

# Ambiguity attitudes, framing, and consistency

Alex Voorhoeve<sup>1</sup>  · Ken Binmore<sup>2</sup> ·  
Arnaldur Stefansson<sup>3</sup> · Lisa Stewart<sup>4</sup>

© The Author(s) 2016. This article is published with open access at Springerlink.com

**Abstract** We use probability-matching variations on Ellsberg's single-urn experiment to assess three questions: (1) How sensitive are ambiguity attitudes to changes from a gain to a loss frame? (2) How sensitive are ambiguity attitudes to making ambiguity easier to recognize? (3) What is the relation between subjects' consistency of choice and the ambiguity attitudes their choices display? Contrary to most other studies, we find that a switch from a gain to a loss frame does not lead to a switch from ambiguity aversion to ambiguity neutrality and/or ambiguity seeking. We also find that making ambiguity easier to recognize has little effect. Finally, we find that while ambiguity aversion does not depend on consistency, other attitudes do: consistent choosers are much more likely to be ambiguity neutral, while ambiguity seeking is much more frequent among highly inconsistent choosers.

**Keywords** Ambiguity attitudes · Framing · Consistency of choice · Ellsberg paradox

**JEL Classification** C91 · D03

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s11238-016-9544-1](https://doi.org/10.1007/s11238-016-9544-1)) contains supplementary material, which is available to authorized users.

---

✉ Alex Voorhoeve  
a.e.voorhoeve@lse.ac.uk

<sup>1</sup> Philosophy, Logic and Scientific Method, LSE, Houghton Street, London WC2A 2AE, UK

<sup>2</sup> Economics, UCL, Gower Street, London WC1E 6BT, UK

<sup>3</sup> Economics, Uppsala University, P.O. Box 513, 751 20 Uppsala, Sweden

<sup>4</sup> Research School of Psychology, The Australian National University, Building 39, Science Road, Canberra, ACT 2601, Australia

## 1 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 Kuylen 2016). 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 hypothesis is supported.

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 that the higher the level of consistency we require, the more ambiguity neutral and the less ambiguity seeking our subjects become. Ambiguity aversion, however, remains moderate and constant across all levels of consistency.

This paper proceeds as follows. Section 2 surveys the literature. Section 3 summarizes our experiments. Section 4 reports the between-subject results on framing. Section 5 reports the results on within-subject consistency and ambiguity attitudes. Section 6 concludes.

## 2 Three questions about ambiguity attitudes

In Ellsberg's (1961) famous single-urn experiment, a ball is randomly drawn from an urn containing 10 red balls and 20 other balls, of which it is known only that each is either white or black. Subjects are asked to choose between betting on red, which yields known probability of winning ( $R_k$ ) and betting on black, which yields an ambiguous probability of winning ( $B_a$ ). They are also asked to choose between a bet in which both white and black win, which yields a known probability of winning ( $W \& B_k$ ) and a bet in which both red and white win, which yields an ambiguous probability of winning ( $R \& W_a$ ). Ellsberg predicted that most people will strictly prefer  $R_k$  to  $B_a$  and  $W \& B_k$  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 gambles (with a known probability of winning) 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  $R_k$  and  $B_a$  and between  $W \& B_k$  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 2016; Charness et al. 2013; Wakker 2010). By contrast, for moderate-to-high probability losses, only a minority find ambiguity aversion.<sup>1</sup> Some studies find ambiguity neutrality;<sup>2</sup> others report a predominance of ambiguity seeking;<sup>3</sup> some find a transition from approximate neutrality to ambiguity seeking as the probability of loss increases.<sup>4</sup> A recent review concludes that “there is clear evidence for an effect of the outcome domain on ambiguity attitude” (Trautmann and Van der Kuijlen 2016, p. 22; cf. Wakker 2010, p. 354).<sup>5</sup>

(ii) *Difficulty in recognizing ambiguity?* In the space of moderate-to-high probability gains, ambiguity aversion is less prevalent in experiments based on Ellsberg’s single-urn setup than in experiments based on Ellsberg’s two-urn setup, 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.<sup>6</sup>

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 risky  $R_k$  over ambiguous  $B_a$  when the known probability of the former is  $1/3$  and choose risky  $W \& B_k$  over ambiguous  $R \& W_a$  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 2016; Kocher et al. 2015). A subject who fails to register ambiguity aversion in the single-urn experiment (because she fails to choose  $R_k$  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  $W \& B_k$  and  $R \& W_a$  than in the single-winning-colour choice between  $R_k$  and  $B_a$ . 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,  $W \& B_k$  is the option with known probabilities

<sup>1</sup> 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.

<sup>2</sup> 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 (2015).

<sup>3</sup> See Abdellaoui et al. (2005), Chakravarty and Roy (2009), and Baillon and Bleichrodt (2015).

<sup>4</sup> See Kahn and Sarin (1988), Viscusi and Chesson (1999), Di Mauro and Maffioletti (2004), and Vieider et al. (2012).

<sup>5</sup> See Viscusi and Chesson (1999) and Bradley (2016) for proposed psychological explanations for such a shift.

<sup>6</sup> 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 66 %; across 7 studies using the single-urn experiment, the average proportion of ambiguity aversion is 27 %. Stahl (2014) also finds lower ambiguity aversion in a single-urn setup.

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.<sup>7</sup> This raises a question about the relationship between subjects' consistency and their ambiguity attitude. This question has not been much studied, and the results so far are conflicting. Two recent studies suggest that ambiguity neutrality is predominant among consistent choosers. If one excludes subjects labelled "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.<sup>8</sup> 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.

