Theory and Decision (2006) 60: 257–282
DOI 10.1007/s11238-005-4573-1

MICHAEL BLEANEY and STEVEN J. HUMPHREY

AN EXPERIMENTAL TEST OF GENERALIZED AMBIGUITY AVERSION USING LOTTERY PRICING TASKS

ABSTRACT. We report the results of an experiment which investigates the impact of the manner in which likelihood information is presented to decision-makers on valuations assigned to lotteries. We find that subjects who observe representative sequences of outcomes attach higher valuations to lotteries than those who are given only a verbal description of a probability distribution. We interpret this in terms of a reduction in ambiguity about the possible lottery outcomes. These findings suggest that ambiguity aversion may be a confounding factor in reported experimental violations of expected utility theory based on verbal descriptions of probability distributions.

KEY WORDS: ambiguity, probability learning, competence and comprehension hypotheses, experiment, lottery valuations

JEL CLASSIFICATION: D81, C91

1. INTRODUCTION

Since Ellsberg's (1961) pioneering work there has been a significant quantity of empirical and theoretical research aimed at increasing understanding of the impact of ambiguity on individual choice (Camerer and Weber (1992) provide a review). Although there is general agreement about the existence of ambiguity aversion, a notable feature of this research is the many different definitions of ambiguity it has produced. For example, Ellsberg (1961, p. 647) defines ambiguity as the "quality, type, reliability, and 'unanimity' of information giving rise to one's degree of confidence in an estimate of relative likelihoods". Einhorn and Hogarth (1986, p. 227) refer to ambiguity as "... uncertainty about uncertainties ..." Camerer (1995, p. 645) prefers, "... ambiguity is *known-to-be-missing information*, or not knowing relevant information that could be known." Irrespective of which definition is preferred, tests of the influence of ambiguity on choice have typically used a variant of Ellsberg's (1961) two-urn problem. In one urn there are 50 red and 50 black balls, and in the other urn there are a total of 100 red and black balls in unknown composition. A ball is randomly drawn from each urn in turn, and the decision-maker is paid either £x or nothing, depending on whether the colour of the ball drawn is correctly predicted. In this task individuals prefer to bet on the outcome of the draw from the urn of known composition—a result that is usually interpreted as ambiguity aversion. As outlined by Camerer (1995, p. 646), it is ambiguity aversion which causes a gap between individuals' beliefs about the likelihood of an event and their willingness to bet on it. The observed ambiguity aversion implies that probability assessments of black and red are greater in the first urn than the second and so cannot sum to unity for both urns. This is a violation of expected utility theory.

Ambiguity aversion is a potentially important confounding factor in experimental tests of expected utility theory. For example, if the tests involve a comparison of a simple and a complex lottery of equal expected utility, subjects may prefer the former simply because their understanding of it is clearer. Most experimental tests are based on the assumption that a mathematically complete description of a probability distribution is sufficient to engender understanding. In fact subjects may not regard such a description as fully informative—they may still feel uncertain about the practical interpretation of statements such as "there is a probability of 0.2 of outcome A and a probability of 0.8 of outcome B." Indeed this is strongly suggested by some recent research in psychology that is discussed below. Such uncertainty is likely to engender ambiguity aversion in the valuation of the lottery. There may be an element of a sense of incompetence in such ambiguity aversion, as suggested by Fox and Tversky (1995).

The psychological research indicates that people have a more accurate understanding of likelihood information that is presented in a frequency format (i.e. as a representative sequence of outcomes e.g. A, B, B, B, A, B, ...) rather than as a statement of probabilities. If that is the case, the outcomes of many experiments—particularly those involving a choice between a certainty and a lottery—might be affected by alternative forms of presentation of the same likelihood information.

This paper reports the results of an experiment designed to test the hypothesis that verbal descriptions of probability distributions are vulnerable to ambiguity aversion. In one treatment subjects value lotteries on the basis of information governing outcomes presented in the conventional manner—a concise but mathematically complete description of the probability distribution. In a second treatment subjects perform the same task, but are additionally presented with a sequence of ten representative resolutions of the risk in each lottery prior to valuing those lotteries.[1] Our hypothesis is that the latter method of presenting likelihood information enhances individuals' understanding of that information and thereby reduces ambiguity, resulting in higher valuations of the lotteries. We motivate our hypothesis by proposing a new generalized definition of ambiguity that encompasses deficiencies in the understanding of information. As we discuss below, if support can be found for our hypothesis, it would suggest that lotteries may be ambiguous even when mathematically complete likelihood information is provided. This would suggest that individuals do not view a complete, but merely probabilistic, description of lotteries as fully informative. In the next section we discuss the most important related literature and detail the hypothesis we test. Section 3 describes the design of our experiment. Section 4 presents and discusses the implications of our results. Section 5 offers some conclusions.

2. THEORY AND RELATED EVIDENCE

2.1. *Some psychology of probability learning*

The psychology literature reports that the manner in which probability information is presented to experimental subjects has a substantial impact on the observed prevalence of biases in probabilistic judgements. Following the arguments of Estes (1976a, b), Gigerenzer and Hoffrage (1995) explain that, in order to minimise cognitive effort, individuals employ mental procedures which better reflect Bayesian algorithms when likelihoods are represented by frequencies of event occurrence rather than probabilities. Cognitive biases are thus reduced when likelihood information is presented in a frequency format (e.g. *absolute* frequency such as '50 investors enjoyed positive returns' or *relative* frequency such as '50 investors in 1000 enjoyed positive returns') rather than a probability format (e.g. 'the chance of positive returns in 0.05').[2] These observations may be an evolutionary adaptation to the high correlation between observed event frequency and probability in natural environments (Pelham et al., 1994).[3] An implication of the evidence regarding the psychology of probability learning is that individuals have a deeper understanding of the *meaning* of likelihood information, in terms of how it translates choices into outcomes, when it is presented in frequency formats rather than in probability formats. The hypothesis which we test is that, if outcomes are presented in cognitively accessible frequency formats *as well as* the usual less accessible formats, valuations of lotteries are higher. We interpret this as an effect of ambiguity aversion.

