Author(s) Stewart, Neil; Chater, Nick; Stott, Henry P.; Reimers, Stian

Article Title: Prospect relativity: How choice options influence decision under risk.

Year of publication: 2003

Link to published version: http://dx.doi.org/10.1037/0096-3445.132.1.23

Running Head: PROSPECT RELATIVITY

Prospect Relativity: How Choice Options Influence Decision Under Risk

Neil Stewart, Nick Chater, Henry P. Stott, and Stian Reimers

University of Warwick

Stewart, N., Chater, N., Stott, H. P., & Reimers, S. (2003). Prospect relativity: How choice options influence decision under risk. *Journal of Experimental Psychology: General, 132,* 23-46.

Abstract

In many theories of decision under risk (e.g., expected utility theory, rank-dependent utility theory, and prospect theory), the utility of a prospect is independent of other options in the choice set. The experiments presented here show a large effect of the available options, suggesting instead that prospects are valued relative to one another. The judged certainty equivalent for a prospect is strongly influenced by the options available. Similarly, the selection of a preferred prospect is strongly influenced by the prospects available. Alternative theories of decision under risk (e.g., the stochastic difference model, multialternative decision field theory, and range frequency theory), where prospects are valued relative to one another, can provide an account of these context effects.

Prospect Relativity: How Choice Options Influence Decision Under Risk

Decisions almost always involve trading off risk and reward. In crossing the road, one balances the risk of accident against the reward of saving time; in choosing a shot in tennis, one balances the risk of an unforced error against the reward of winning. Choosing a career, a life partner, or whether to have children involves trading off different balances between the risks and returns of the prospects available. Understanding how people decide between different levels of risk and return is, therefore, a central question for psychology.

Understanding how people trade off risk and return is also a central issue for economics. The foundations of economic theory are rooted in models of individual decision making. For example, to explain the behavior of markets one needs a model of the decision-making behavior of buyers and sellers in those markets. Most interesting economic decisions involve risk. Thus, an economic understanding of markets for insurance, of risky assets such as stocks and shares, of the lending and borrowing of money itself, and indeed of the economy at large requires understanding how people trade risk and reward.

In both psychology and economics, the starting point for investigating how people make decisions involving risk has not been empirical data on human behavior. Instead, the starting point has been a normative theory of decision making, expected utility (hereafter, EU) theory (axiomatized by von Neumann & Morgenstern, 1947), which specifies how people ought to make decisions and plays a key role in theories of rational choice (for a review, see Shafir & LeBoeuf, 2002). The assumption has then been that, to an approximation, people do make decisions as they ought to, that is, that EU theory can be viewed as a descriptive, as well as a normative, theory of human behavior (Arrow, 1971; Friedman & Savage, 1948). At the core of EU theory are the assumptions that people make choices that maximize their utility and that they value a risky option by the EU (in a probabilistic sense of expectation) that it will provide. In general, the prospect $(x_1, p_1; x_2, p_2; \ldots ; x_n, p_n)$, where outcome x_i occurs with probability p_i, and $p_1 + p_2 + \ldots + p_n = 1$, has EU

$$U(x_1, p_1; x_2, p_2; \ldots; x_n, p_n) = p_1 u(x_1) + p_2 u(x_2) + \ldots + p_n u(x_n) \qquad (1)$$

(The function *U* gives the utility of a risky prospect. The function *u* is reserved for the utility of certain outcomes only.)

In psychology and experimental economics, there has been considerable interest in probing the limits of this approximation, that is, in finding divergence or agreement between EU theory and actual behavior (see, e.g., Kagel & Roth, 1995; Kahneman, Slovic, & Tversky, 1982; Kahneman & Tversky, 2000). It is now well established that people systematically violate the axioms of EU theory (see Camerer, 1995; Luce, 2000; Schoemaker, 1982, for reviews). In economics more broadly, there has been interest in how robustly economic theory copes with anomalies between EU theory and observed behavior (for a range of views, see, e.g., Akerlof & Yellen, 1985; Cyert & de Groot, 1974; de Canio, 1979; Friedman, 1953; Nelson & Winter, 1982; Simon, 1959, 1992).

The present article demonstrates a new and large anomaly for EU theory in decision making under risk. Specifically, we report results that seem to indicate that people do not possess a well-defined notion of the utility of a risky prospect and hence, a fortiori, do not view such utilities in terms of EU. Instead, people's perceived utility for a risky prospect appears highly context sensitive. We call this phenomenon *prospect relativity*.

Motivation From Psychophysics

In judging risky prospects, people must assess the magnitudes of risk and return that the prospects comprise. The motivation for the experiments presented here was the idea that some of the factors that determine how people assess these magnitudes might be analogous to factors underlying assessment of psychophysical magnitudes, such as loudness or weight. Specifically, people appear poor at providing stable absolute judgments of such magnitudes and are heavily influenced by the options available to them. For example, Garner (1954) asked participants to judge whether tones were more or less than half as loud as a 90-dB reference loudness. Participants' judgments were entirely determined by the range of tones played to them. Participants played tones in the range 55-65 dB had a half-loudness point, where their judgments were more than half as loud 50% of the time and less than half as loud 50% of the time, of about 60 dB. Another group, which received tones in the range 65-75 dB, had a half-loudness point of about 70 dB. A final group, who heard tones in the range 75-85 dB, had a half-loudness point of about 80 dB. Laming (1997) provided an extensive discussion of other similar findings. Context effects, like those found by Garner, are consistent with participants making perceptual judgments on the basis of relative magnitude information, rather than absolute magnitude information (Laming, 1984, 1997; Stewart, Brown, & Chater, 2002a, 2002b). If the attributes of risky prospects behave like those of perceptual stimuli, then similar context effects should be expected in risky decision making. This hypothesis motivated the experiments in this article, experiments that were loosely based on Garner's procedure.

Existing Experimental Investigations

A few experiments have already investigated the effect of the set of available options in decision under risk. Mellers, Ordóñez, and Birnbaum (1992) measured participants' attractiveness ratings and buying prices (i.e., the price that a participant would pay to obtain a single chance to play the prospect and have a chance of receiving the outcome) for a set of simple binary prospects of the form "p chance of x." These experimental prospects were presented with one of two sets of filler prospects. For one set of filler prospects, the distribution of expected values was positively skewed, and for the other set, the distribution of expected values was negatively skewed. Attractiveness ratings of the experimental prospects were significantly influenced by the filler prospects, with higher ratings for prospects in the positive skew condition than for the same prospects in the negative skew condition. However, context had only a very small effect on buying price. With more complicated prospects of the form "p chance of x otherwise y," the effect of skew on buying price was slightly larger. The large effect that the set of options available had on attractiveness ratings and the much smaller effect on buying price are consistent with a similar demonstration by Janiszewski and

Lichtenstein (1999). They gave participants a set of prices for different brands of the same product to study. The prices varied in range. The range had an effect on judgments of the attractiveness of a new price but not on the amount participants reported that they would expect to pay for a new product.

The set of options available as potential certainty equivalents (hereafter, CEs) has been shown to affect the choice of CE for risky prospects. In making a CE judgment, participants suggest or select from a set of options the amount of money for certain that is worth the same to them as a single chance to play the prospect. We considered CE judgments extensively in our experiments. Birnbaum (1992) demonstrated that skewing the distribution of options offered as CEs for simple prospects, while holding the maximum and minimum constant, influenced the selection of a CE. When the options were positively skewed (i.e., when most values were small) prospects were undervalued compared to when the options were negatively skewed (i.e., when most values were large).

MacCrimmon, Stanbury, and Wehrung (1980) presented some evidence that the set in which a prospect is embedded can affect judgments about the prospect. They presented participants with two sets of five prospects to be ranked in order of attractiveness. The expected value of each prospect was constant across all prospects and both sets. There were two prospects in common between the two sets. If context provided by the other prospects in a set did not affect the attractiveness of a prospect, each participant should have consistently ranked one prospect as more attractive than the other in both sets. MacCrimmon et al. found that 9 of a total of 40 participants had a different ordering of the two prospects in the two sets. They argued that this was not merely inconsistency, because these participants made consistent rankings within a set, but instead reflected the influence of the other prospects in the choice set. Following this logic, however, it would take only one participant who had a different ordering of the two prospects but who otherwise behaved consistently to conclude that there was an effect of choice set. A random fluctuation in risk aversion between sets might produce this result. With such a small number of data points and no concrete null hypothesis allowing a significance test to be made any conclusion based on this results must be very tentative.

In summary, there is an effect of previously considered prospects on the attractiveness rating assigned to a current prospect and also a small effect on buying price. Moreover, the context provided by a set of values from which a CE is to be chosen affects CE judgments. Finally, in choosing between prospects, there is a suggestion that other prospects in the choice set may reverse preferences between identical pairs of prospects. In the experiments reported below we found large and systematic effects of choice set (both potential CEs and accompanying prospects) on the valuation of individual prospects. These effects are not compatible with EU theory or some of its most influential variants, according to which the value of a prospect is independent of other available options. These results are, though, compatible with a variety of models that discard this independence assumption. We consider such models in the General Discussion.

Summary of Experiments

As indicated above, in this study we adapt methodologies from psychophysics to investigate the possibility that context effects influence decision under risk. The aim of Experiments 1A-1D was to determine whether options offered as potential CEs influenced estimates of a prospect's CE. We consistently found substantial effects. In Experiment 2, to investigate these effects, we introduce a new procedure in which, under certain assumptions, it was optimal for participants to provide truthful CEs. In Experiment 3, we examined whether these effects were similar to those observed in magnitude estimation tasks. In the remaining two experiments, Experiments 4 and 5, we investigated whether the effect of available options arose in choices between prospects as well as in CE judgments about prospects.