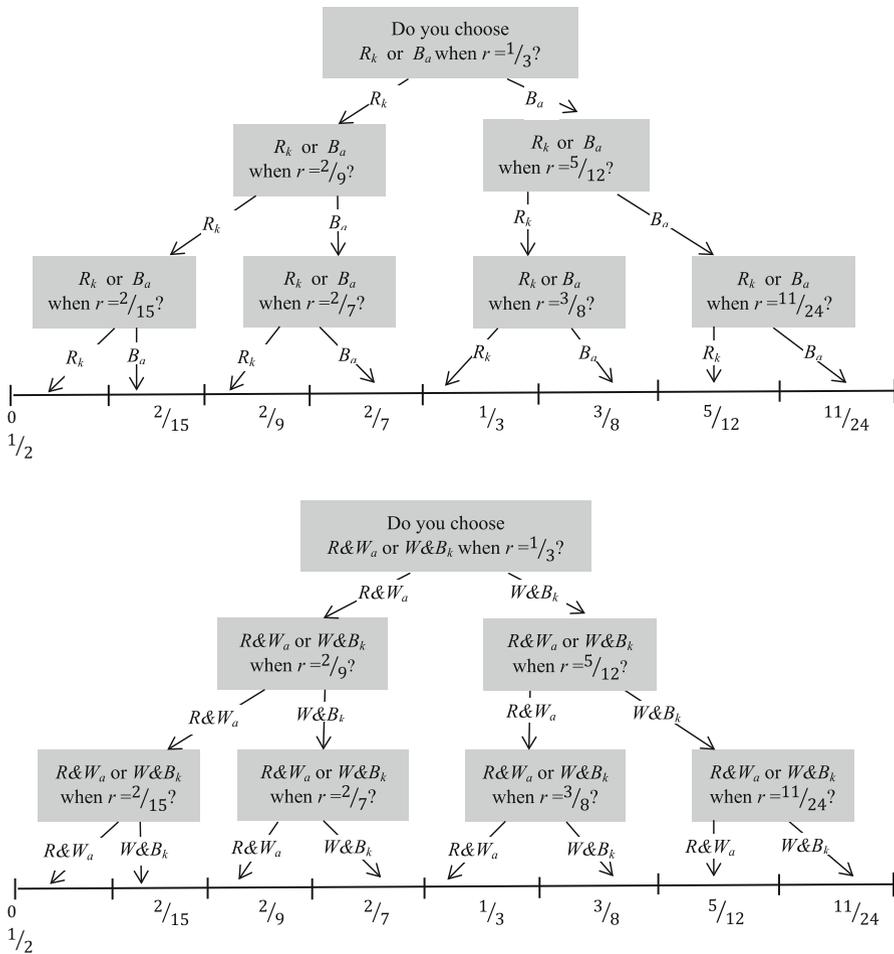
### 3 Experiments

Our experiments use decks of cards rather than urns and employ a titration to estimate, first, the value  $r_1$  of the probability of drawing a red card that makes a particular subject indifferent between  $R_k$  and  $B_a$  and second, the value  $r_2$  of the probability of drawing red that makes the same subject indifferent between  $W \& B_k$  and  $R \& W_a$ .<sup>9</sup> The aim of

<sup>7</sup> 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. \(2015\)](#) find that up to 45 % do not choose according to their previously stated preferences.

<sup>8</sup> 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.

<sup>9</sup> A probability-matching approach is also used in [MacCrimmon and Larsson \(1979\)](#), [Kahn and Sarin \(1988\)](#), and [Viscusi and Magat \(1992\)](#). [Dimmock et al. \(2015\)](#) provide theoretical support for this method.

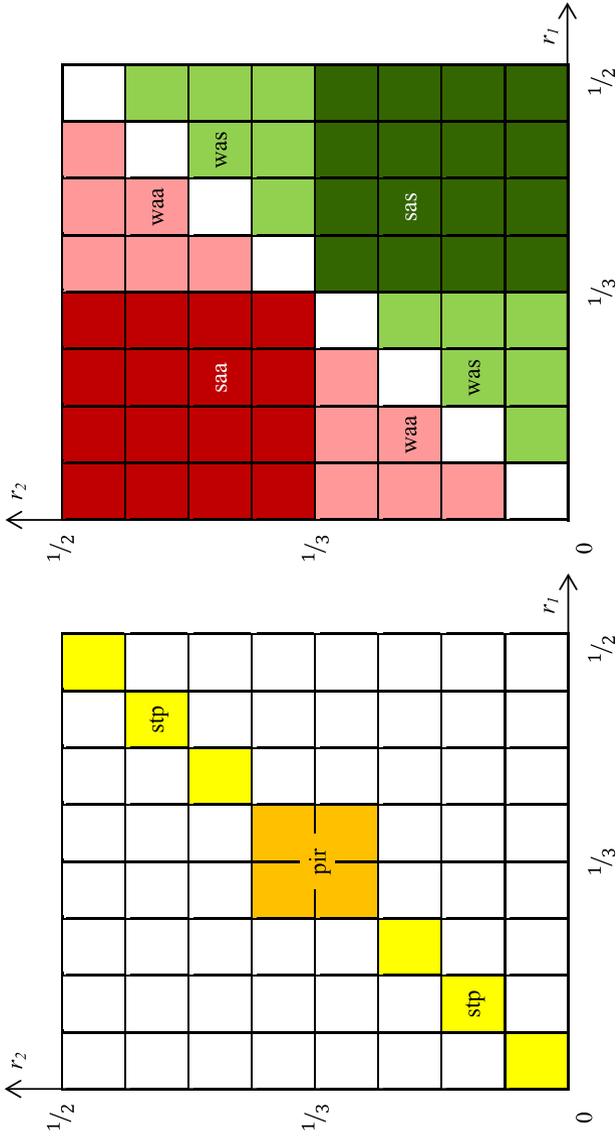


**Fig. 1** The *top tree* shows the questions asked about the subjects’ preferences when the probability of red is  $r$  in order to locate  $r_1$  in one of eight subintervals of  $[0, 1/2]$ . The *bottom tree* is used to so locate  $r_2$ . For example, suppose a subject chooses as follows. In the *top titration*, in sequence,  $R_k; B_a; R_k; B_a$ . This subject is then assigned a value of  $r_1$  satisfying  $2/9 \leq r_1 \leq 2/7$ . In the *bottom titration*, in sequence,  $W \& B_k; R \& W_a; W \& B_k$ . This subject is assigned a value of  $r_2$  satisfying  $3/8 \leq r_2 \leq 5/12$ . This subject’s  $(r_1, r_2)$  would therefore be represented by the cell which is in the third column from the left and the third row from the top in Fig. 2, in the strongly ambiguity averse (saa) region

these titrations is to locate the values of  $r_1$  and  $r_2$  within eight subintervals of  $[0, 1/2]$  using the scheme illustrated in Fig. 1.

Our titrations locate 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.

**Ambiguity neutrality.** We allow for two kinds of ambiguity-neutral behaviour. *Sure-thing principle* (stp): A subject who honours Savage’s (1954, p. 21) sure-thing principle has  $r_1 = r_2$ .