2.2. *Some important research regarding ambiguity aversion*

Heath and Tversky (1991) show that people with little knowledge of sports prefer to bet on non-sports events of known likelihood rather than on sports events even when the latter were judged to be equally likely. The authors' explanation is that a lack of knowledge of sports triggers feelings of incompetence, thus making the likelihoods of sports events ambiguous and undermining the willingness to bet on these events. They conclude that ambiguity aversion is driven by feelings of incompetence (knowledge, skill and—of particular relevance here—comprehension) regarding the assessment of event likelihoods. In order to explain the contribution of feelings of incompetence to ambiguity aversion Fox and Tversky (1995, p. 587) consider, "...what conditions produce this state of mind". They propose and provide evidence supporting the *comparative ignorance* hypothesis. This hypothesis emphasises the role of comparing vague events with more familiar events, or comparing one's own knowledge against that of more knowledgeable individuals. Fox and Tversky (1995) interpret their results as showing a reluctance to act on inferior knowledge, the salience of which emerges from a comparison with superior knowledge. This, they argue, reveals ambiguity aversion to be a special case of *source preference* where individuals prefer to bet on events where the uncertainty is from one source (say, a familiar domain) rather than another source (say, an unfamiliar domain). We follow a broadly similar strategy to Fox and Tversky (1995), although the details differ, by providing an account of how incompetence may give rise to ambiguity aversion.

2.3. *A new definition of ambiguity and its novel implications*

Our account of ambiguity is based on the proposition that source is not the only feature of uncertainty over which individuals can plausibly have preferences. The psychological evidence discussed above suggests that the manner in which likelihood information is provided is also important to decision-makers in terms of their understanding of the meaning of that information. We call this *event presentation preference*. The source preference for risk over uncertainty illustrated by the Ellsberg (1961) tasks suggests that not knowing relevant information that could be known is sufficient to render a prospect ambiguous (Camerer 1995). Event presentation preference, on the other hand, suggests this may not be necessary. Simply knowing relevant (likelihood) information may in fact be largely irrelevant to mitigating ambiguity if the meaning of that information is not fully understood. For example, a decision-maker may know that a lottery is [$100, 0.25; zero, 0.75]. They may also understand that on a single trial they may win either $100 or zero. They may, however, be insufficiently trained in probability or too cognitively constrained to run the mental simulations required to realise that in a large number of trials they will win $100 approximately one quarter of the time. In this case the decision-maker fails to fully understand the meaning of the likelihood information. This may undermine their willingness to bet on the lottery. Event presentation preference therefore suggests a new definition of ambiguity: *Generalized ambiguity* is *relevant information that is either not known or not fully* ***understood*** (this is a generalization of Camerer's (1995) preferred definition to allow for deficient understanding of information).

A novel implication of generalized ambiguity is that prospects may be ambiguous under mathematically complete information regarding the probability distribution which governs outcomes, if likelihood information is presented in such a way that its meaning is less than fully comprehended. Note that this implication is contrary to the conventional definitions used in the literature. Choice contexts which involve complete probability information are typically defined as being risky and not as ambiguous. Indeed, Einhorn and Hogarth (1986, p. 229) explicitly state that, "…well-known random processes (such as flipping coins or dice) are uncertain but not ambiguous since the probabilities are well specified". It is possible, however, that the manifestations of ambiguity aversion observed in comparisons of risky and uncertain events have antecedents which are fundamental properties of human decision-making behaviour. If so, broader definitions of behaviours may assist in identifying common antecedents to phenomena which have previously been considered to be distinct.

We now propose *the comprehension hypothesis* which details the manner in which event presentation preference operates to engender incompetence and thereby cause generalized ambiguity aversion (we henceforth refer to generalized ambiguity simply as ambiguity).[4] The hypothesis says that due to the cognitive limitations of human decision-makers and the manner in which the psychology of probability learning has developed (possibly adaptively) to deal with these constraints, some forms of presentation of probabilistic events are not fully understood. Previous research suggests that this applies to information presented in a probability format more than in a frequency format. Since the meaning of event probability information is not as easily comprehended as the meaning of event frequency information, probability formats engender incompetence. Consequently, information presented in a relatively cognitively inaccessible format (e.g. a probability format) is open to ambiguity aversion effects to a greater degree than information presented in a cognitively accessible format (e.g., a frequency format). An implication of the comprehension hypothesis is that ambiguity-averse decision-makers should assign higher values to lotteries when information is presented in a frequency format in addition to a probability format. One interpretation of the comprehension hypothesis is as an extension to Fox and Tversky's (1995) work and suggests that source is not the only facet over which individuals may have preferences: Decision-makers may also have preferences over how likelihood information is presented. These preferences can undermine competence and engender ambiguity aversion. This hypothesis is the essence of the experimental test we describe in the next section.