Experiment 1A

Following a similar logic to Garner's (1954) loudness judgment experiment described above, we gave participants a set of four options as possible CEs for a series of prospects. Participants were asked to decide on a CE for the prospect and then select the option closest to their estimate. For each prospect, options were either all lower in value than the mean free choice CE (given by another group of participants) or all higher.[1] If participants were not influenced by the set of options, then their choice of option should have been that nearest to their free-choice CE. A key prediction of this hypothesis is that either the highest option of the low options (L_4), the lowest option of the high options (H_1), or both should be chosen more than half of the time. Consider the distribution of free-choice CEs illustrated in Figure 1A. If H_1 is to be selected less than half of the time, then the area to the left of the H_1-H_2 bound must be less than one half. This area corresponds to the proportion of times that the free choice CE is nearest to H_1. Thus, the area to the right (labeled Area A) must be greater than one half. This means that the proportion of times that L_4 will be selected from the low options must be greater than one half, as this proportion corresponds to the sum of area to the right of the H_1-H_2 bound (Area A - which was greater that half) and the area between the L_3-L_4 and H_1-H_2 bounds (Area B - which depends on the exact probability density function, but must be greater than or equal to zero). This argument is true for any probability density function. Similarly, if L_4 is selected less than half of the time, H_1 must be selected more than half of the time (Figure 1B). The selection of L_4 less than half of the time and H_1 less than half of the time is not consistent with any distribution of free-choice CEs that is not affected by context. If participants' responses were solely determined by the set of options presented to them, however, then the distribution of responses across options should have been the same for both the low- and the high-value options.

Method

Participants. Free-choice CEs were given by 14 psychology undergraduates from the University of Warwick. Another 16 psychology undergraduates chose CEs from sets of options. All participated for course credit. Ages ranged from 18 to 20 years. All but three participants were female.

Design. A set of 20 prospects, each of the form "*p* chance of *x*," was created by crossing the amounts £200, £400, £600, £800, and £1000 with the probabilities .2, .4, .6, and .8. In a pretest, 14 participants were asked to provide a CE for each prospect. The means and standard deviations of the free choice CEs were calculated for each prospect (see Appendix A).

For each prospect, two sets of options were created as follows. In the low-options condition, the options were 1/6, 2/6, 3/6, and 4/6 of a standard deviation below the mean. In the high options condition, options were 1/6, 2/6, 3/6, and 4/6 of a standard deviation above the mean. Thus, the range of each set was half a standard deviation. Options were rounded to have familiar, easy-to-deal-with values.

Sixteen new participants were presented with the prospects and options and asked to select the option closest to their CE from a set of four. Eight participants received the low options for every prospect. The other eight received the high options for every prospect. Note that this method, in which a range of potential CEs is presented, is not uncommon in other experimental work in this area (see, e.g., Tversky & Kahneman, 1992).

Procedure. Participants were asked to imagine choosing between "£30 for certain" or a "50% chance of £100" to illustrate that prospects could have a value. They were told they would be asked to value a series of prospects. It was explained that the purpose of the experiment was to investigate how much they thought the prospects were worth and that there were no correct answers. Participants were asked to choose the option nearest the value they thought the prospect was worth to them.

Each prospect was presented on a separate page of a 20-page booklet. The ordering of the prospects was random and different for each participant. Probabilities were always presented as percentages. Options were always presented in numerical order, as in the following example of a low-option set:

How much is the gamble

"60% chance of £400"

worth?

Is it: £60 £80 £100 £120

When collecting pretest free-choice CEs, the options were omitted, and a blank line on which participants could write their CE was added.

Results

Participants took approximately 5 minutes to complete the task. Under free-choice conditions, the average CE (see Appendix A) increased with both probability of winning and the amount that could be won, demonstrating that participants were sensitive to manipulations of both. The chosen CE was an approximately linear function of the independent effects of prospect amount and prospect probability.

The proportion of times each option was chosen is plotted in Figure 2. The distribution of options is approximately the same for the two conditions. L_4 was chosen significantly less than half the time, $t(7) = 4.21$, $p = .0040$ ($\eta^2 = .72$). The same was true of H_1, $t(7) = 5.26$, $p = .0012$ ($\eta^2 = .80$). Thus the hypothesis that participants' CEs would be unaffected by context can be rejected. (An alpha level of .05 was used for all statistical tests in this article, but for informational value, we also report the exact *p* value of each test.[2])

Discussion

CE judgments were strongly influenced by the CE options offered. These data therefore appear to illustrate prospect relativity: People do not seem to form a stable absolute judgment of the value of a prospect but instead choose an option relative to the options available. The preference for central options in each set may be an example of extremeness aversion (also called the compromise effect; see Simonson & Tversky, 1992). Indeed, these data may reflect a more general tendency to prefer central options that is seen when choosing among identical options (e.g., products on a supermarket shelf; Christenfeld, 1995).

Experiment 1B

A natural explanation of the effect of the set of options available in Experiment 1A is that, on a given trial, the options available affected a participant's judgment. However, an alternative and, for our purposes, less interesting explanation is that when participants were repeatedly presented with trials containing too-high or too-low options, they learned to readjust their judgments to fit their responses within the alternatives given. One way of ruling out this alternative explanation was to use a within-participants design. In this design each participant was presented with some trials on which all the options were lower than the free-choice CE and others on which all the options were higher. If the effect seen previously had been caused by participants learning to adjust their judgments up or down to fit into the response scale, the effect should now have disappeared. However, if the effect was caused by the options available on that trial only, then the pattern of results demonstrated in Experiment 1A should have been be replicated.

Method

Participants. Free-choice CEs were given by 35 volunteers. Twenty-eight different volunteers chose CEs from sets of options. All participants were undergraduates or postgraduates from the University of Warwick, and none had participated in Experiment 1A. Ages ranged from 18 to 30 years, with a mean of 22 years. Two thirds of participants were female. Participants were paid £5 for taking part in this and other related experiments.

Design and procedure. The design was the same as in Experiment 1 except that for each participant, 10 trials were randomly selected to have options all higher than the pretest mean for that prospect, the other 10 having options all below the mean. Trials were randomly intermixed. Free-choice CEs and corresponding options are given in Appendix B. The procedure was the same as in Experiment 1A.

Results

Participants took approximately 5 minutes to complete the task. One participant's data were excluded from the analysis because he had been given a misleading answer to a query about the task that would have led to an inappropriate response strategy. As before, under free choice, the CE increased approximately linearly with both the amount that could be won and the probability of winning.

Figure 3 shows the proportion of choices of each option. In both conditions, the proportion of responses increases with proximity to the mean free-choice CE. Planned *t* tests showed that the proportions of L_4 and H_1 responses were significantly below .5, $t(26) = 2.65$, $p = .0135$ ($\eta^2 = .21$), and $t(26) = 3.81$, $p = .0008$ ($\eta^2 = .36$), respectively.

A further issue is whether the context effects found in the main analysis applied to all participants or whether some people showed a larger context effect than others. For each participant, two scores were constructed, one for each condition. Participants were awarded one point for each time they chose the lowest option, two for the next lowest, three for the second highest, and four for the highest (i.e., scores were the rank of the options selected in each condition). Those showing no effect of the option set should have chosen the L_4 and H_1. However, those who based their judgment entirely on the option set would choose midrange options. Thus, a negative correlation between low-choice CEs and high-choice CEs would be evidence that people varied in the extent to which context influenced their CE decisions. Far from being negatively correlated, there was a significant positive correlation ($r^2 = .45$, $p = .0001$). One interpretation of this positive correlation is that participants had a tendency to choose an option with the same rank across the low- and high-choice CE conditions. (Note that this analysis could not be run for Experiment 1A because the set of options was manipulated between participants.)

Perhaps surprisingly, we found no evidence that the options offered on the previous trial influenced the option selected on the current trial. One might imagine that, say, offering low options on the previous trial might cause participants who were trying to be consistent to have selected an option lower than they otherwise would on the current trial. However, the rank (as described above) of the option selected within the set did not depend on the previous option set offered (mean rank = 2.63, SE = 0.11, for prior low options vs. mean rank = 2.62, SE = 0.13, for prior high options), $t(26) = 0.23$, $p = .8407$. This test is sufficiently sensitive to be able to detect a difference of 0.21 between the mean ranks with a power of 80%.[3]

Discussion

As in Experiment 1A, participants' CE judgments in Experiment 1B were influenced, at least in part, by the options from which the CE had to be chosen. In Experiment 1B, there was a tendency to prefer higher options from the low CEs set and lower options from the high CEs set. (This contrasts with Experiment 1A, where there was no such tendency.) Overall, the pattern of results shown in Experiment 1B is intermediate between that expected under the hypothesis that the available options are irrelevant and that expected if the available options are the only determinant of responses. The tendency in Experiment 1B can be accounted for in three ways. First, the effect in Experiment 1A may have been partly caused by participants readjusting their responses to fit in with the options available. In Experiment 1B, such an adjustment need not have been made as the option set was manipulated within participants, and thus a smaller context effect was observed. However, this explanation still requires an additional factor, such as prospect relativity, to contribute to the effects seen in Experiments 1A and 1B.