**Fig. 2** Regions corresponding to ambiguity attitudes. Note that the North-West cell of *pir* overlaps with the most South-Easterly cell of *sas*; and the South-East cell of *pir* overlaps with the most North-Westerly cell of *sas*

*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  $r_1 = r_2 = 1/3$ . Since  $1/3$  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  $(1/3, 1/3)$  is counted as lying within pir. This area is therefore a sub-set of stp.<sup>10</sup>

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

*Weak ambiguity sensitivity:* We say that  $r_1 < r_2$  is consistent with weak ambiguity aversion (waa), because it implies that  $R_k$  is preferred to  $B_a$  and  $W \& B_k$  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  $r_1 < 1/3$  and  $r_2 > 1/3$  as consistent with strong ambiguity aversion (saa). We treat pairs  $(r_1, r_2)$  with  $r_1 > 1/3$  and  $r_2 < 1/3$  as consistent with strong ambiguity seeking (sas).

**Common properties of the experiments.** The experiment took place in a room 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  $R_k$  and  $B_a$  and between versions of  $W \& B_k$  and  $R \& W_a$  in ignorance of the mixture of white and black 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 white or black cards, the latter marked on the screen with a “?” on the back. They were then told that the “?” cards could be any mixture of white and black cards, with three subsequent screens inviting them to move the mouse over the “?” cards, revealing three illustrative mixtures.

We sought to allay suspicion of skulduggery (and thereby eliminate one possible source of uncontrolled ambiguity) 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 white or black cards, and counted out the number of red and white or black 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 colour.

After the introduction, the main experiment consisted of four rounds, each with two parts. In part one of the first round, the subjects were navigated through the top tree of Fig. 1 to estimate the interval in which to locate the value  $r_1$  that makes a subject indifferent between  $R_k$  and  $B_a$ . In part two of the first round, subjects were then navigated through the bottom tree in Fig. 1 to estimate the value  $r_2$  that makes a subject indifferent between  $W \& B_k$  and  $R \& W_a$ . (Each round therefore consisted of

<sup>10</sup> In this respect, our definition of stp departs from the definition in Binmore et al. (2012), which included only choices on the diagonal.

six choices.) 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.

**Incentives.** Subjects were informed at the start of the main experiment that of the 24 bets they would face, 2 would be randomly selected to be played for real money at the end of the experiment. This took place as follows. When a given subject finished all their choices, the computer randomly selected two of the bets this subject had faced and displayed them on the screen, listing the composition of each bet in terms of the number of red cards and the number of ambiguous cards and which colour(s) the subject had chosen as their winning colour(s). The subject then came to the front of the room, where the experimenter constructed, with the subject watching, each of these two bets from the displayed boxes of red and ambiguously mixed cards. These two bets were then played for real, using the shuffling machine. Each subject therefore played two tailor-made bets. Subjects left with between £5 ( $\approx$ US\$ 8) and £25 ( $\approx$ US\$ 40).

Two questions arise about our procedure. The first is whether it gives some subjects a reason to misrepresent their preferences in the very first choice, because they would thereby get access to alternatives they regard as more favourable (Harrison 1986). The monetary payoffs associated with each bet were chosen to make such misrepresentation unprofitable for risk-neutral subjects who honour the principle of insufficient reason, but some ambiguity-averse subjects would profit from it. For example, some ambiguity-averse subjects would prefer to incur the expected cost of choosing  $B_a$  against their preference in the first choice in the top tree in Fig. 1 in order to gain access to two subsequent bets with a higher proportion of red cards; an ambiguity-averse subject who chose in this manner would then be classified as ambiguity neutral and/or ambiguity seeking. (For analogous reasons, some ambiguity-seeking subjects would prefer to choose  $R_k$  in the first choice in the top tree in Fig. 1.) However, we believe it is unlikely that such misrepresentation occurred on a significant scale. Subjects were not informed about the titrations they would face and so lacked the knowledge required to exploit the opportunity for misrepresentation. And a subject who chose in line with their true preferences in round 1 would not learn what would have happened if they had chosen differently. It is possible that some inquisitive subjects strayed from their preferred choices in rounds 2 and 3 and made correct conjectures about the decision tree which would then inform their behaviour in round 4. Some subjects who displayed their true ambiguity aversion in round 1 would then have made choices consistent with ambiguity neutrality or ambiguity seeking in round 4. However, as we report in our Online Appendix, we found no evidence of a systematic shift away from ambiguity aversion between round 1 and round 4 across our experiments.

The second question is whether paying subjects for a random selection of their choices allows them to hedge across choices, thereby distorting the representation of their ambiguity attitudes (Oechssler and Roomets 2014; Bade 2015). In our experiment, subjects could not hedge between rounds, because they were informed that each round used new decks. However, as noted in Binmore et al. (2012, p. 228), subjects who knew the decision trees they faced could hedge within the two parts of a given round by choosing  $B_a$  when offered the choice between  $R_k$  and  $B_a$ , and choosing  $R \& W_a$  when offered the choice between  $W \& B_k$  and  $R \& W_a$ . Since each choice is equally likely to be played for real, this is equivalent to turning down an equiprobable lottery between  $R_k$  and  $W \& B_k$  in favour of an equiprobable lottery between  $B_a$  and

$R&W_a$ . The latter has a probability  $1/2$  of winning. No appeal to the principle of insufficient reason is then necessary to justify playing according to its tenets.

We believe it is unlikely that subjects employed such a hedging strategy. In the first round, subjects lacked the requisite knowledge of the decision trees they faced and, as mentioned, we found no evidence of a systematic decrease in ambiguity aversion in subsequent rounds. Moreover, the strategy is rather complex (it involves thinking several choices ahead and matching one's behaviour in a one-winning-colour choice with one's later behaviour in a corresponding two-winning colour choice). Indeed, in Binmore et al. (2012), we compared a version of our experiments which allowed a simpler form of hedging with one in which this opportunity was eliminated, and our statistical tests were not able to distinguish the two data sets.

**Particular frames.** In Binmore et al. (2012), we reported the effects of variations in the presentation of alternatives in three earlier versions of this experiment. Those 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 the third version from that paper, which we here refer to as Version GI<sub>3</sub> (denoting that it uses a gain framing, that ambiguity was implicit and that it was the third version of our previous experiment) and revised it to study the effects of switching to a loss frame and of making explicit any ambiguity present in an alternative.

Version LI represented a loss frame in which ambiguity remained implicit. Upon entering, subjects were given 25 plastic £1 coins. 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 20 of these coins. Every screen that presented a "for real" choice informed them that they would *keep* their money if they won and that they would *lose* their money otherwise.