3. EXPERIMENT DESIGN AND IMPLEMENTATION

The experiment involves subjects assigning valuations to the twenty lotteries described in Table I. The experiment involves two treatments. Treatment 1 is the control treatment, where subjects are asked to place a money value on a simple lottery which involves likelihood information presented in a weak frequency format. Subjects in this treatment are hence termed the *c*-group.

TABLE I
Lottery Parameters[a]

Lottery	Outcomes			Probability			Expected value
	1	2	3	1	2	3	
1	21	9	0	0.2	0.5	0.3	8.7
2	–	9	0	0	0.9	0.1	8.1
3	21	–	0	0.2	–	0.8	4.2
4	–	9	0	–	0.4	0.6	3.6
5	21	–	0	0.7	–	0.3	14.7
6	21	21	–	0.4	0.3	0.3	14.7
7	21	9	0	0.5	0.4	0.1	14.1
8	29	10	0	0.2	0.5	0.3	10.8
9	–	10	–	–	1.0	–	10.0
10	29	–	0	0.2	–	0.8	5.8
11	–	10	0	–	0.5	0.5	5.0
12	10	10	0	0.3	0.2	0.5	5.0
13	29	–	0	0.7	–	0.3	20.3
14	29	10	–	0.5	0.5	–	19.5
15	18	–	0	0.5	–	0.5	9.0
16	–	11	0	–	0.7	0.3	7.7
17	11	10.5	0	0.5	0.2	0.3	7.6
18	24	–	0	0.4	–	0.6	9.6
19	–	13	0	–	0.6	0.4	7.8
20	13	12	0	0.4	0.2	0.4	7.6

[a]All money outcomes and expected values are in pounds sterling.

In Treatment 2 subjects perform the same task as the *c*-group, and with the same information, but first observe a sequence of 10 resolutions of the risk in the lottery. The observation sequence provides likelihood information in a relatively cognitively more accessible (strong) frequency format which mimics natural sampling. Treatment 2 subjects are hence termed the *f*-group. Subjects in the *f*-group saw both representations of probability information since this provides the sharpest test of

LOTTERY:	Lottery Tickets	1 to 2	pay you	£21
	Lottery Tickets	3 to 7	pay you	£9
	Lottery Ticket	8 to 10	pay you	zero

YARDSTICK:	Lottery Tickets	1 to 10	pay you	??.??

Draw:	1	2	3	4	5	6	7	8	9	10
Winnings:	21	0	9	9	21	0	9	9	0	9

Enter a value for ??.?? which would make the LOTTERY and the YARDSTICK equally attractive to you.
The up and down cursor keys select a value. Press <ENTER>to confirm.

Figure 1. Task Display.

ambiguity aversion. For example, if the *f*-group saw likelihood information in *only* the strong frequency format and we subsequently observed differences in valuations between the groups, this difference would be less readily attributable to ambiguity inherent in relatively cognitively inaccessible formats than would be the case in our design.

Figure 1 shows an example of a valuation screen from the experimental software (using lottery 1 from table I). The top 'lottery' box shows the lottery to be valued.[5] All lotteries were expressed in terms of 10 lottery tickets. The 'yardstick' is the vehicle through which the lottery is valued. Subjects were told that they should value the lottery by entering an amount in the small box at the bottom of the screen (which would also appear in the small box within the yardstick) which makes them indifferent between the lottery and the yardstick.[6] The *c*-group saw only the top two boxes in Figure 1. The *f*-group saw these two boxes together with additional event frequency information presented in the third-from-top box identified with 'draw' and 'winnings'. Subjects in the *f*-group would first see the lottery then, when ready, press 'enter' to reveal the observation box showing draws 1–10 and the empty winnings row. Pressing 'enter' again would start the observation sequence wherein the computer would reveal the outcome of a single draw of the lottery under draw 1 in the winnings row, pause and then repeat the process up to draw 10. In terms of Figure 1 the first draw gave £21, the second gave zero and so on. After the observation sequence had finished the valuation message would appear on the screen and the subject would proceed to value the lottery.

The observation sequence for the *f*-group was such that each outcome occurred with the exact frequency suggested by its probability. As can be seen in Figure 1 the probability of winning £9 is 0.5 and the number of times £9 occurs in the observation sequence is 5 out of 10. This feature of our design controls for unrepresentative sequences (which may emerge from genuinely random sequences of draws) affecting lottery valuations. We are interested in the impact of frequency information on ambiguity aversion and not in how differences between frequency and probability information may introduce decision biases. This latter question has been dealt with elsewhere (Humphrey, 1999).[7]

The experiment employed an incentive system which meant that it was in subjects' financial interests to assign valuations to lotteries according to their preference *ordering* over the set of lotteries. Prior to valuing the twenty lotteries, subjects were informed that at the end of the experiment two lotteries would be randomly selected by drawing two numbered discs from a bag containing twenty consecutively numbered discs. The valuations assigned to these two lotteries would be compared and the risk in the lottery to which the subject had assigned the higher valuation would be resolved (by drawing a disc from a bag containing ten consecutively numbered discs) to determine their payment for participation in the experiment. If the two randomly selected lotteries were equally valued, the payment lottery was determined by flipping a coin. Because subjects did not know until after they had valued all twenty lotteries which two would be compared to determine their payment lottery, they could only be sure of playing out their genuinely preferred lottery from any pair by assigning valuations according to their preference ordering over the entire set. Note that since this incentive mechanism only requires valuation orderings to reflect preference orderings, apparently irrational behaviour such as valuing lotteries higher than the largest outcome they offer is incentive compatible. Subjects could in fact perform any monotonic transformation on their genuine *absolute* valuations and still reflect their true preferences over the set of lotteries. Subjects were told, however, that one way to be sure of playing out their truly preferred lottery from the randomly selected pair would be to consider each of the lotteries carefully and assign them genuine absolute valuations.