A second explanation of the tendency is that the absolute spacing of the options differed between Experiments 1A and 1B. In both experiments, the spacing of CE options was set at 1/6 of a free-choice SD. However, the free-choice SDs were smaller in Experiment 1B, and thus, the absolute values of the options were actually more closely spaced in Experiment 1B.[4] For this reason, we replicated Experiments 1A and 1B as conditions of a single experiment, using the spacings from Experiment 1B throughout. The between-participants result was similar to those of Experiment 1A and the within-participants result similar to those of Experiment 1B, so we have not presented them here. An explanation of the different pattern of preferences in Experiments 1A and 1B in terms of using different options can thus be ruled out.

A third explanation of the difference between Experiments 1A and 1B, and one that we favor, is that participants may have been trying to be consistent with their responses to previously completed questions. Thus, encountering both low and high CE options caused participants to seem less affected by context. Such a consistency effect cannot be due to the immediately preceding trial alone as there was no effect of the immediately preceding trial. Instead, if this kind of explanation is correct, the tendency must be due to some larger window of previous trials.

Experiment 1C

The effect of available options in Experiments 1A and 1B appears to create difficulties for EU and related theories as descriptive accounts of decision under risk. Yet these difficulties may be less pressing if the effects demonstrated thus far arose only because the options presented as CEs were simply too close together, and that participants were roughly indifferent between them. (Although note that this ought to lead to a u-shaped preference across the options, rather than an n-shape.[5]) If this were the case then these effects should disappear when the options are more widely spaced and hence are no longer indifferent between them. In Experiment 1C, we investigated the effect of increasing the spacing of the options.

Method

The method was the same as in Experiment 1B (i.e., low and high options were presented within participants) except that there were three between-participants spacing conditions. In the narrow condition, options were spaced at 1/6 of a free-choice standard deviation as in Experiments 1A and B. In the wide condition, the spacing between adjacent options was doubled to one third of a free-choice standard deviation, and thus the wide condition differed from the narrow condition both in the spacing of the options and in the means of the option sets. A gap condition was introduced, with the option spacing of the narrow condition and the option group means of the wide condition. Thus, the gap condition differed from the narrow condition only on the mean of the sets and from the wide condition only on the spacing of the options. Eighteen undergraduate and postgraduate students from the University of Warwick took part in the gap condition, and 19 in each of the narrow and wide conditions. Ages ranged from 18 to 30 years, with a mean of 22 years. Two thirds of participants were female.

Results

The proportion of times each option was selected is shown in Figure 4. In the narrow condition, option L_4 was selected significantly less than half the time, $t(18) = 4.02$, $p = .0008$ ($\eta^2 = .47$), as was the H_1 option, $t(18) = 3.42$, $p = .0031$ ($\eta^2 = .39$), replicating the results of Experiment 1B. In the wide condition, option L_4 was selected less than half the time, $t(18) = 2.04$, $p = .0565$ ($\eta^2 = .19$), although this difference is only marginally significant. The H_1 option was selected significantly less than half of the time, $t(18) = 3.52$, $p = .0024$ ($\eta^2 = .41$). The key result is that doubling the spacing of the options did not eliminate the context effect. In the gap condition, L_4 was not chosen significantly less than half of the time, $t(17) = 0.17$, $p = .8665$, but H_1 was, $t(17) = 2.99$, $p = .0082$ ($\eta^2 = .34$).

The differences between the conditions were examined with a two-way (Condition x Option Set) univariate analysis of variance (ANOVA), with the mean rank of the option selected as the dependent measure. (It was not possible to run a Condition x Option Set x Option ANOVA, as the selection of options was not independent: The proportion of times each option was selected must sum to 1.) There was no significant main effect of condition, $F(2, 53) < 1.00$. There was a significant main effect of set, $F(1, 53) = 56.16$, $p < .0001$ ($\eta^2 = .51$). Participants chose options with higher ranks in the low-options condition compared with the high-options condition. The interaction was also significant, $F(2, 53) = 4.00$, $p = .0241$ ($\eta^2 = .13$). The tendency to prefer L_4 and H_1 was larger in the wide and gap condition than the other conditions. To examine the interaction further, we ran a one-way univariate ANOVA with condition as a factor and the difference in ranks between the low and high options as the dependent measure. Ryan REGWQ post hoc tests revealed that the tendency to respond with high-ranking low options and low-ranking high options (i.e., the central tendency) was significantly smaller for the narrow condition, with no difference between the wide and gap conditions.

There was a marginally significant positive correlation between the mean rank of the options selected in the low- and high-options trials in the narrow-option spacing condition, $r^2 = .20$, $t(17) = 2.07$, $p = .0543$. This replicates the correlation seen in Experiment 1B. This correlation was not seen in the wide, $r^2 = .01$, $t(17) = 0.45$, $p = .6574$, and gap conditions, $r^2 = .05$, $t(16) = 0.95$, $p = .3570$. A correlation of $r^2 \geq .35$ can be detected with 80% power in this design. Despite the failure to find positive correlations in the wide and gap conditions, there is reason to think the positive correlations seen in the narrow condition of this experiment and in Experiment 1B are reliable: We have replicated the correlation in the replication of Experiments 1A and 1B reported in the discussion of Experiment 1B and also in a further, unpublished study.

As in Experiment 1B, we investigated whether there was an effect of the option set offered on the previous trial on the option selected on the current trial. A two-way (Condition x Previous Option Set) ANOVA was run, with the rank of the option selected on the current trial as the dependent measure. There was no significant main effect of condition, $F(2, 53) = 1.38$, $p = .2601$. As in Experiment 1B, there was no significant main effect of the previous option set, $F(1, 53) < 1.00$ (mean rank of option selected on current trial with low options on the previous trial = 2.51, $SE = 0.05$, vs. a mean rank = 2.50, $SE = 0.07$, with high options of the previous trial). (A single *t* test for the difference in mean ranks between the low- and high-option sets can detect a difference of .17 with a power of 80%.) There was no significant interaction, $F(2, 53) = 2.12$, $p = .1301$.

Discussion

The effect of the options offered as CEs seen in Experiment 1B was replicated in the narrow condition of Experiment 1C. Doubling the spacing of the options in the wide condition did not eliminate the effect. However, there was a greater tendency to select lower options from the high set and higher options from the low set in the wide and gap conditions when the difference between the means of the option sets was larger. We consider two possible accounts of this result. First, increasing the set mean spacing may have caused participants to rely more on some representation of the absolute utility of each prospect. Perhaps this might be because they realized that the options were quite different from some approximate representation of the utilities of the prospects. However, in this case one might have expected the effect to be larger in the wide rather than the gap condition as the wide condition contained the more disparate options. In fact, the difference, although nonsignificant, was in the opposite direction. A second explanation is that increasing the spacing of the set means should have increased a consistency effect where participants tried to give consistent answers across the low- and high-option sets. Both explanations offer an account of the lack of a correlation between the rank of the options selected across the low- and high-option sets for the wide and gap conditions. Consistency between the low- and high-option sets should have caused participants to select low options from the high set and high options from the low set, as should an increased reliance upon the prospects' absolute utilities. Thus, consistency should have caused a negative correlation that would have acted in opposition to the positive correlation observed in Experiment 1B and the narrow condition here, leaving a net zero correlation.

Experiment 1D

Experiment 1D aimed to discriminate between the two explanations of the tendency to select L_4 options from the low set and H_1 options from the high set seen in the wide and gap conditions of Experiment 1C. If the tendency was due to a consistency effect between the low and high sets, then repeating the experiment with option set as a between-participants variable should have eliminated the effect. Alternatively, if the wider spacing somehow caused participants to rely more on some representation of absolute utilities, then the greater central tendency should also have been seen in the wide spacing condition even if a given participant saw only low options or only high options. For this reason, in Experiment 1D, option set manipulated between participants.

Method

Participants. Sixty professionals attending a conference at the University of Warwick were approached on the campus and asked to participate. Participants' ages ranged between 20 and 40 years, with a mean of 27 years. An approximately equal number of male and female participants took part.

Design and procedure. Spacing (narrow or wide) and option set (low or high) were manipulated between participants: Fifteen participants took part in each condition. In the wide condition, the low options were set at .10, .20, .30, and .40 of the expected value of each gamble. The high options were set at .60, .70, .80, and .90 of the expected value. In the narrow condition, the interval between options was halved. The low options were set at .30, .35, .40, and .45 of the expected value. The high options were set at .55, .60, .65, and .70 of the expected value. In other respects, the procedure was the same as in previous experiments.[6]

Results

The proportion of times each option was selected is shown in Figure 5. Overall, there was a preference for the lowest options. This was caused by some participants stating that they did not gamble and then selecting the lowest option on every trial. In the narrow condition, option L_4 was selected significantly less than half the time, $t(14) = 5.69$, $p = .0001$ ($\eta^2 = .70$), as was the H_1 option, $t(14) = 2.31$, $p = .0368$ ($\eta^2 = .28$), replicating the results of Experiment 1A. In the wide condition, option L_4 was selected less than half the time, $t(14) = 2.17$, $p = .0477$ ($\eta^2 = .25$), as was the H_1 option, $t(14) = 2.33$, $p = .0354$ ($\eta^2 = .28$). Doubling the spacing of the options did not eliminate the context effect. As in Experiment 1C, a two-way (Spacing x Option Set) ANOVA was run, with the mean rank of the option selected as the dependent measure. There was no main effect of spacing, $F(1, 56) < 1$. In contrast with Experiment 1C, there was no main effect of option set, $F(1, 56) < 1$, and no interaction, $F(1, 56) < 1$. As option set was manipulated between participants the correlation across option sets and the effect of the option set on the previous trial could not be examined.