Version LE was identical to Version LI, 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 option 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 WHITE & BLACK, your chance of winning is 71 % (because the number of cards that are WHITE or BLACK 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  $W&B_k$  and  $R&W_a$ , subjects were informed that if they chose to bet on  $W&B_k$ , then their chance of winning was 67 %, while if they chose to bet on  $R&W_a$ , then their chance of winning was anywhere from 33 to 100 %.

Version GE reverted to the gain frame, but made explicit in the manner just outlined which alternative was ambiguous and how ambiguous it was. Our Online Appendix provides the introductions and two sample screens for each version.

The different versions were performed consecutively, so subjects were not randomized across versions. (Subjects who had enrolled in an earlier version could not enrol in a subsequent one.) Version GI<sub>3</sub> 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 advertise to and recruit from the same pool, primarily of central London students. Experiments carried out at UCL had a higher proportion of students (90 %) than those carried out at LSE (78 %). We therefore confirmed that our central conclusions from Sect. 4 were robust to excluding all non-students. Further demographic data and analysis, including this robustness analysis, are available in our Online Appendix.

## 4 Results on framing

There was no evidence that subjects systematically 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 Fig. 3. The results sorted by ambiguity attitudes appear in Table 1.

Across all versions, roughly 45 % of subjects' choices are consistent with weak ambiguity aversion (waa) and roughly 30 % are consistent with strong ambiguity aversion (saa). Choice in line with ambiguity neutrality is also common, but more variable: the share of choices consistent with the sure-thing principle (stp) varies from 36 to 52 % and with the principle of insufficient reason from 24 to 38 %. Ambiguity seeking is least common in all versions. The share of subjects' choices that is consistent with weak ambiguity seeking (was) is quite constant at around 21 %, but the share in line with strong ambiguity seeking (sas) shows more variability, fluctuating from 7 to 16 %.

The hypotheses under consideration regarding framing are:

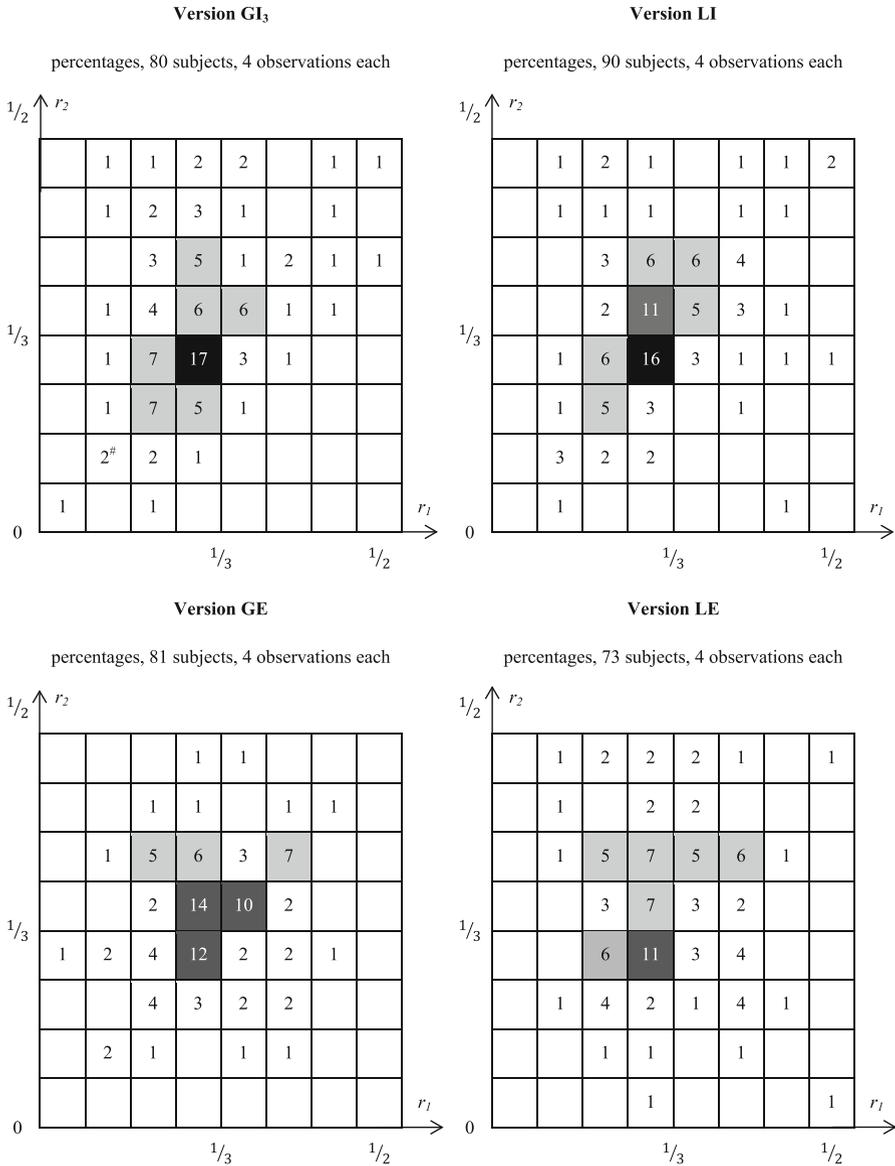
- (i) When compared to gains, for losses there will be less ambiguity aversion and substantially more ambiguity-neutral or ambiguity-seeking behaviour.
- (ii) Making the ambiguous option clearer will magnify the response to ambiguity; it will generate more ambiguity-averse behaviour for gains and more ambiguity-neutral and/or ambiguity-seeking behaviour for losses.

We employ three complementary tests of these hypotheses: a test at the level of each attitude taken separately (Sect. 4.1); a version of the Kolmogorov–Smirnov test applied to the distribution across all 64 squares (Sect. 4.2); and a model of the data (Sect. 4.3).

### 4.1 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 across all versions under the null hypothesis that observations from each subject



**Fig. 3** Aggregate results by frame. *Shaded squares* indicate a particularly high concentration of responses. *Hash symbol* This entry was omitted in error in Binmore et al. (2012, p. 226). (The analysis in that paper was done with the correct data.)

are 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.