The incentive system was illustrated to subjects using the responses they provided to two example valuation tasks. These valuations were compared (on the computer screen) and it was explained to subjects that the lottery they had valued highest would be taken to be their preferred lottery out of the two. Subjects were then told that if these example lotteries were the two lotteries randomly selected at the end of the experiment to determine actual winnings, it would, therefore, be the outcome of the higher valued lottery which would constitute payment.

The incentive system described above is a combination of two standard experimental incentive mechanisms. The *random lottery incentive mechanism*, whereby subjects are paid according to a task selected randomly after all tasks have been completed, controls for possible wealth effects which may arise if subjects are paid for more than one task whilst simultaneously making it in subjects interests to consider each task as if it were for real money. The *ordinal payoff scheme* (Tversky et al., 1990) is used in valuation task experiments to elicit preference orderings over lotteries and to circumvent problems associated with other devices which elicit absolute valuations of lotteries, such as the Becker et al. (1964) mechanism (BDM). The BDM device elicits genuine absolute valuations of lotteries by comparing stated valuations with a randomly generated offer. If the offer is less than the valuation the subject plays out the lottery. If the offer is equal to or greater than the valuation the subject receives the offer. It has been shown theoretically (Karni and Safra, 1987; Segal, 1988) that the BDM mechanism can only be trusted to yield genuine absolute valuations if the independence axiom of expected utility theory is true; and there is plenty of evidence documenting its violation (e.g. see Camerer, 1995). By not requiring the BDM device, experiments which employ an ordinal payoff scheme control for this possible source of design bias.

A similar theoretical criticism is levelled at the random lottery incentive system by Holt (1986). He shows that if subjects regard multiple task experiments as a single large decision problem involving compound lotteries which have been generated by reduction according to the calculus of probabilities, it cannot be inferred that the experiment has elicited true preferences unless the independence axiom holds. The essence of Holt's (1986) argument is that if the independence axiom does not hold then on any particular task behaviour may have been contaminated by behaviour over the other tasks in the experiment, and cause differences between that behaviour and the behaviour dictated by preferences in, for example, a single-task experiment.[8] Holt's (1986) argument is theoretically plausible. Starmer and Sugden (1991), however, conduct a direct test of Holt's (1986) hypothesis using different experimental designs for problems over which independence is commonly violated (as the argument requires), and find no evidence to suggest that random lottery experiments do not elicit genuine preferences. They conclude that such designs are legitimate and appropriate. Cubitt et al. (1998) conduct a test of a milder form of Holt's (1986) argument (that there may be contamination between tasks in random lottery experiments) and reach a similar conclusion.[9]

The experiment was conducted at the Centre for Decision Research and Experimental Economics (CeDEx) laboratory at the University of Nottingham. Subjects were recruited by an e-mail message (sent to a CeDEx mailbase of pre-registered undergraduate volunteers) which asked them to reserve a place in one of a number of prearranged sessions. A total of

203 subjects took part in the experiment and were randomly allocated to the *c*- and *f*-groups.

Each session took about one hour to complete, including exhaustive instructions from the experiment organiser and the on-screen example valuation tasks mentioned above. Average subject payment was £12.97, an amount significantly above a UK undergraduate's marginal wage rate. Each subject faced the twenty valuation tasks in random order to control for any possible order effects. Similarly, for *f*-group subjects the order in which outcomes appeared in the observation sequence was randomised. No time limit was imposed.

4. RESULTS

Table II shows the summary statistics of valuations by lottery and by group. For the *c*-group, the mean valuation was greater than the expected value of the lottery in eleven cases, and less than the expected value in nine cases. The *c*-group's median valuation was never greater than the expected value; it was below it in ten cases and equal to it in ten cases. By contrast, the *f*-group's mean valuation was higher than the expected value for nineteen lotteries, and 0.01 below it for the twentieth. The *f*-group's median valuation was equal to the expected value of the lottery in 16 cases, and twice each above or below it. For both groups the median valuation was close to, and often equal to, the expected value of the lottery.[10] The tendency for mean valuations to be greater than median valuations indicates some positive skewness. This is also reflected in the maxima and minima. A relatively small number of subjects gave some very high valuations. In the subsequent analysis we check that our results are robust to the exclusion of these individuals from the sample.

Table III presents some regression analysis of each group's median and mean valuations as a function of the expected value and the standard deviation of the probability distribution of the twenty lotteries. There is evidence of risk aversion,

TABLE II
Summary statistics of lottery valuations[a]