Discussion

In Experiment 1C, where option set was manipulated within participants, there was a greater tendency to prefer L_4 and H_1 when the options were widely spaced. However, such a tendency was not evident here, when option set was manipulated across participants. Thus, not all of the central tendency evident in Experiment 1C can be attributed to participants relying more on some representation of absolute utilities: Instead, at least some of this effect must be due to consistency across option sets.

Summary of Experiments 1A-1D

To sum up thus far, Experiment 1A demonstrated that the options presented as CEs had a large effect on CE judgments for simple prospects. In Experiment 1B, the set of options was manipulated within participants to investigate whether the results of Experiment 1A might be due to participants adjusting their CE estimates over the course of the experiment to fit in with the options offered. However, the context effect remained. In Experiment 1C, the spacing of the options was increased to investigate the limits of this prospect relativity. However, even when the options were widely spaced, there was an effect of option set. A greater tendency to prefer high options from the set of low options and low options from the set of high options was seen when options were more widely spaced. In Experiment 1D, the greater tendency was not observed when the different option sets were presented across participants. This suggests that the tendency was due, at least in part, to participants attempting to be consistent across conditions in the within-participant designs, rather than participants relying more on some representation of absolute utilities.

Experiment 2

In Experiments 1A-1D, the set of options offered as CEs affected the CE selected. Experiment 2 was designed to demonstrate the same effect of restricting the range of CE options in a task where it was optimal for participants to report CEs truthfully. Although psychologists typically assume that participants are honest when providing hypothetical CEs, economists are typically concerned with providing a system of incentives that ensures it is optimal for participants to provide truthful CEs. Hence, the results above may be criticized from an economic perspective.

There is evidence that psychologists are correct in their assumption that participants are typically honest in their judgments. For example, Lichtenstein and Slovic (1971) demonstrated preference reversals in choices between two prospects and in CE estimates for those prospects in situations where decisions were hypothetical and in situations where there was an incentive system (see also Tversky, Slovic, & Kahneman, 1990). Preference reversals have also been demonstrated with ordinary gamblers playing for high stakes in Las Vegas (Lichtenstein & Slovic, 1973; see also Grether & Plott, 1979). For further discussion of these and other similar findings, see Hertwig and Ortmann (2001) and Luce (2000, pp. 15-16). However, because of the potential importance of the findings from Experiments 1A-1D for models of economics, we include this experiment where the incentive system had been designed to motivate participants to provide truthful CEs.

The design follows a solution to the cake-cutting problem, where a cake must be divided fairly between two children. One solution is to allow one child to cut the cake into two pieces and the other child to select the piece. The first child should cut the cake exactly in half, otherwise the other child will take the larger piece of cake, leaving the first child with the smaller piece.

In this experiment, participants divided a sum of money into an amount for certain and an amount that could be be won with a known, fixed probability. For example, they might split £1,000 into a sure amount of £300 and the prospect "60% chance of £700." Participants knew that the other person (who was not the experimenter) would select either the prospect or the sure amount, taking the better of the two, leaving the participant with the other. Thus, it was optimal for participants to split the given amount so that they had no preference between the resulting fixed amount and the resulting prospect. Note that this procedure works only if each participant assumes that the chooser has the same level of risk preference as himself or herself. To this end, participants were told to assume that the chooser did have the same risk preference as they did.

This procedure is more simple than other methods used to elicit truthful CEs, for example, the first price auction, or the Becker, DeGroot, and Marschak (BDM; 1964) procedure. In the BDM procedure, participants are given the chance to play a prospect and are asked to state the minimum price at which they would sell the prospect. A buying price is then randomly generated by the experimenter, and if it exceeds the selling price, then the prospect is bought from the participant. If not, then the participant plays the prospect. It is the case that it is optimal for participants to state a selling price that is the CE of the prospect, though it is unlikely that many participants realize this.

Method

Participants. Participants were psychology undergraduates from the University of Warwick who had not participated in Experiments 1A-1D. Ages ranged from 18 to 25 years, with a mean age of 20 years. The majority of participants were female. Seventeen participants took part in the free-choice condition of the experiment. Nineteen further participants took part in the restricted-choice conditions. Participants were paid £5, plus performance-related winnings of up to £4.

Design. In each trial of the free-choice condition, participants divided a given amount of money x into two smaller amounts y and z to make one fixed amount (y) and the prospect "*p* chance of *z*." Probability *p* was known to participants before splitting amount *x*. Participants were told that one trial in the experiment would be selected at random at the end of the experiment and that a second person would take either the fixed amount or the prospect for himself, leaving the participant with the other option. Under the assumption that the chooser had the same risk preference as they did, it was explained to participants that the chooser would choose the option with greater utility, leaving the participants with their less preferred option, if they did not split the amount to make options of equal utility. It was therefore optimal for participants to split the amount *x* into amounts *y* and *z* such that *y* and a "*p* chance of *z*" were equivalent for them, that is, that *y* was the CE for the prospect "*p* chance of *z*."

The restricted-choice conditions differed by offering participants a choice from a set of four pre-split options rather than giving them a completely free choice. That is, values for *y* and *z* were presented, and participants selected one pair that could be played at the end of the experiment. It could be argued that participants might reasonably think that the pairs of *y* and *z* values presented might provide them with information about the chooser's risk preferences. For this reason, two people were used in running the experiment. One person was responsible for administering the tests and the other for making the choice at the end of the experiment. The intention was to keep the roles of the experimenter and the person making the choice at the end of the experiment separate in participants' minds to minimize the degree to which participants would think that the options provided information about the chooser's risk preference.

It was hypothesized, as in Experiments 1A-1D, that the set of pairs of values for *y* and *z* presented would influence participants' choices. To investigate this, we varied one between-participants factor. The set of values for *y* and *z* were selected such that either *y* was always greater than the mean free-choice value of *y* and *z* was smaller than the mean free-choice value of *z*, or vice versa. These corresponded to the low- and high-option sets in Experiments

1A-1D. The precise option sets were constructed as follows. The mean and standard deviation of the free-choice amount were calculated for each prospect (see Appendix C). The two sets of equally spaced options (for the high-value and low-value conditions) were calculated as described for Experiment 1A. As in Experiment 1A, if participants were not influenced by the set of choices, then the distribution of responses across the options should not have been biased towards the free-choice splitting. There were 32 trials in the experiment, made by crossing four values of p (.2, .4, .6, and .8) with eight values of x (£250, £500, ... , £2,000).

Procedure. All conditions of the experiment began with written instructions. It was explained to participants that they were playing a gambling game and that they should try to win as much money as possible. They were told that a single trial would be randomly selected at the end of the experiment and used to determine their bonus. They were told the purpose of the experiment was to investigate how much people thought various prospects were worth. It was emphasized that it was optimal for the participants to split the money so they thought the amount for certain was equal in worth to a chance on the prospect. Participants were told that if they allocated funds so that either the amount for certain was worth more than the prospect or vice versa, then the chooser would take the better option, leaving them with less than if they had allocated the money so the prospect was worth the certain amount. They were told that although they could not be certain what the chooser would do, they should assume that the chooser would behave as they would themselves.

Participants were given five practice trials. One of the trials was chosen at random, and it was explained that if the chooser chose the fixed amount, then the prospect would be played, and the participants would get the winnings. They were also told that if instead the chooser took the prospect, they would get the fixed amount. Note that this discussion was hypothetical and that participants were not actually told what the chooser's preference would be.

After the practice participants completed a booklet of options. The pairs of options were presented in a random order to each participant. An example page from a free-choice condition booklet is shown in Figure 6A. In the restricted-choice conditions, presplit options were presented as in Figure 6B. When the experiment was completed, one trial was chosen at random and played to determine each participant's bonus (using an experiment exchange rate of 0.0024).

Results

Participants took between half an hour and one hour to complete the booklet. It seems that the introduction of a bonus caused participants to deliberate on their answers for much longer than in Experiments 1A-1D. One participant from the free-choice condition was eliminated from subsequent analysis for showing a completely different pattern of results to other participants, suggesting he had misunderstood the task. The participant had decreased the value of the fixed amount *y* as the chance of the prospect amount *p* increased (i.e., he had responded as if prospects with a higher chance of winning were worth less to him). Fourteen out of 512 trials (16 participants x 32 trials) where the initial amount had been incorrectly split, were deleted and treated as missing data.

For the free-choice splits, as the total amount *x* increased, participants' allocation of the fixed amount *y* increased. As the probability *p* of winning the prospect increased, participants' estimates of the value of the prospect *y* also increased. Thus, participants' responses seemed lawful and sensible, indicating that the task made sense to them.

The choices made in the restricted-choice conditions are shown in Figure 7. Participants did prefer end options over central options in both conditions, as would be expected if participants were not influenced by the option set. However, if there were no effect of context, L_4 and H_1 should have been chosen over half of the time. L_4 was chosen significantly less than half of the time, $t(9) = 3.47$, $p = .0070$ ($\eta^2 = .57$), as was H_1, $t(8) = 4.20$, $p = .0023$ ($\eta^2 = .69$). Thus, the proportion of times each option was selected differed significantly from the proportions expected under the assumption that participants were not influenced by the options available.

Discussion

The results of the restricted-choice conditions in Experiment 2 replicate the prospect relativity finding shown in Experiment 1 under a more rigorous procedure. When participants were presented with a range of presplit total amounts, so that the CE options were either always lower or always higher than the free-choice value, the context provided by the presplit options influenced their choice of CE.