**Table 1** Percentage shares of ambiguity attitudes in different versions

		GI <sub>3</sub>	LI	GE	LE	Diff. a	G	L	Diff. b	I	E	Diff. c
Averse	waa	43	44	44	49		43	46	0.330	43	46	0.327
	saa	28	27	31	31		30	29	0.838	28	31	0.227
Neutral	stp	46	51	52	36	0.043** (GE & LE) 0.048** (LI & LE)	49	44	0.087*	49	44	0.103
	pir	32	36	38	24	0.055* (GE & LE) 0.092* (LI & LE)	35	30	0.070*	34	31	0.258
Seeking	was	20	19	21	25		20	22	0.443	20	23	0.161
	sas	7	8	12	16	0.084* (GI <sub>3</sub> & LE)	9	12	0.201	8	14	0.000***

Entries in columns GI<sub>3</sub>; LI; GE; LE; G; L; I; and E report fractions of the population in each version (or combination of versions) consistent with each of the ambiguity attitudes we distinguish. Note that the totals of waa, stp and was do not add to 100 % because one cell in our 8 × 8 grid is consistent with both waa and stp, and one cell is consistent with both stp and was

**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. **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 (G) or a loss (L) frame. **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 (I) or explicit (E)

\* Taking the null hypothesis to be that the prevalence of each attitude is independent of the version of the experiment, the difference between versions is significant at the 10 % level

\*\* Difference is significant at the 5 % level

\*\*\* Difference is significant at the 1 % level

- 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.
- 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 Online Appendix).

**Gains versus losses:** The mere switch from gains to losses while leaving subjects to infer the ambiguous option in every choice (the switch from GI<sub>3</sub> to LI) makes no significant difference for any ambiguity attitude. The mere switch from gains to losses while making the ambiguous option explicit (from GE to LE) decreases behaviour

consistent with ambiguity neutrality (the decrease in choices in line with *stp* is significant at the 5 % level; in choices in line with *pir* at the 10 % level), but has no significant effect on other ambiguity attitudes, because the decrease in choices consistent with ambiguity neutrality is “spread out” among other ambiguity attitudes. Comparing the aggregate of both gain versions with the aggregate of both loss versions (G to L), we find only a modest decrease in ambiguity-neutral behaviour, which is significant only at the 10 % level. In sum, our results are not in line with the common finding embodied in hypothesis (i) that changing from gains to losses leads to a reduction in ambiguity aversion and an increase in ambiguity-neutral and/or ambiguity-seeking behaviour.

**Making ambiguity easier to recognize:** Making the ambiguous option explicit in every choice has no effect in a gain frame (GI<sub>3</sub> versus GE). Our experiments therefore do not support hypothesis (ii)’s claim that making ambiguity explicit would yield greater ambiguity aversion for gains. In a loss frame (LI versus LE), there is a significant decrease in ambiguity-neutral behaviour (at the 5 % level for *stp* and the 10 % level for *pir*). There is also the hypothesized increase in ambiguity-loving behaviour, but the latter is too small to be significant.<sup>11</sup> Our experiments therefore do not offer substantial support for hypothesis (ii)’s claim that making ambiguity explicit in a loss frame would increase ambiguity neutrality and/or ambiguity seeking.

We also compared the aggregate of versions that leave ambiguity implicit with the aggregate of versions that make it explicit (I versus E). 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. There is no significant difference, except that the “explicit” aggregate has a higher share of strong ambiguity seeking (significant at the 1 % level).

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

## 4.2 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 are 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

<sup>11</sup> Excluding all non-students makes the difference in strong ambiguity seeking between LI and LE significant at the 10 % level. See our Online Appendix.

all possible orderings of the data points. Such a procedure might be appropriate when the data are unstructured, but we exploit the underlying structure of our problem by applying the one-dimensional K-S test separately to three distributions of interest in our  $8 \times 8$  grids of Fig. 2:

- (i) the sums of the columns (this checks for differences in the distribution of  $r_1$ );
- (ii) the sums of the rows (this checks for differences in the distribution of  $r_2$ );
- (iii) the sums of the parallels to the diagonal  $r_1 = r_2$ , moving from North-West to South-East (since ambiguity-averse behaviour lies North-West of this diagonal and ambiguity-seeking behaviour lies to its South-East, this checks for differences in the distribution of ambiguity attitudes).

For each comparison, Table 2 reports the largest of these three 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 switch from gains to losses while leaving ambiguity implicit (from  $GI_3$  to LI) makes no significant difference. The switch from gains to losses while making ambiguity explicit (from GE to LE) equally has no significant effect. Nor does

**Table 2** Significantly different distributions?

	LI	GE	LE		L	I
$GI_3$	0.07	0.13	0.14	G	0.08	
LI		0.06	0.11	E		0.10
GE			0.09			

For each comparison between versions, we compute three K-S statistics comparing, respectively: (i) the distribution of  $r_1$ ; (ii) the distribution of  $r_2$ ; (iii) the distribution of the sums of the parallels to the  $r_1 = r_2$  diagonal in our  $8 \times 8$  chessboard, going from NW to SE. 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.)

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 (GI<sub>3</sub> versus GE) or a loss frame (LI versus LE). Nor does the K-S test reveal any significant difference between the “implicit” and “explicit” groupings (I versus E).

In sum, the K-S test does not permit us to say with confidence of any version that it generates different behaviour from any other version. Just as the analysis of Sect. 4.1, it therefore offers no support for either of our two hypotheses.

### 4.3 Modelling the data

Our third test involves fitting the econometric model from [Binmore et al. \(2012\)](#) to each version. The model assumes that subjects enter our titration with a “baseline preference” in line with the sure-thing principle, but can diverge from this preference when answering the questions in our titration in either an ambiguity-averse or an ambiguity-loving direction.

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 a subject faces a choice in our titration between an option that is consistent with the subject’s ambiguity-neutral baseline preference and an option that is inconsistent with their baseline preference but consistent with ambiguity-seeking behaviour, the model assumes that the subject chooses in line with their baseline preference with probability  $a$ , where  $a \leq 1$ . If  $a$  were 1, then no subject would display ambiguity-seeking behaviour; if  $a$  were 0, then subjects would diverge from the sure-thing principle in an ambiguity-seeking direction whenever they had an opportunity to do so. We therefore take  $1 - a$  as a measure of ambiguity seeking.

When a subject faces a choice between an option that is consistent with their ambiguity-neutral baseline preference and an option that is inconsistent with their baseline preference but consistent with ambiguity-averse behaviour, we assume that the subject responds in line with their baseline preference with probability  $d$ , where  $d \leq 1$ . If  $d$  were 1, then no subject would display ambiguity-averse behaviour; if  $d$  were 0, then subjects would diverge from the sure-thing principle in an ambiguity-averse direction at every opportunity. We therefore take  $1 - d$  as a measure of ambiguity aversion. Observe that, 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 therefore take  $a - d$  as a measure of the degree to which ambiguity aversion is more prevalent than ambiguity seeking.