Lottery statistics			Valuation statistics										
			c-group					*f*-group					
Lottery	Expected Value	Standard Deviation	Median	Mean	S.Dev.	Max	Min	Median	Mean	S.Dev.	Max	Min	RRP
1	8.8	7.28	8.70	9.67	3.57	22.00	4.20	9.60	10.42	5.68	48.20	1.50	1.08
2	8.1	2.70	8.10	8.28	2.59	20.00	2.50	8.10	10.50	10.16	72.00	1.00	1.27
3	4.2	8.40	4.00	5.86	12.12	99.90	0.50	4.20	5.10	3.74	20.50	1.00	0.87
4	3.6	4.41	3.60	3.91	1.66	9.00	0.20	3.60	4.29	2.41	16.50	1.00	1.10
5	14.7	9.62	14.70	13.70	4.41	30.00	3.60	14.70	15.22	11.53	99.90	3.00	1.11
6	14.7	9.62	14.70	14.03	4.24	30.50	4.70	14.70	15.64	11.95	99.90	2.00	1.11
7	14.1	7.35	13.00	12.99	4.14	30.00	5.00	14.10	15.18	10.06	75.30	3.50	1.17
8	10.8	10.08	10.80	11.51	5.08	38.00	2.80	11.00	13.28	7.97	60.00	3.00	1.13
9	10.0	0.00	10.00	11.61	5.62	45.00	3.00	10.00	12.71	6.77	50.00	6.00	1.12
10	5.8	11.60	5.00	7.21	9.97	75.00	0.30	5.00	7.40	6.45	35.00	1.00	0.85
11	5.0	5.00	5.00	5.57	5.74	5.20	0.20	5.00	5.84	2.82	20.00	1.50	0.84
12	5.0	5.00	5.00	5.35	1.81	10.00	0.60	5.00	6.18	2.87	15.50	1.90	1.14
13	20.3	13.29	19.00	17.36	6.29	40.00	4.00	20.30	20.29	7.21	50.00	6.20	1.10

14	19.5	9.50	19.00	18.16	6.20	40.00	5.00	19.50	21.84	11.66	99.90	9.00	1.21
15	9.0	9.00	8.50	8.21	3.66	25.00	0.60	8.90	9.48	7.07	60.00	3.00	1.27
16	7.7	5.04	7.70	7.64	2.15	18.00	3.00	7.70	8.67	6.10	49.00	1.50	1.18
17	7.6	4.98	7.60	7.86	2.17	18.50	3.40	7.60	8.96	7.24	58.00	0.00	0.94
18	9.6	11.76	8.50	8.64	4.20	22.00	0.30	9.60	10.58	6.26	45.00	3.00	1.02
19	7.8	6.37	7.20	7.28	2.40	17.00	0.30	7.80	8.34	3.07	20.10	2.00	1.13
20	7.6	6.22	7.60	7.68	2.55	18.00	1.10	7.60	8.73	6.71	60.00	2.70	1.04

[a]"Lottery Statistics" refer to the probability distribution of the lottery as presented to the subjects (for lottery details see Table I); "Valuation Statistics" refer to the valuation of each lottery by the subjects. "RRP" is the relative risk premium for each lottery defined as (mean f-group valuation/expected value)/(mean c-group valuation/expected value).

TABLE III

Regressions of median and mean valuations of each group on expected value and standard deviation of lottery[a]

	Dependent Variable			
	c-group		f-group	
	Mean	Median	Mean	Median
Constant	1.93	0.38	1.69	0.046
	(4.21)	(1.75)	(3.89)	(0.28)
Expected value	0.840	0.983	1.037	1.015
	(20.2)	(49.5)	(26.3)	(68.5)
Standard deviation	−0.061	−0.072	−0.111	−0.025
	(−1.02)	(−2.53)	(−1.96)	(−1.18)
Sample size	20	20	20	20
R-squared	0.968	0.994	0.980	0.997
Standard error	0.749	0.358	0.711	0.267
F-test(2,17)	12.96	6.05	1.91	0.83
[p-value]	[0.000]	[0.010]	[0.178]	[0.455]

[a]Figures in parentheses are t-statistics. The dependent variables are measures of central tendency for the twenty lotteries for each group. The F-statistic refers to the null hypothesis that the coefficient of the lottery expected value is one and the other coefficients are zero.

in the sense that the standard deviation always has a negative coefficient: for a given expected value, lotteries with higher standard deviations received lower mean and median valuations (but this effect is only statistically significant for the c-group median and—marginally—for the f-group mean). The coefficient of the expected value is significantly different from unity only for the c-group mean, for which the coefficient is 0.84. This reflects the tendency for the c-group mean to be above the expected value for low-value lotteries and below it for high-value lotteries. The last row of the table gives the F-statistic for the test that the group's mean (or median) valuation is precisely equal to the expected value (i.e.

that the coefficient of the expected value is unity and that the other coefficients are zero). The null hypothesis is rejected at the 0.01 level in the case of the *c*-group (both for the mean and the median), but not rejected even at the 0.10 level for the *f*-group (either for the mean or the median).

Table IV shows the difference in the average lottery valuations of each group, together with the *t*-statistic of this difference and the average of each of these across the twenty lotteries. A few subjects tended to choose particularly high valuations on some questions. Although this is not irrational given the incentives provided, the last two columns of Table IV. show the same statistics with eight "outliers" excluded from the sample, as a check that the results are not entirely driven by a few individuals. For the full sample, the *f*-group valued the lottery higher on average than the *c*-group in 19 out of 20 lotteries. The chance of this result, under the null hypothesis of no difference between the two groups, is about one in 50,000. The average difference in valuation is 1.33, which is significantly different from zero at the 0.001 level. It is slightly preferable to correct for the variation in accuracy of the estimates of the differences in mean across lotteries by calculating the average *t*-statistic across lotteries. This gives an even stronger result (a *t*-statistic of 7.80 rather than 5.98). If the eight outlying subjects are excluded from the analysis, the difference between the two groups is smaller, but the null hypothesis of no difference is still rejected by all three tests at the 0.001 level. The *f*-group gives a higher valuation in 16 cases, and the two *t*-statistics are 3.95 for the difference of means and 4.78 for the mean *t*-statistic of this difference.