Experiment 3

The demonstration of apparent prospect relativity in risky decision making suggests that the representation of the utility dimension may be similar to that of perceptual psychological dimensions, where context effects have also been demonstrated. Empirical investigations in absolute identification (Garner, 1953; Holland & Lockhead, 1968; Hu, 1997; Lacouture, 1997; Lockhead, 1984; Long, 1937; Luce, Nosofsky, Green, & Smith, 1982; Purks, Callahan, Braida, & Durlach, 1980; Staddon, King, & Lockhead, 1980; Stewart, 2001; Stewart et al., 2002a; Ward & Lockhead, 1970, 1971), magnitude estimation, matching tasks, and relative intensity judgment (see, e.g., Jesteadt, Luce, & Green, 1977; Lockhead & King, 1983; Stevens, 1975, p. 275) have shown that perceptual judgments of stimuli varying along a single psychological continuum are strongly influenced by the preceding material. If the representation of utility is similar to the representation of these simple perceptual dimensions, then preceding material might be expected to influence current judgments, as it does in the perceptual case.

Simonson and Tversky (1992) provided several cases where preceding material does indeed influence current judgments in decision making. For example, when choosing between pairs of computers that vary in price and amount of memory, the trade-off between the two attributes in the previous choice affects the current choice. Indeed, preference reversals can be obtained by varying the preceding products. In Experiment 3, we considered the effect of preceding material on judgments concerning a single dimension (utility) rather than the trade-off between two dimensions. Participants simply provided CEs for simple prospects of the form "*p* chance of *x*." We then examined the effect of preceding prospects on the CE assigned to the current prospect.

Method

Participants. Fourteen undergraduates from the University of Warwick participated for payment of £3. All participants had previously taken part in the free-choice condition of Experiment 2.

Design. Participants were asked to state the value of a series of prospects. Participants had previously taken part in a task where estimating the value of prospects truthfully optimized their reward, compared with overestimating or underestimating the value of a prospect. Participants were instructed to continue providing CEs in the same way.

Ten sets of 10 simple prospects of the form "*p* chance of *x*" were constructed. Figure 8 shows the values of *p* and *x* for each prospect. Each set of prospects lying on the same curve (these are hyperbolas) shares a common expected value. (The slight deviation of the crosses from the line is caused by rounding the values of *p* and *x*.) Prospects were chosen in this fashion simply because it produces an equal number of prospects with each expected value. The order in which prospects were presented was random and different for each participant.

We hypothesized that the preceding prospect should affect the value assigned to the current prospect as follows. If the previous prospect had a low expected value, then we expected that the current prospect would be overvalued. Conversely, if the previous prospect had a high expected value, then we expected that the current prospect would be undervalued. This prediction was motivated by the contrast effects observed in the analogous perceptual task, magnitude estimation.

Procedure. Participants were told that they would be asked to value prospects and that they should do this in the same way as in the previous experiment (the free-choice condition of

Experiment 2). Participants completed a booklet with a separate prospect on each page, together with a space for their valuation.

Results

Figure 9 plots the mean value of prospects, as a function of the expected value of the previous prospect, for each possible expected value of the current prospect. CEs given to a prospect increased as the expected value of the prospect increased. The response was, on average, 97% of the expected value (*SD* = 36%) showing slight risk aversion, on average. The expected value of the previous prospect had no effect on the value assigned to the current prospect (i.e., the lines in Figure 9 are flat).

To examine possible sequence effects more closely, we completed a linear regression for each participant to see what proportion of the variability in the current response could be explained by the previous prospect's expected value after the effects of the attributes of the current prospect had been partialled out. The previous prospect's expected value was a significant predictor of the current response for just 1 of the 14 participants, no more than would be expected by chance. For this participant, $r^2 = .04$, and for all other participants $r^2 < .04$. Similar analyses for the previous (a) response, (b) x, and (c) p also showed no sequential dependencies. (For this design, an $r^2 \geq .08$ can be detected with a power of 80%.) Across all participants, the mean slope of the best fitting regression lines did not differ significantly from zero for any predictor from the previous trial.

Discussion

In perceptual tasks where a series of stimuli are presented and a judgment is made after each stimulus, the response to the current stimulus is shown to depend on the stimuli (or responses; the two are normally highly correlated) on previous trials. In other words, the response on the current trial is systematically biased by (some aspect of) the previous trial. Some authors (e.g., Birnbaum, 1992) have suggested that judgments about risky prospects might be similarly affected. Here, CE judgments for simple prospects did not show sequential dependencies like those shown in the analogous perceptual judgment tasks. There was little carryover of information from one trial to the next. This finding is consistent with Mellers et al.'s (1992) result, where the buying prices of a set of critical prospects were at most only slightly influenced by the expected value of (previously encountered) filler prospects.

Experiment 4

Experiments 1 to 3 investigated the effect of context in CE judgment tasks. Careful discussion by Luce (2000) highlighted the difference between judged CEs, where participants provide a single judgment of the value of a prospect, and choice CEs, derived from a series of choices between risky prospects and fixed amounts. For example, for the kinds of prospects with large amounts and moderate probabilities used here, judged CEs are larger than choice CEs (see, e.g., Bostic, Herrnstein, & Luce, 1990). The preference reversal phenomenon (see, e.g., Lichtenstein & Slovic, 1971) is evidence that there is often a discrepancy between choice-based and CE-based methods of assessing utility (see also Tversky, Sattath, & Slovic, 1988). Indeed, Luce went as far as to advocate developing separate theories for judged and choice CEs.

Experiment 4 investigated context effects in a choice-based procedure rather than a judged CE-based procedure. Participants made a single choice from a set of simple prospects of the form "p chance of x" where the probability of winning was traded off against the amount that could be won. The context was provided by manipulating the range of values of p (and therefore of x) offered.

Method

Participants. Ninety-one undergraduate and postgraduate students from the University of Warwick took part. Ages ranged from 18 to 40 years, with a mean age of 26 years. An approximately equal number of male and female participants took part. None had previously participated in any other experiment described in this article. Payment was determined by playing the prospect selected by each participant, and winnings were between £0 and £2.

Design. Participants were each offered a set of simple prospects of the form "*p* chance of *x*." Within the set, the probability of winning and the win amount were traded off against one another, and thus the choice was between a large probability of winning a small amount through to a smaller probability of winning a larger amount. Ten prospects were used: "50% chance of £50," "55% chance of £45," ..., "95% chance of £5."

If utility is assumed to be a simple power function of x, as is standardly assumed in economics, with exponent γ - that is $u(x) = x^{\gamma}$ - the EU of each prospect can be calculated.[7] Figure 10 plots utility as a function of the probability of winning p for different curvatures of the utility function (values of γ). For a risk-neutral person ($\gamma = 1.0$), for whom utility is proportional to monetary value, the probability of winning for the prospect with the maximum utility is $p = .5$. For a risk-averse person ($\gamma < 1.0$), the prospect with maximum utility has a larger probability of winning a smaller amount; the maximum falls at higher p for lower γ. The key observation is that the prospect with maximum utility in the set is determined by the level of risk aversion γ. Thus, a participant's choice of prospect can be mapped directly onto a degree of risk aversion.

There were three between-participants conditions in the experiment. In the free-choice condition, all 10 prospects were presented. In two other conditions, the choice of prospects was limited to either the first or the second half of the prospects available in the free-choice condition. In the more risky condition, the prospects ranged from a "50% chance of £50" to a "70% chance of £30." In the less risky condition, the prospects ranged from a "75% chance of £25" to a "95% chance of £5."

Procedure. Each participant was presented with a sheet listing a set of prospects. The prospects were presented in an ordered table, with a row for each prospect and columns headed "chance of winning" and "amount to win." Probabilities were presented as percentages. Participants were asked to choose one prospect from the set to play. Before making their choice, they were given an explanation of how the prospect would be played. The selected prospect was played, and participants were paid according to its outcome, multiplied by an experiment exchange rate (0.002).

Results

The results are displayed in Table 1. Blank cells indicate that the prospect was not available for selection in that condition. Two hypotheses were tested. The first was that participants are sensitive to the absolute values of prospect attributes and are unaffected by the choice options. According to this hypothesis, in the restricted choice conditions, participants should have chosen the prospect in the set that was nearest to the prospect they would have chosen under free-choice conditions. For example, in the more risky condition, only 5 participants selected the prospect "70% chance of £30" (less risky prospects were not available), but in the free-choice condition, 3 + 6 + 3 + 1 + 5 + 1 = 19 participants selected the prospect "70% chance of £30" or one less risky. A 2 x 2 contingency table was constructed to test the hypothesis that there was no difference in the proportion of people selecting the prospect "70% chance of £30" or one less risky between the two conditions. The difference in proportions was significant, Fisher's exact $p = .0002$. An analogous table was constructed to test the difference between the less risky and free choice conditions. Again, the difference in proportions was significant, Fisher's exact $p = .0029$. In conclusion, we can reject the hypothesis that participants in the restricted-choice conditions chose the prospect in the set nearest to the prospect they would have chosen under free-choice conditions and were otherwise uninfluenced by the set of options.

The second hypothesis tested was that, although there may be some effect of the choices available, there would still be some effect of the absolute magnitude of prospect attributes. If so, we would expect a tendency for participants in the most risky condition to choose the least risky prospect available and vice versa. However, if participants' choices were determined solely by the set of available prospects, then the distribution of responses across options (from the most risky to the least risky) should not differ across the more risky and less risky conditions. There was no significant difference, $\chi^2(4, N = 61) = 2.89$, $p = .5767$. In other words, there is no evidence that the absolute riskiness of a prospect had any influence on the choices made in each of the restricted-choice conditions. For this chi-squared test, a difference in J. Cohen's (1988) $w = 0.45$ - which corresponds to a $\chi^2(4) = 27.45$ - can be detected with 80% power.