We therefore have a model with four parameters:  $\mu$ ,  $\sigma$ ,  $a$ ,  $d$ . 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 cannot confidently be rejected as a representation of the distribution of ambiguity attitudes in the data, a test of the effects of framing is this: does the best-fitting model for version

**Table 3** The best-fitting models

Model parameters	GI <sub>3</sub>	LI	GE	LE	G	L
$1 - a$ (likelihood of diverging from stp in an ambiguity-seeking direction when given the opportunity)	0.09	0.11	0.12	0.10	0.11	0.11
$1 - d$ (likelihood of diverging from stp in an ambiguity-averse direction when given the opportunity)	0.22	0.20	0.21	0.20	0.23	0.21
$a - d$ (indicator of the prevalence of ambiguity aversion over seeking)	0.13	0.09	0.09	0.10	0.12	0.10
$\mu$ (mean of the distribution along the $r_1 = r_2$ diagonal)	0.31	0.32	0.33	0.33	0.32	0.32
$\sigma$ (standard deviation)	0.035	0.045	0.056	0.060	0.045	0.051
<b>Fit</b>						
$K_I$ ( $r_1 = r_2$ diagonal)	0.06	0.08	0.05	0.10	0.05	0.09
$K_{II}$ (parallels to $r_1 = r_2$ diagonal)	0.05	0.03	0.02	0.07	0.02	0.04

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). These 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

*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  $K_{II}$ , for a comparison between the model and each version (or combination of versions). We use the observed data on the  $r_1 = r_2$  diagonal for  $K_I$  and the sums of parallels to this diagonal for  $K_{II}$ . (As mentioned above, the latter is an indicator of the distribution of ambiguity attitudes.) Low values of  $K_I$  indicate that the observed data points on the  $r_1 = r_2$  diagonal of our chessboard are consistent with our model. Low values of  $K_{II}$  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 cannot be confidently rejected as representations of the distribution of ambiguity attitudes in 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 throughout. 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.

#### 4.4 Conclusions with regard to framing

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

*Hypothesis (i) There will be less ambiguity-averse behaviour and 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 do not support 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 substantial effect. This does not support 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 offer no support for hypothesis (ii) for losses. We note that the failure to find support for this hypothesis in the gain frame means that we lack an explanation of the finding, reported in Sect. 2, that ambiguity aversion appears less prevalent in single-urn experiments.

### 5 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 start by defining a “representative pair” of values  $(r_1, r_2)$  for subject  $n$  as a pair that minimizes the total number of horizontal and vertical “moves”  $M_n$  in our checkerboard between this representative pair and the subject’s four estimated pairs of  $(r_1, r_2)$ . We then take  $M_n$  to be our measure of inconsistency. For a completely consistent subject,  $M_n$  is zero; the more inconsistent a subject, the higher the value of  $M_n$ . More precisely, we define the “distance” from square  $(g, h)$  to square  $(i, j)$  by  $d = |g - i| + |h - j|$ . Given our four data points  $(i_1, j_1), (i_2, j_2), (i_3, j_3), (i_4, j_4)$  for individual  $n$ :

$$m(g, h) = \sum_{s=1}^4 \{|g - i_s| + |h - j_s|\}$$

**Table 4** The percentage share of subjects for each level of inconsistency, for each version

$M_n$	Version					Aggregate	Inconsistency
	GI <sub>2</sub>	GI <sub>3</sub>	LI	GE	LE		
0	9	3	16	11	6	9	} Low: 29% ( $N = 113$ )
1	4	15	9	20	6	11	
2	12	9	9	5	8	9	
3	17	15	19	11	10	15	} Moderate: 44% ( $N = 177$ )
4	13	16	18	10	26	16	
5	21	11	9	17	8	13	
6	15	8	8	7	11	10	} High: 28% ( $N = 110$ )
7	4	6	1	6	6	4	
8	3	8	2	4	8	5	
9	0	4	3	4	4	3	
$\geq 10$	3	6	7	5	8	6	
Median $M_n$	4	4	3	4	4	4	
$N$	76	80	90	81	83	400	

$M_n$  is our measure of inconsistency; 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 GI<sub>2</sub> through GE give the percentage of the population for each version for each level of inconsistency; the next column reports the percentages in the aggregate of these versions

for each  $(g, h)$ . The measure  $M_n$  of inconsistency for each individual is then

$$M_n = \min m(g, h)$$

taken over all  $(g, h)$ .<sup>12</sup>

Table 4 reports the percentage share of subjects at each level of inconsistency 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 permit. We therefore employ the K-S test outlined in Sect. 4.2 for all six versions of our experiment (the four reported here and two earlier ones reported in Binmore et al. 2012, denoted as GI<sub>1</sub> and GI<sub>2</sub>) as follows:

1. We exclude a version from our aggregated data if the largest of our three K-S statistics indicates that it is different at the 10 % significance level from any other version. This excludes only GI<sub>1</sub>.
2. For the remaining versions, we assess whether the K-S test prohibits adding various salient combinations of versions to other such combinations, e.g. we compare “all gain versions” (GI<sub>2</sub>, GI<sub>3</sub>, and GE) to “all loss versions” (LI and LE). No version or combination of versions is excluded by this test (see our Online Appendix).

The upshot is that this procedure does not rule out aggregating versions GI<sub>2</sub>, GI<sub>3</sub>, LI, GE, and LE of our experiment. However, since we saw in Sect. 4.1 that Version

<sup>12</sup> There will often be more than one representative pair  $(g, h)$  that achieves the minimum  $M_n$ . Whenever it matters, we take the average of these values.