As a further check on the robustness of our results, we ranked the valuations of each lottery in ascending order (i.e. with the lowest valuation ranked one) and compared the average valuation ranking of the *f*- and *c*-groups. This method of analysis might also be considered appropriate for our incentive system which requires valuations to satisfy preference orderings. Table V shows these results, again for both the full sample and with the outliers excluded. As in Table IV the average ranking is higher for the *f*-group in nineteen lotteries

TABLE IV

Difference in mean Valuation of the *f*-group and the *c*-group[a]

Lottery no:	All observations		Excluding 8 "outliers"	
	Diffrence in mean valuations	*t*-statistic of difference	Difference in mean valuations	*t*-statistic of difference
1	0.75	(0.91)	−0.13	(−0.24)
2	2.22	(1.74)	0.20	(0.46)
3	−0.76	(−0.49)	0.49	(0.78)
4	0.39	(1.08)	0.15	(0.06)
5	1.52	(1.01)	−0.09	(−0.13)
6	1.61	(1.04)	−0.26	(−0.39)
7	2.20	(1.65)	0.55	(0.85)
8	1.76	(1.53)	1.39	(1.51)
9	1.15	(1.08)	0.98	(1.25)
10	0.19	(0.13)	1.18	(1.14)
11	0.28	(0.36)	0.95	(2.38)
12	0.83	(2.01)	0.85	(2.06)
13	2.93	(2.52)	2.92	(2.81)
14	3.67	(2.28)	2.52	(2.55)
15	1.27	(1.30)	0.48	(0.84)
16	1.04	(1.31)	0.06	(0.16)
17	1.11	(1.20)	−0.09	(−0.20)
18	1.93	(2.11)	1.61	(2.15)
19	1.06	(2.23)	0.98	(2.30)
20	1.06	(1.21)	0.44	(1.08)
Mean	1.33	1.31	0.76	1.07
(*t*-statistic)	(5.98)	(7.80)	(3.95)	(4.78)
Share of plus signs [*p*-value]	19/20 [.000]		16/20 [.006]	

[a]The dependent variable is the valuation of the specified lottery by an individual subject. The table shows the *f*-group valuation minus the *c*-group valuation. The penultimate row shows the mean of the figures in that column across the 20 lotteries, and the *t*-statistic for the null hypothesis that this mean is zero. In the bottom row the *p*-value is based on a two-tailed test of the null hypothesis that plus and minus signs are equally likely.

with the full sample and in sixteen lotteries with the restricted sample. The average t- statistic for the difference of means is 1.08 for the full sample (17% smaller than for the valuations themselves), but the null hypothesis that this average is zero is still comfortably rejected ($t = 6.40$). In the restricted sample, the results are actually stronger than in Table IV (although still weaker than for the full sample).

The comprehension hypothesis developed earlier suggested that subjects' understanding of a probability distribution would be enhanced if they saw a series of observations drawn from that distribution as well as receiving a mathematically complete verbal description. With greater understanding of the distribution, subjects would experience less ambiguity and would feel more competent in their decisions, so that they would value a lottery more highly. Our findings are consistent with this hypothesis: when decision-makers observe a sequence of outcomes which exactly reflect the probability distribution which govern the risky prospect they are valuing, valuations are significantly higher than when no such observations are experienced.

5. CONCLUSIONS

We have shown that lotteries are valued significantly more highly when information is presented to subjects in a frequency as well as in a probability format. It is likely that this is a consequence of reduced ambiguity about the probability distribution of possible outcomes. We interpret our results in terms of preferences over the presentation of frequency information and the effect this has on feeling of (in)competence in decision-making. In related work, Fox and Tversky (1995) have suggested, on the basis of a series of experiments, that ambiguity aversion is present only in a comparative context in which a person evaluates both clear and vague prospects, and seems to disappear in a non-comparative context where a person evaluates only one of these prospects in isolation. They point out that their results do not sit easily with

TABLE V

Difference in mean rankings of lottery by group[a]

Lottery no:	All observations		Excluding 8 "outliers"	
	Diffrence in mean valuations	t-statistic of difference	Difference in mean valuations	t-statistic of difference
1	5.88	(0.88)	2.23	(0.34)
2	0.27	(0.04)	−3.47	(−0.53)
3	8.19	(1.23)	6.88	(1.05)
4	6.13	(0.92)	4.25	(0.65)
5	1.36	(0.20)	−0.88	(−0.13)
6	−1.36	(−0.20)	−3.74	(−0.57)
7	9.31	(1.40)	7.24	(1.11)
8	8.34	(1.24)	10.02	(1.51)
9	7.05	(1.22)	6.08	(1.08)
10	5.88	(0.87)	7.40	(1.11)
11	11.50	(1.76)	12.31	(1.92)
12	10.90	(1.67)	10.47	(1.63)
13	15.71	(2.37)	16.14	(2.48)
14	15.40	(2.32)	15.50	(2.38)
15	5.73	(0.86)	3.88	(0.60)
16	3.72	(0.55)	0.41	(0.06)
17	0.06	(0.01)	−3.63	(−0.55)
18	11.86	(1.77)	11.84	(1.80)
19	13.27	(1.99)	12.18	(1.85)
20	3.55	(0.53)	3.00	(0.45)
Mean	7.14	1.08	5.91	0.91
(t-statistic)	(6.38)	(6.40)	(4.25)	(4.28)
Share of plus signs [p-value]	19/20 [.000]		16/20 [.006]	