Discussion

Participants were asked to select a single prospect from a set to play. Within the set, the probability of winning a prospect was reduced as the amount that could be won was increased. Thus, each participant faced a choice between prospects offering small amounts with high probability through to larger amounts with a lower probability. In the restricted conditions, participants were offered only a subset of the prospects available. If participants' preferences were unaffected by the set of options provided, they should simply have chosen the prospect closest to the prospect they would select under free-choice conditions. However, the distribution of choices differed significantly from those expected under this prediction. Instead, the set of options available seemed to determine participants' preferences, and there was no significant evidence that participants were sensitive to the absolute level of risk implicit in a prospect. In conclusion, the level of risk aversion shown by a participant was shown here to be a function of the set of prospects offered.

We know of only one other experiment where the effects of the context provided by the choice set has been shown to affect the prospect chosen. In an unpublished study by Payne, Bettman, and Simonson (reported in Simonson & Tversky, 1992), participants were asked to make a choice between a pair of three-outcome prospects. Adding a third prospect that was dominated by one of the original prospects but not the other significantly increased the proportion of times the (original) dominating prospect was selected over the (original) nondominating prospect. This effect has also been seen when making nonrisky decisions where, for example, participants chose between $6 or a famous brand pen. The introduction of a pen from a lesser known brandname increased the proportion of participants selecting the famous brand pen and reduced the proportion selecting the $6 (see Simonson & Tversky, 1992, for this and other examples). The notion of trade-off contrast, where participants who are assumed to have little knowledge about the trade-off between two properties, deduce what the average trade-off is from the current or earlier choice sets, can account for this type of data.

However, it is not immediately obvious how the notion of trade-off contrast might account for the results of Experiment 4. Across both contexts, the trade-off between probability and amount was constant (as the chance of winning the prospect was increased by 5%, the amount to win fell by £5). Instead, it seems that participants had no absolute grip of the level of risk implicit in each prospect in the choice set and instead chose a prospect with reference to its riskiness relative to the other prospects in the set. This demonstration of prospect relativity in choice is consistent with that described in earlier experiments, where CE judgments were used.

Experiment 5

Our final experiment was designed to investigate the extent to which a choice between two prospects is affected by preceding context. Thus, this experiment mirrors Experiment 3, but with actual choices rather than CE judgments. On each trial, participants chose between a sure amount of money and a prospect offering a larger amount with a known probability. Let us informally call a trial risky (or safe) to the degree that participants are expected to prefer the risky prospect (or the sure amount of money). For example, we expected people to prefer the risky prospect over the sure amount more often in (a) a "50% chance of £100" or £10 for sure compared with (b) a "50% chance of £100" or £40 for sure. Half of the trials, the common trials, were given to all participants and were designed so that the sure amount was such that a typical, moderately risk-averse participant would be indifferent between the sure amount and the risky prospect. The other half of the trials were filler trials, and their properties were manipulated between participants. For half of the participants, the filler trials were constructed so that only a very risk-averse individual would be indifferent to the sure amount and the risky prospect. For these risky trials, most participants should have favored the risky prospect. For the other half of the participants, the filler trials were constructed so that only relatively risk-neutral participants would favor the risky prospect. For these safe trials, most participants should have favored the sure amount. The intention was to assess whether the riskiness of the filler trials would affect choices on the common trials. If participants represented current prospects relative to previous prospects, then the common trials should have seemed relatively safe if the filler trials were risky, and participants should have favored the safe, sure amount. Conversely, if the filler trials were safe, then the common trials should have seemed relatively risky, and participants should have favored the risky prospect.

Method

Participants. Thirty-five undergraduate and postgraduate students from the University of Warwick took part in the experiment and were paid £5 for participating in this and three other related experiments. Ages ranged from 18 to 30 years, with a mean of 22 years. Two thirds of participants were female.

Design. Thirty-six trials were generated, each of which comprised a simple prospect of the form "p chance of x" and an amount offered for certain. The amounts £100, £200, £300, £400, £500, and £600 were crossed with the win probabilities of .1, .2, .4, .6, .8, and .9 to give 36 prospects. Half of the trials were selected at random and consistently used to set the context. For half of the participants, the fixed amount offered on these trials was low, and for the other half of the participants, the fixed amount was high. The other half of the trials was common to both groups, and the fixed amounts were at an intermediate level.

A sure amount was generated by using Equation 2.

$$y = x\, p^{\frac{1}{\gamma}} \qquad (2)$$

where y is the sure amount and the prospect is a "p chance of x." γ describes the curvature of a hypothetical power law utility function, $u(x) = x^{\gamma}$. $\gamma = 1$ for a risk-neutral person. Smaller values of γ denote greater risk aversion. For each condition, six values of γ were used. The values 0.50, 0.55, and 0.60 were used to generate sure amounts for the common trials. Risky fillers were generated using the values 0.35, 0.40, and 0.45, which made the prospects on the experimental trials seem comparatively unattractive. Safe fillers were generated using the values of 0.65, 0.70, and 0.75. (For the population used in this experiment, we observed values of γ in this range in an unpublished study from our laboratory. The values of γ were deduced from choices between simple prospects and sure amounts.) The assignment of values of γ to trials was such that a given value of γ occurred only once for each probability and only once for each prospect amount. Otherwise, the assignment was random and the same for all participants.

Procedure. Participants were given brief oral instructions. They were told that they would have to imagine making choices between playing a prospect to receive an amount of money and taking a smaller amount for sure. Each pair of options was presented on a separate page of a 36-page booklet and appeared as follows:

Which option do you prefer?

10% chance of £300

£12

Participants were told to mark the option they would prefer and move on to the next page. They were also made aware that there was no objective right answer and that choice was a matter of personal preference.

Results

The dependent measure was the mean proportion of trials on which the prospect was preferred to the sure amount. With safe fillers, participants selected the risky prospect significantly more often in the experimental trials (mean = .53, *SE* = .04) than in the filler trials (mean = .40, *SE* = .05) as hypothesized, $t(16) = 7.10$, $p < .0001$. With risky fillers, participants selected the risky prospect less often in the experimental trials (mean = .47, *SE* = .05) than in the filler trials (mean = .67, *SE* = .04), again, as hypothesized, $t(17) = 7.39$, $p < .0001$. The comparison of interest is performance on the common trials across the safe and risky conditions. For the common trials, the risky prospect was selected slightly more often in the condition in which it was designed to look more attractive, but the difference did not approach significance, $t(33) = 0.8$, $p = .4305$. This design can detect a difference of .14 with a power of 80%.

Discussion

Imagine being presented with a choice between a sure amount and a simple risky prospect. If, in previous trials, the sure amount was low compared with the prospect, then the sure amount on this trial might seem quite appealing. Conversely, if, in previous trials, the sure amount was high compared with the prospect, then the prospect on this trial might seem quite appealing. However, this experiment found no evidence that the properties of preceding trials affected people's judgments between a prospect and a certain amount. The results of this experiment point further to the notion that context effects are much more potent within a trial than between trials and that this is the case for CE judgments (Experiment 3) and choice paradigms (Experiment 5).

We were quite surprised by the small or lacking effect of previously considered prospects on the current choice and so have conducted a meta-analysis of sequential effects in other choice experiments from our laboratory. The experiments involved making choices between two prospects, each of the form "*p* chance of *x* otherwise *y*," where $x > y > 0$. In each pair, one prospect was always more risky than the other (i.e., the probability of winning was smaller) but had a higher expected value. Thus, the choice was always between a comparatively more likely but smaller amount versus a larger but less likely amount. Trials were split into two groups according to whether the total expected value of the prospects on the previous trial was more or less than the median amount. The proportion of risky picks on the current trial did differ significantly between the two groups, $t(95) = 1.99$, $p = .0422$ ($\eta^2 = .04$), although the actual difference in proportions was very small (.39 when the previous expected value was high vs. .41 when the previous expected value was low). It seems that this small effect was largely caused by the prospect with the smaller expected value on the previous trial, as a median split of current trials on this attribute led to a slightly larger significant difference (.39 vs. .42), $t(95) = 1.99$, $p = .0079$ ($\eta^2 = .04$). Splitting by other attributes of the previous trial (e.g., the difference in expected value, the maximum outcome, the higher expected value, the maximum probability of a zero outcome, and the probability of the maximum outcome) did not produce significant differences. In conclusion, it seems that the effects of previous choices between risky prospects on current choices are small in comparison to the within-trial effects.

General Discussion

Together, the results presented in this article suggest that prospects are judged relative to accompanying prospects, a phenomenon that we call prospect relativity. In Experiments 1A-1D, the set of options offered as potential CEs for simple prospects had a large effect on the CE selected. In Experiment 2, this effect was replicated despite monetary incentives designed to encourage participants to deliver accurate CEs. In Experiment 4, the set from which a simple prospect was selected was shown to have a large effect on the prospect that was chosen. In two further experiments, Experiments 3 and 5, previously considered prospects had little or no effect on judgments about the current prospect. It seems that the context provided by items that are considered simultaneously does affect decisions about risky choice but that the context provided by previously considered risky choices, even if they are very recent, has little effect. We call this effect the simultaneous consideration effect.

In the following section, we briefly review existing theories of decision under risk and investigate what account they may offer, if any, of the prospect relativity phenomena presented. Existing theories can be divided roughly into two classes: (a) those where the utility or value of a prospect depends only on the attributes of the prospect and (b) those where prospect attributes are compared against those of other competing prospects.