**Table 5** Ambiguity attitude by level of inconsistency

		Low inconsistency	Moderate inconsistency	High inconsistency	Low-High Difference Significant?
<b>Averse</b>	waa	45	44	43	0.967
	saa	33	27	28	0.724
<b>Neutral</b>	stp	63	45	34	0.001***
	pir	48	33	21	0.002***
<b>Seeking</b>	was	11	22	33	0.004***
	sas	4	10	19	0.007***

The first three columns show the fractions of the population from the low, moderate and high inconsistency groups compatible with each of our six ambiguity attitudes. The fourth column is the probability (*p* value) of seeing the observed or greater difference between the low and high fractions under the null hypothesis that the degree of inconsistency does not influence ambiguity attitudes

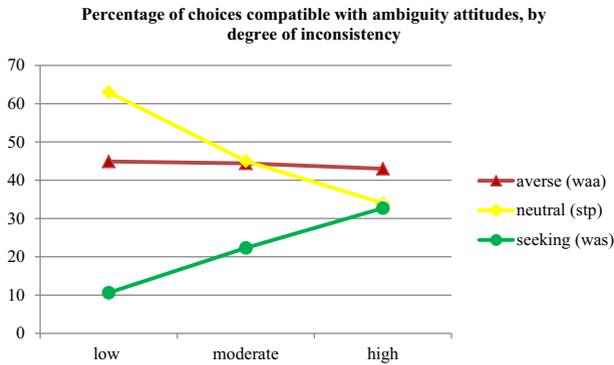
\*\*\* Low-high difference between groups is significant at the 1 % level

LE is most different from our other versions, we confirmed that our results regarding the relationship between consistency and ambiguity attitude are robust to removing LE from our data. (See our Online Appendix.)

Next, we classify subjects into three groups: low inconsistency ( $0 \leq M_n \leq 2$ ); moderate inconsistency ( $3 \leq M_n \leq 5$ ); and high inconsistency ( $6 \leq M_n$ ). We employ this classification because it both brings together subjects who are plausibly similar in terms of inconsistency of choice and creates groupings that are similar in size to ensure statistical power. On this classification, 29 % of our subjects display low inconsistency (are consistent); 44 % are moderately inconsistent; and 28 % are highly inconsistent. 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).<sup>13</sup>

So defined, how does consistency relate to ambiguity attitude? Table 5 and Fig. 4 give our results. Three things are apparent. First, the prevalence of ambiguity aversion is nearly identical in all three groups. Second, ambiguity-neutral behaviour is roughly twice as prevalent among consistent (low inconsistency) subjects as it is among highly inconsistent subjects, with moderately inconsistent subjects in the middle. Third, ambiguity seeking increases as subjects become more inconsistent: it is 3–4 times more prevalent among highly inconsistent subjects than among consistent subjects, with moderately inconsistent subjects again in the middle.

<sup>13</sup> See Charness et al. (2013), Duersch et al. (2013), Dimmock et al. (2015), and Stahl (2014).



**Fig. 4** Ambiguity attitudes and inconsistency. The figure shows how the prevalence of choices compatible with, respectively, weak ambiguity aversion (waa); ambiguity neutrality (stp); and weak ambiguity seeking (sas) relates to subjects' degree of inconsistency. (Recall that the totals of waa, stp, and was need not add to 100 % because one cell in our  $8 \times 8$  grid is consistent with both waa and stp, and one cell is consistent with both stp and was.) Consistent individuals (low inconsistency) are much more likely to be ambiguity neutral; highly inconsistent individuals are much more likely to be ambiguity seeking. (As Table 5 demonstrates, the same pattern exists for saa, pir, and sas.)

To establish whether these differences are significant, we employ tests similar to those of Sect. 4: a test for the prevalence of each ambiguity attitude (Sect. 5.1); the Kolmogorov–Smirnov test for differences between the distributions for each consistency group over our  $64$ -square grid (Sect. 5.2); and a test involving the best-fitting model (Sect. 5.3).

### 5.1 Is the prevalence of any attitude significantly different?

The fact that the fraction of observations satisfying stp decreases markedly with our measure of inconsistency leads us to compute the probability ( $p$  value) of seeing the observed or a greater difference between the low and high fractions under the null hypothesis that inconsistency does not affect ambiguity attitudes. In particular, the deviations in the low and high fractions from the moderate fraction are assumed independent. We condition this probability on the observed ordering of the fractions because this ordering was instrumental in our choice of 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.

### 5.2 Kolmogorov–Smirnov test

To assess whether the distributions are different, we employ the K-S test as outlined in Sect. 4.2. Table 6 reports the results. While the K-S test does not allow us to conclude that consistent individuals (low inconsistency) and moderately inconsistent individuals choose differently, we can be very confident that consistent individuals choose differently from highly inconsistent individuals. We can also be reasonably

**Table 6** Are the distributions of individuals displaying a low, moderate, and high degree of inconsistency significantly different?

	Moderate	High
Low	0.12	0.23***
Moderate		0.16*

For each comparison, we compute three K-S statistics comparing, respectively: (i) the distribution of  $r_1$ ; (ii) the distribution of  $r_2$ ; (iii) the distribution of the sums of the parallels to the  $r_1 = r_2$  diagonal in our  $8 \times 8$  chessboard, going from NW to SE. We report the largest of these. We can be reasonably confident that moderate and high are different distributions; we can be very confident that low and high are different  
 \*\* Less than 10 % chance of wrongly rejecting the hypothesis that the distributions are drawn from the same underlying distribution  
 \*\*\* Less than 1 % chance of wrongly rejecting the hypothesis that the distributions are drawn from the same underlying distribution

confident that the moderately inconsistent group differs from the highly inconsistent group. This is in line with the results reported in Table 5.

### 5.3 Modelling the data

Finally, we find the best-fitting parameter values for the model outlined in Sect. 4.3. Table 7 reports these parameters. None of these models can be confidently rejected as a representation of our data, according to our K-S tests; moreover, the parameter range within which this result can be maintained is small. Our key finding is that highly inconsistent individuals are most likely to diverge from the sure-thing principle because ambiguity seeking is much more common among them. In addition, moderately inconsistent and consistent choosers are quite similar. This analysis therefore confirms the results of the previous two tests.

### 5.4 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 choose in line with ambiguity aversion also contrasts with Charness et al. (2013), who find very little ambiguity aversion among consistent choosers.

On the other hand, our conclusions that consistent choosers are much more likely to be ambiguity neutral and, indeed, that behaviour in line with ambiguity neutrality

**Table 7** Best-fitting models

Model parameters	Degree of inconsistency		
	Low	Moderate	High
$1 - a$ (likelihood of diverging from stp in an ambiguity-seeking direction when given the opportunity)	0.08	0.12	0.24
$1 - d$ (likelihood of diverging from stp in an ambiguity-averse direction when given the opportunity)	0.22	0.22	0.28
$a - d$ (indicator of the prevalence of ambiguity aversion over seeking)	0.14	0.10	0.04
$\mu$ (mean of the distribution along the $r_1 = r_2$ diagonal)	0.32	0.32	0.35
$\sigma$ (standard deviation)	0.032	0.033	0.076

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

is dominant among such subjects are entirely in line with [Stahl \(2014\)](#) and [Charness et al. \(2013\)](#). Our finding that ambiguity-seeking subjects are likely to be highly 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.)