[a]The dependent variable is the ranking of the valuation of the specified lottery by an individual subject (lowest valuation ranked one). The table shows the f-group mean valuation ranking minus the c-group mean valuation ranking. The penultimate row shows the mean of the figures in that column across the 20 lotteries, and the t-statistic for the null hypothesis that this mean is zero. In the bottom row the p-value is based on a two-tailed test of the null hypothesis that plus and minus signs are equally likely.

models of risky choice which involve decision weights, because in decision-weighting models there is no distinction between comparative and non-comparative evaluations. The results of our experiment demonstrate the existence of ambiguity aversion even in a non-comparative context. Subjects in the *c*-group were not provided with the opportunity to compare their understanding of different types of information presentation. Their understanding of this information may be inferior to that of the (*f*-group) frequency format information, but there was no comparison to render this inferiority salient. On Fox and Tversky's (1995) view there would therefore be no catalyst for ambiguity aversion in the *c*-group.

Secondly, our results suggest interesting lines of inquiry which may have relevance to the debate over whether expected utility theory should be replaced by an alternative theoretical account of risky choice. Many experiments designed to test competing theories of risky choice have employed probability format presentations of likelihood information. In the light of our results it is possible that these experiments may have confounded the influence of ambiguity and the other features of individual behaviour with which the tests were primarily concerned. Although our data show that less ambiguous representations of likelihood information increase the valuations of individual lotteries, it is unclear how less ambiguous representations would affect behaviour over sets of multiple lotteries such as those involved in, for example, the common consequence and common ratio effects. It is these violations of expected utility which have often been invoked as evidence to undermine the case for maintaining that theory. If, however, these experiments were re-run with a cognitively accessible frequency format presentation of likelihood information and it was found that violations of expected utility theory diminished, then perhaps the debate would swing towards those in favour of maintaining expected utility theory. Frequency formats are after all often present in natural environments (e.g. investors observe a time series of gains and losses on stocks and estimates of likelihood stem from the frequencies of these) and external validity is usually a potent argument. This is, of course, presently little more than speculation, but speculation which it would seem worthwhile to pursue with additional empirical investigation.

Extensions to our work to address these empirical questions might also consider variations to our experimental design. For example, in the *f*-group we have used a particular frequency format presentation of likelihood information. We employed this format for two reasons. First, we felt it conveys some external validity regarding the observation of time series of outcomes (stock performances, sports results etc.). Second, our reading of the psychological literature, particularly the work of Estes (1976a, b), suggested that this format would give the effect we have tested a fair chance to emerge. It remains to be seen whether other frequency format presentations (such as revealing the outcomes of *n* resolutions of risk simultaneously rather than sequentially) render lotteries more or less ambiguous than is the case with the frequency format discussed above. Finally, we have also used a particular method of preference elicitation. One feature of the valuation task we used is that it allows subjects to concentrate on one lottery at a time and avoid the cognitively costly compensatory comparisons involved in choice tasks. This arguably renders the task substantially simpler, particularly for the *f*-group who were confronted with the additional need to attend to the observation sequence. Our results may of course vary with other preference elicitation methods. Pairwise choice tasks and lottery ranking tasks would be two interesting places to start this investigation.

ACKNOWLEDGEMENTS

The authors are grateful for financial support from The Nuffield Foundation, award no. SGS/LB/0295, to David Lawson for running the experiments, Chris Starmer, Robin Cubitt, participants at the 2004 FUR XI meetings in Paris (particularly Jerome Busemeyer, Mark Machina and George Wu) and an anonymous referee for valuable comments, and to Olga Babouchkina for excellent research assistance.

NOTES

1. Each sequence shown was unbiased in the sense that it had the same expected value as the probability distribution described.
2. See Harries and Harvey (2000) for a recent discussion of some of this evidence.
3. The adaptive origins of event frequency-based probability learning may be related to Kiseilius and Sternthal's (1986) availability–valence hypothesis. This states that more frequently observed outcomes are more available in memory and are therefore easier to imagine as future outcomes.
4. Note that whereas Heath and Tversky (1991) refer to 'feelings of incompetence' leading to ambiguity aversion, we refer simply to incompetence. This is because we do not require decision-makers to be consciously aware of their incompetence as is suggested by the use of the word 'feelings'.
5. Note that the description of event likelihoods in Figure 1 is in a frequency format (and not a *pure* probability format such as, for example, '£21 occurs with probability 0.2'). We employed this format in order to enhance understanding of the task in the context of the incentive scheme we used (described below). Our test of the comprehension hypothesis is not affected by this presentation. All that is required for our test is that the frequency information provided to the *f*-group enhances the understanding of the likelihood information in relation to the presentation experienced by the *c*-group.
6. Subjects were told that indifference meant that after they had entered their valuation they would not mind whether they received either the lottery or the yardstick. The valuation was made by using the 'up' and 'down' cursor keys on the keyboard. Pressing the up key replaced the question marks with £00.00, pressing again incremented this to £00.10 and so on. The down key generated 10 pence decrements. There was no upper bound on valuations and a zero lower bound. The valuation was confirmed by pressing 'enter' followed by a chance to change it or move on to the next problem.
7. Our test relies on subjects in the *f*-group attending to the information in the observation sequence. However, in order to provide the strongest test of the comprehension hypothesis, we did not want our instructions to add any objectively irrelevant importance to this information such that the f group might be unduly cued towards it rather than the initially stated probability information in the description of the lottery. To this end *f*-group subjects were simply told that, "...before you value each lottery in the experiment we are going to show you what the outcome of ten draws might look like...the outcomes of the observation sequence are purely

illustrative". It was then explained—in exactly the same manner as it was explained to the *c*-group—that winnings from the experiment would be determined at the end of the experiment by randomly selecting one plastic disc from a bag containing ten consecutively numbered discs and that the prize given by each numbered disc is that described in the LOTTERY on the computer screen.