Independent Prospect Evaluation Theories

EU theory, rank-dependent utility theory (Quiggin, 1982, 1993; see Diecidue & Wakker, 2001, for an intuitive introduction), configural weight models (Birnbaum, Patton, & Lott, 1999), and prospect and cumulative prospect theories (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) all assign a risky prospect with a value or utility that depends only on the attributes of that prospect. For this reason, this entire class of theory must fail to offer an account of the effect of the set of options from which a CE is chosen, as in Experiments 1A-1D and 2. This class of theory also provides no account of the effect of the set of prospects from which a prospect is chosen on that choice, as in Experiment 4.

Dependent Prospect Evaluation Theories

In the following theories, the utility or value of a prospect is not independent of the other prospects in the choice set. Thus these theories are potential candidates in accounting for the findings in this article.

Regret theory. According to regret theory (Loomes & Sugden, 1982), people anticipate feelings of regret they may have on experiencing the outcome of a prospect. Anticipated feelings of regret modify the utility of an outcome that results from a particular choice with respect to the outcomes that would have resulted from other choices. For the simple gamble "p chance of x," the CE is such that the utilities of the outcomes, once modified by regret and summed over all world states, are equal for the prospect and the CE. As in the independent theories, the CE options on offer simply do not enter into the equation, and thus, regret theory cannot account for the results of Experiments 1A-1D and 2.

Because, in regret theory, the utility of a prospect is not independent of the other prospects in the choice set, it seems that regret theory might be able to offer an account of the context effect in Experiment 4. Unfortunately, with 10 independent binary prospects (as in the free-choice condition), there are $2^{10} = 1{,}024$ possible world states, each with a different pattern of possible outcomes, and thus, it is not obvious what the predictions of regret theory would be. We therefore simulated the results of Experiment 4, assuming utility to be a power function of money and regret a power function of the difference in the actual outcome and the best outcome that could have occurred (following Quiggin, 1994; but see Loomes & Sugden, 1982, for an alternative extension to multiple prospects.) For every point in the parameter space, if regret theory predicts a midset prospect is preferred in one restricted set, then, in the other set, the nearest extreme prospect is preferred. Roughly, the pattern of preference for a restricted choice set can always be predicted from the pattern across a free choice of all prospects. In summary, at least for this implementation of regret theory, the context effects in Experiment 4 cannot be predicted. This is primarily because although outcomes are judged, at least in part, relative to an anchor defined by the choice set, probabilities are not.

Stochastic difference model. In the stochastic difference model (González-Vallejo, 2002), prospects are judged relative to one another. For simplicity, González-Vallejo assumed that subjective prospect attributes are the actual prospect attributes and that the function comparing attributes gives the difference between them as a proportion of the larger attribute. This proportional difference strategy is a special case of the stochastic difference model. (No other instantiations of the theory have been investigated.) The proportions are summed over all attributes to give the overall preference for one prospect over another. This model can account for violations of stochastic dominance, independence, and stochastic transitivity and thus seems a plausible candidate model to account for the context effects presented in this article.

The stochastic difference model is primarily a model of choice. It is not obvious how it could be extended to produce CEs. Here, we assume that the CE is a prospect of the form "*y* for certain" where the model predicts no preference for the CE over the prospect "*p* chance of *x*" under consideration. There is no preference for the prospect over the CE when the proportion difference in the probabilities is equal to the proportion difference in amounts. Thus, the model predicts risk neutrality where the CE is the expected value of the prospect and offers no account of the data from Experiments 1A-1D and 2.

Preliminary suggestions have been given (González-Vallejo, 2002, p. 152) as to how the model might be extended to choice among multiple prospects using the notion of trade-off contrast (Simonson & Tversky, 1992) in a two-step procedure. First, the strengths of preference for one prospect over another are calculated for all pairwise comparisons within the set of prospects. The overall preference for a given prospect is then the sum of all of the pairwise strengths where that prospect was favored. The extended model can be applied to our Experiment 4 as follows.

The stochastic difference model predicts that, for any pair of prospects from Experiment 4 (from a "50% chance of £50" to a "5% chance of £95"), the more risky prospect would be favored. This is because the proportional difference in probabilities is smaller than the proportional difference in money for all pairwise combinations of prospects. Averaging across all pairwise combinations in the free-choice condition, the model predicts a skew in preferences toward the more risky prospects, with a "60% chance of £40" most preferred. In the restricted-choice conditions, the skew remains, with the most risky prospect being preferred most in each case. These predictions are independent of the decision threshold (which modulates the weight placed on each attribute). However, given the closeness of the overall preference values, we think that it is unlikely that this prediction is independent of the form of the functions mapping actual attribute values into subjective values or the choice of generalization to the multiple prospect case. Thus, we conjecture that the stochastic difference model may be flexible enough to accommodate our data.

Multialternative decision field theory. Roe, Busemeyer, and Townsend (2001) extended decision field theory (Busemeyer & Townsend, 1993) to scenarios with multiple alternatives to offer an account of three key results. Consider a binary choice between two options, A and B, that vary on two dimensions, where one option might be higher on one dimension and the other option higher on the other dimension. In the similarity effect (see, e.g., Tversky, 1972), the addition of a new competitive option that is highly similar to Option A, but not Option B, can reverse a preference for A in the binary case to a preference for B in the ternary case. The attraction effect (see, e.g., Huber, Payne, & Puto, 1982) describes the increase in preference for a dominating option when an asymmetrically dominated option is added to the binary set. In the compromise effect (see, e.g., Simonson, 1989), an option that represents a compromise between two alternatives may be preferred over the alternatives in the ternary choice even though it was not preferred in either pairwise binary choice.

Multialternative decision field theory is able to offer an account of the similarity, attraction, and compromise effects using two key mechanisms. First, attribute values are compared across options, and these (weighted) differences are summed across dimensions to produce what Roe et al. (2001) termed momentary valences for each option. The relative weight for each dimension is assumed to vary over time. Preferences are constructed for each option by integrating valences over time. This process contrasts with the accumulation of absolute attribute values. Instead, valences represent the "comparative affective evaluations" (Roe et al., 2001, p. 387). Thus, the choice between options is made in relative rather than absolute terms, as in the stochastic difference model. The second key mechanism is the competition of valences via lateral inhibitory connections such that preferences for more similar options compete more.

There are two natural representations of the simple prospects used in Experiment 4.[8] First, the probability of winning and the amount to win can be considered as separate attributes for each prospect. In this case, the valences for the less risky set (when attending to either the win amount or the win probability) are the same as those for the more risky set. This is because it is the location of the prospects in the space relative to one another that determines their associated valences rather than their absolute location. Thus, multialternative decision field theory predicts that the pattern of preferences should be the same across the less risky and more risky conditions. In other words, the theory predicts pure context effects. Multialternative decision field theory also predicts a tendency to prefer the central prospects in a set in the same way that it predicts the compromise effect.

The second natural representation of the prospects uses a single attribute representing the subjective EU of each prospect. This representation might be considered more plausible as it seems rather odd, perhaps, to represent the probability of an outcome happening in the same way as actual outcome attributes (e.g., price, quality). Following the original decision field theory, weights no longer represent the strength of attention to an attribute at a given moment. Instead, weights represent the strength of attention to a world state (as the subjective probabilities in subjective EU theory do). Thus for a given prospect, the valence is the difference in subjective EU for that prospect and the average subjective EU for all remaining prospects. This implies that within a given context, the pattern of valences is the same as the pattern of actual subjective EUs. Thus, in the same way that EU cannot predict any context effects, neither can multialternative decision field theory using this second representation.

The componential-context model. Tversky and Simonson (1993) presented a model of context dependent preference that is a generalization of the contingent weighting model (Tversky et al., 1988). The model was devised to provide an account of trade-off contrast and extremeness aversion (Simonson & Tversky, 1992). Each attribute of an object has a subjective value depending on its magnitude. The value of an option is a weighted sum of its attribute values. The effect of a previous choice set (i.e., the background context) is to modify the weighting of each attribute according to the trade off between attributes implicit in the background context. The value of an option is then modified by the relative value of the option averaged over pairwise comparisons with the other options in the choice set.

Tversky and Simonson (1993) did not apply their model to choices between risky prospects. We consider the representation where probability is simply represented as any other option attribute, as we did for multialternative decision field theory. The effects of choice set in Experiment 4 can then be accounted for as another example of extremeness aversion. Specifically, the componential-context model explains the pattern by assuming that losses on the value of one attribute loom larger than gains in the value of another attribute as the two attributes are traded off, and thus, a central compromise option, where the overall loss is minimized, is preferred. An alternative representation with a single dimension for the outcome and probabilities determining the weighting of that outcome reduces to something like regret theory, and therefore, we do not consider it further.

Range frequency theory. Range frequency theory (Parducci, 1965, 1974) predicts how items that vary along a single dimension will be valued or rated. The subjective value given to an attribute is a function of its position within the overall range of attributes and its rank. Thus, attributes are judged purely in relation to one another. Specifically, the subjective value $S(x)$ of an attribute x from the ordered set $\{x_1, ..., x_i, ..., x_n\}$ is given by

$$S(x) = c\frac{i-1}{n-1} + (1-c)\frac{x_i - x_1}{x_n - x_1} \quad , \qquad (3)$$

where c is a parameter that specifies the relative contributions of rank and range. Increasing all of the attributes by a constant value or, alternatively, increasing the spacing of all of the attributes would not change their position within the range or their rank, and thus, according to range frequency theory, their subjective value should remain unaltered.