## 6 Concluding discussion

We assessed both the sensitivity of ambiguity attitudes to framing and the relationship between these attitudes and the consistency with which subjects choose within a single frame. We summarize our key findings, along with some tentative explanations.

1. Contrary to the predominant (though not universal) finding in the literature, a switch from a gain to a loss frame does not generate a shift from ambiguity aversion towards ambiguity neutrality or ambiguity seeking.
2. Making ambiguity easier to recognize also has little effect.
3. As choosers become more consistent, ambiguity neutrality comes to dominate behaviour 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, the distribution of ambiguity attitudes 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 depart less frequently from ambiguity neutrality, because they very rarely do so in an ambiguity-seeking direction.

What might explain the difference between our finding that ambiguity aversion and neutrality together predominate even in a loss frame and studies that report a predominance of ambiguity seeking for losses? Our separate result that ambiguity seeking is most common among inconsistent choosers, as well as [Charness et al. \(2013\)](#)'s finding that ambiguity seekers can easily be persuaded to change their mind, suggest the following hypothesis. A higher proportion of ambiguity seeking in a loss frame may indicate a higher proportion of subjects who are unsure how to choose and are simply trying something out without conviction or choosing on a whim. In future work on ambiguity attitudes for losses, it would therefore be worth examining the steadfastness of subjects' attitudes. The fact that we do not find much of a shift towards ambiguity seeking for losses may be due to the fact that, in each of our experiments, the decision situation was comparatively transparent to most subjects, which is a point we elaborate on next.

The dominant role of ambiguity neutrality among consistent choosers in our experiments is broadly in line with findings reported in [Charness et al. \(2013\)](#) and [Stahl \(2014\)](#). These papers take care, as we do, to minimize the ambiguity that stems from subjects' doubts about the experimenter's intentions or from their unawareness of the way that gambles are constructed (for the importance of these factors, see [Pulford 2009](#)). We conjecture that this plays a part in generating the relatively high levels of ambiguity-neutral behaviour observed compared to the literature as a whole (especially among consistent choosers), because it enables more subjects to represent the decision situation they face as what [Savage \(1954, p. 16\)](#) called a "small world". A small world is a decision situation in which: (i) the subject knows every possible aspect of the situation, or state of the world, which is unaffected by their choice and which, together with their choice, will determine the consequences they want to take into account; and (ii) the subject knows the relevant consequences of every possible act in every such state of the world. If the Ellsberg experiment is conceived of as a small-world problem, it should generate rational behaviour that is ambiguity neutral ([Raiffa 1961](#); [Al-Najjar and Weinstein 2009](#)). Our finding that the most consistent subjects are predominantly ambiguity neutral is therefore at least to some extent in line with the orthodox theory of rationality. But [Savage](#) stressed that one should not expect Bayesian decision theory to apply outside of a small world (see also [Binmore 2009](#)). We therefore see no conflict between our results and reports of greater ambiguity sensitivity in experiments in which subjects cannot readily represent the decision situation as a small world.

In sum, we find that ambiguity attitudes are little affected by gain/loss framing, but correlate markedly with subjects' consistency of choice. Two hypotheses that emerge as worth further investigation are that ambiguity seeking is a sign of subjects who have not made up their mind, and that the more transparent and controlled the source of ambiguity, the more ambiguity neutrality will be observed.

**Acknowledgments** This paper was presented at the LSE, UCL, Uppsala University, and Warwick University. We thank those present as well as Richard Bradley, Antonio Cabrales, Martin Cripps, Casey Helgeson, Stefan Trautmann, Peter Wakker, an editor and a referee for this journal for comments. We gratefully acknowledge support from the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD), the British Arts and Humanities Research Council (AHRC) 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 its facilities.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## 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.
- Al-Najjar, N., & Weinstein, J. (2009). The ambiguity aversion literature: A critical assessment. *Economics and Philosophy*, *25*, 249–284.
- Bade, S. (2015). Randomization devices and the elicitation of ambiguity-averse preferences. *Journal of Economic Theory*, *159*, 221–235.
- Baillon, A., & Bleichrodt, H. (2015). Testing ambiguity models through the measurement of probabilities for gains and losses. *American Economic Journal: Micro*, *7*(2), 77–100.
- 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. (2016). Ellsberg's paradox and the value of chances. *Economics and Philosophy* (in press).
- 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 Sep 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.
- Di Mauro, C., & Maffioletti, A. (2004). Attitudes to risk and attitudes to uncertainty: experimental evidence. *Applied Economics*, *36*, 357–372.
- Dimmock, S., Kouwenberg, R., & Wakker, P. (2015). Ambiguity attitudes in a large representative sample. *Management Science* (in press).
- 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*, 225–250.
- 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.
- Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, *75*, 503–536.

- Harrison, G. (1986). An experimental test for risk aversion. *Economics Letters*, *21*, 7–11.
- 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., Trautmann, S. (2015). Ambiguity aversion is the exception. CESifo Working Paper Series Number 5261. Accessed Jan 2016.
- Lopes, R., Reid, I., & Hobbes, P. (2007). The two-dimensional Kolmogorov–Smirnov test. In *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: Riedel.
- Mangelsdorff, L., & Weber, M. (1994). Testing Choquet expected utility. *Journal of Economic Behavior and Organization*, *25*, 437–457.
- Oechssler, J., & Roomets, A. (2014). Unintended hedging in ambiguity experiments. *Economics Letters*, *122*, 243–246.
- Pulford, B. (2009). Is luck on my side? Optimism, pessimism, and ambiguity aversion. *Quarterly Journal of Experimental Psychology*, *62*, 1079–1087.
- 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. (2016). Ambiguity attitudes. In G. Keren & G. Wu (Eds.), *Blackwell handbook of judgment and decision-making* (pp. 89–116). Oxford: Blackwell.
- Trautmann, S., & Wakker, P. (2015). Making the Anscombe–Aumann approach to ambiguity suited for descriptive applications. Working paper on <http://people.few.eur.nl/wakker/pdf/aadescr.pdf>. Accessed Jan 2016.
- Vieider, 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](http://www.ferdinandvieider.com/stake_effects_ambiguity.pdf). Accessed Sep 2014.
- Viscusi, W. K., & Chesson, H. (1999). Hopes and fears: The conflicting effects of 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.