8. Assume our experiment contained only two lotteries, A and B, and V(.) denotes the valuations assigned to these lotteries. $V(A) > V(B)$ would be taken as evidence of a strict preference for lottery A over lottery B. In an experiment with twenty lotteries Holt's (1986) hypothesis views $V(A) > V(B)$ as equivalent to choosing the compound lottery L_1: [A,p; Z, $1 - p$], where p is the probability of lotteries A and B being randomly selected and their valuations compared to determine the payment lottery at the end of the experiment, and Z is give by behaviour over the remaining 18 lotteries, Similarly $V(B) > V(A)$ is equivalent to selecting L_2: [B, p; Z, $1 - p$] Holt (1986) is right, and the independence axiom of expected utility theory does not hold, then observing $V(A) > V(B)$ in an experiment with additional tasks cannot be taken as evidence of a genuine preference for A over B because those preferences may have been disrupted by the influence of the common term in Z.
9. Moreover, if we find evidence supporting our hypothesis, it could not be explained by any generalisation of expected utility theory which relaxes independence and embodies choice between gambles depending only on outcomes and probabilities, even if independence is violated and even if there is contamination of the revelation of relative valuations. Any such theory would view the incentive scheme faced by our two groups as completely equivalent and so could not explain any difference between groups. See Cubitt et al. (2004) for a similar argument.
10. This observation is consistent with the suggestion that although our incentive system only requires subjects to report valuations which respect their preference ordering over the set of twenty lotteries, subjects generally did not monotonically transform their valuations. If so we may conclude that observed valuations are similar to the genuine *absolute* valuations elicited by other incentive schemes such as the Becker–DeGroot–Marschak mechanism.

REFERENCES

Becker, G., DeGroot, M. and Marschak, J. (1964), Measuring utility by a single-response sequential method, *Behavioral Science* 9, 226–232.

Camerer, C. (1995), Individual decision making, in Kagel J. H. and Roth, A.E. (eds.), *The Handbook of Experimental Economics* (Princeton University Press, Princeton).

Camerer, C. and Weber, M. (1992), Recent developments in modelling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty* 5, 325–370.

Cubitt, R., Starmer, C. and Sugden, R. (1998), On the validity of the random lottery incentive system. *Experimental Economics* 1,115–131.

Cubitt, R., Munro, A. and Starmer, C. (2004), Testing explanations of preference reversal, *Economic Journal* 114, 707–726.

Einhorn, H. J. and Hogarth, R. M. (1986), Decision making under ambiguity, *Journal of Business* 59, 225–249.

Ellsberg, D. (1961), Risk, ambiguity and the Savage axioms, *Quarterly Journal of Economics* 75, 643–696.

Estes, W. K. (1976a), The cognitive side of probability learning, *Psychological Review* 83, 37–64.

Estes, W. K. (1976b), Some functions of memory in probability learning and choice behavior, in Bower, G. H. (ed.), *The Psychology of Learning and Motivation* vol. 10, Academic Press, New York.

Fox, C. and Tversky, A. (1995), Ambiguity aversion and comparative ignorance, *Quarterly Journal of Economics* 110, 585–603.

Gigerenzer, G. and Hoffrage, U. (1995), How to improve Bayesian reasoning without instruction: Frequency formats, *Psychological Review* 102, 684–704.

Harries, C. and Harvey, N. (2000), Are absolute frequencies, relative frequencies, or both effective in reducing cognitive biases?, *Journal of Behavioral Decision Making* 13, 431–444.

Heath, C. and Tversky, A. (1991), Preference and belief: Ambiguity and competence in choice under Uncertainty, *Journal of Risk and Uncertainty* 4, 5–28.

Holt, C. (1986), Preference reversals and the independence axiom, *American Economic Review* 76, 508–515.

Humphrey, S. J. (1999), Probability learning event-splitting effects and the economic theory of choice, *Theory and Decision* 46, 51–78.

Karni, E. and Safra, Z. (1987), Preference reversal and the observability of preferences by experimental methods, *Econometrica* 55, 675–685.

Kisielius, J. and Sternthal, B. (1986), Examining the vividness controversy: An availability valence interpretation. *Journal of Consumer Research* 12, 418–431.

Pelham, B. W., Sumarta, T. T. and Myaskovsky, L. (1994), The easy path from many to much: The numerosity heuristic, *Cognitive Psychology* 26, 103–133.

Segal, U. (1988), Does the preference reversal phenomenon necessarily contradict the independence axiom? *American Economic Review* 78, 233–236.

Starmer, C. and Sugden, R. (1991), Does the random-lottery incentive system elicit true preferences? An experimental investigation, *American Economic Review* 81, 971–978.

Tversky, A., Slovic, P. and Kahneman, D. (1990), The causes of preference reversal, *American Economic Review* 80, 204–217.

Address for correspondence: Steve Humphrey, Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham, Nottingham NG7 2RD, UK. Phone.:+44 115 951 5472; E-mail: steve.humphrey@nottingham.ac.uk