity seeking. Working paper of July 2014.

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.

Mangelsdorf, 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).

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

Viscus, 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)

**$R_r$ versus $B_a$ screen:**

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

| 10 Red cards | 20 White or Black cards |
|--------------|-------------------------|
| ![Red cards](image1) | ![Black cards](image2) |
| 33% Red | 67% White or Black |

Please click on your choice of winning card

- Red
- Black

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

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

| 10 Red cards | 20 White or Black cards |
|--------------|-------------------------|
| ![Red cards](image3) | ![Black cards](image4) |
| 33% Red | 67% White or Black |

Please click on your choice of winning cards

- Red & White
- White & Black

Electronic copy available at: https://ssrn.com/abstract=3885126
Version 4 (losses; subjects must infer ambiguous alternative)

**$R_r$ versus $B_a$ screen:**

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

| 10 Red cards | 20 White or Black cards |
|--------------|-------------------------|
| ![Red cards](red_cards.png) | ![White or Black cards](white_black_cards.png) |

**Please click on your choice of winning card**

- Red
- Black

**$R&W_a$ versus $B&W_r$ screen:**

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

| 10 Red cards | 20 White or Black cards |
|--------------|-------------------------|
| ![Red cards](red_cards.png) | ![White or Black cards](white_black_cards.png) |

**Please click on your choice of winning cards**

- Red & White
- White & Black
Version 5 (losses; subjects are told which alternative is ambiguous)

**R, versus B_a screen:**

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

---

10 Red cards

20 White or Black cards

Please click on your choice of winning card

33% Red  From 0% to 67% Black

---

**R&W_a versus B&W_r screen:**

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

---

10 Red cards

20 White or Black cards

Please click on your choice of winning cards

From 33% to 100% Red or White  67% White or Black

Electronic copy available at: https://ssrn.com/abstract=3885126
Version 6 (gains; subjects are told which alternative is ambiguous)

**R_r versus B_a screen:**

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

---

10 Red cards

20 White or Black cards

---

Please click on your choice of winning card

33% Red From 0% to 67% Black

---

**R&W_a versus B&W_r screen:**

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

---

10 Red cards

20 White or Black cards

---

Please click on your choice of winning cards

From 33% to 100% Red or White 67% White or Black

Electronic copy available at: https://ssrn.com/abstract=3885126