There is some precedent for using range frequency theory to account for context effects in decision under risk. Birnbaum (1992) found his data to be consistent with the theory. Recall that he investigated the effect of skewing the values of options offered as CEs for simple prospects. The subjective value of a given option is larger in the positive skew condition because the option has a higher rank thanks to the presence of many smaller options. This is consistent with the finding that, when options were positively skewed, prospects were assigned smaller CEs, compared with the case where options were negatively skewed. A similar explanation accounts for the results of Mellers et al. (1992) described in the introduction.

In the experiments presented here, the stimuli can be considered to vary along a single risk dimension. For example, in Experiment 1, selecting an option at one end of the range represents a risk-averse choice, whereas selecting an option at the other end represents a less risk-averse choice. In Experiment 4, the prospects in the set varied from risky ("50% chance of £50") through to safe ("95% chance of £5"). If it is assumed that people are poor at making judgments about the absolute risk attached to each choice, then they may instead make relative judgments (cf. the evaluability hypothesis; Hsee, Loewenstein, Blount, & Bazerman, 1999) of the type described by range frequency theory. Although such relative comparisons allow people to evaluate which options are more risky than others, and even by how much, they do not provide information on how risky the overall set is; the options in the set may all be relatively low risk, all be relatively high risk, or span the entire range of risk.

In a range frequency account, an individual's level of risk aversion would be represented by a relative risk aversion parameter (see Wernerfelt, 1995, for a related parameterization based only on rank). This parameter corresponds to a preferred range frequency subjective value. A relatively risk-averse person, for example, would have a low value, suggesting that he or she prefers prospects at a relatively less risky position within the total range of risk, and a low-ranking prospect when prospects are ranked by risk. Because, according to range frequency theory, prospects are judged in relative terms, an individual parameterized in this way would display pure context effects of the sort seen in Experiments 1A and 4.

Range frequency theory is also able to offer an account of several details in the experimental findings presented here. First, there was a tendency to prefer the highest option in the low CEs condition and the lowest option in the high CEs conditions of the within-participants Experiments 1B and 1C that was absent in the between-participants Experiment

1A. If it is assumed that some small memory of previous options is maintained across trials, then, for example, the cumulative effect of all the CE options from the low set is to add an anchor value below options of the high set. This increases the rank and position up the range of the high CE options and thus increases their range frequency scores, causing participants to prefer a lower option than they might otherwise prefer. Second, when we examined between-trials context effects, we found them to be small or nonsignificant. If it is assumed that there is only a slight memory for previous trials (as above) then there is no reason to expect large between-trials effects. In Experiment 3, for example, the cumulative effect of many previous trials would not have been additive, as we supposed in Experiments 1B and 1C because there were many types of trial in Experiment 3 but only two types in Experiments 1B and 1C. In Experiment 5, each participant did see two different types of trial. Here, for example, we might expect a cumulative effect of the safe fillers. Although the difference was not significant, it was in the hypothesized direction. Third, the range frequency account also explains why there should be a correlation between the rank of the options picked between the low- and the high-option sets of Experiments 1B and 1C. If an individual is parameterized by a preferred range frequency subjective value, this should lead to the individual's selecting options with the same range frequency score between the two conditions. An individual with a low parameter value should select the lower risk options in all sets, and an individual with a high value should select the higher risk options in all sets.

It is not obvious how range frequency theory could be applied to the wide variety of key phenomena in the literature. Choice sets do not always involve a trade-off between risk and return that is as transparent as it was in our studies. Furthermore, many studies involve only binary choices, and so, a key challenge is to extend range frequency theory to this situation. Transforming prospect attributes before combining them (as in multialternative decision field theory and the stochastic difference model) is one possibility we are currently investigating.

Summary of Theoretical Accounts

Theories where prospects are valued independently of one another, such as EU theory, prospect theory, configural weight theory, and rank-dependent utility theory, must, by definition, fail to predict context effects of the sort reported here. When prospects are judged in relation to one another, as in the stochastic difference model, multialternative decision field theory, and range frequency theory, the effect of choice set can, under some circumstances, be predicted. These relational theories all have in common the idea that preferences are constructed for a given choice set (see Slovic, 1995). We offered a detailed account in terms of range frequency theory; however, similar mechanisms might well be incorporated into the other models. Regret theory and the componential-context model can be considered hybrid theories where utilities derived independently for each prospect are modified depending on their relationship to other prospects in the choice set.

There are two ways in which to view the challenge to theories of decision under risk that cannot explain the prospect relativity effects shown in this article. First, assume that the theory is correctly representing the underlying decision process and that the context effects demonstrated here represent a biasing of judgments. We discuss this possibility below. However, if people are subject to such biases in making everyday decisions - and we see no reason why they should not be - then the descriptive theories should be revised to provide an account of these effects (see Tversky & Simonson, 1993, for a similar point). Given the large size of the effects, there is a second possibility that should be given some consideration: that the models are inadequate and should be rejected. It is too early to say which of these possibilities is correct. Hybrid models where an underlying EU-type decision process is biased by the context may prove adequate. Alternatively, purely relative models where judgments about prospects are made relative to the choice set and other anchors may be extended to account for the classic phenomena that traditional models describe.

Conversational Pragmatics

An important question concerns whether the prospect relativity effect involves reasoning about the experimenter's intentions. Is it critical that participants view the options they are given as provided by a cooperative experimenter and hence infer that their response should naturally fall within that range? There is some evidence to support this possibility. For example, Schwarz, Hippler, Deutsch, and Strack (1985) asked participants to report the number of hours they spent watching television each day. Half of the participants were given a scale that varied from "up to 1/2 an hour" to "more than 2 1/2 hours," and half were given a scale that varied from "up to 2 1/2 hours" to "more than 4 1/2 hours." Twice as many respondents claimed to have watched less than 2 1/2 hours of television per night with the latter scale (16% vs. 38%). Schwarz (1994) reported that the effect of response alternatives completely disappears when the informational value of the scale is removed (e.g., by saying it is a pretest to explore the adequacy of the response alternatives). If this pragmatic explanation is correct, then we might explain the performance that we observed as follows: People have a weak grip on a notion of the utility of a risky option, but they may take the options available as a clue from the experimenter. They may, for example, assume that the experimenter has chosen the options so that each will be the choice of some experimental participant. Then, if a participant judges that he or she is, for example, slightly happier with risk than the average participant, he or she may decide to choose a value slightly higher than the average option available. Accordingly, context would be expected to play a substantial role in determining participants' choices. This would build connections between the current work and pragmatic theory in linguistic communication (see, e.g., Grice, 1975; Levinson, 1983).[9]

Anchoring Effects

Alternatively, though, it may be that the set of available options merely primes participants' choices in a way that is insensitive to intentional factors. Tversky and Kahneman (1974) have demonstrated large effects of standard or anchor values in judgment. Estimates are typically assimilated toward the anchor provided, even if participants know that the anchors have been randomly selected. Use of randomly selected anchors makes it unlikely that participants take their inclusion to be informative. Furthermore, such effects are evident even for quite implausible anchor values (see, e.g., Chapman & Johnson, 1994). The more uncertain participants are about a judgment the more their estimates are assimilated toward the anchor value (Jacowitz & Kahneman, 1995).

Recent research provides evidence that the anchoring effect is largely caused by the retrieval of relevant semantic information, rather than numerical priming (Mussweiler & Strack, 2001b; but see Wong & Kwong, 2000). When the use of semantic information is prohibited by a change in the target between an initial comparative judgment containing the anchor and a subsequent absolute judgment, only small numerical anchoring effects remain (Mussweiler & Strack, 2001b). When dealing with anchors, participants are hypothesized to create an anchor-consistent mental model of the target (Mussweiler & Strack, 1999, 2000; Strack & Mussweiler, 1997). In Experiments 1A-1D and Experiment 2, options offered as potential CEs had a large effect on the option chosen. If the options were acting as anchors, then, according to Mussweiler and Strack, participants tested the hypothesis that each option was the CE, and this testing process assimilated the judgment of the CE toward the options. It is less obvious how the selective accessibility model might account for the choice results in Experiment 4.

One way to test between anchoring and conversational-pragmatic explanations would be to repeat the experiments here under conditions where participants believe that the ranges of choices are generated randomly (e.g., by spinning a roulette wheel or similar device). Thus, the participants could not reasonably attribute these choices to a cooperative experimenter. If the effects described here are intentionally mediated, we would expect the context effects to be eliminated; if they result from nonintentional factors, then they should remain unchanged.

Methodological Implications

The results presented here have implications for the empirical work on descriptive theories of decision making under risk. There are various procedures that are typically used as part of this research, but here we divide them into two broad categories based on the type of questions they use. First, there are those procedures that require participants to select between two prospects: choice experiments. Second, there are those that elicit a quantity that causes a participant to be indifferent between two prospects: certainty equivalence experiments. A brief directional poll of 38 such studies finds that 19 fall into the choice category, 16 are based on certainty equivalence, and 3 combine the two.

We have found small sequential effects in binary choice experiments from our laboratory (see Experiment 5 and the meta-analysis in the *Discussion* section of Experiment 5). It is possible that the set of prospects used may have had a small effect on participants' choices. We therefore conclude that this should be the subject of further research, particularly when considering the relevance to at least half the existing experimental literature.

For those studies concerned with certainty equivalence, about one third presented participants with an array of options to choose between rather than relying on participants to generate their own amounts. For example, Tversky and Fox (1995) elicited CEs using a series of choices. Participants made a series of choices between a prospect and six sure payments offered one at a time in descending order. These were roughly spaced between the highest outcome available in the risky prospect and $0. Then, seven further options were presented, spanning the narrower range between the lowest payment the participant had accepted and the highest payment rejected by the participant. Although the presentation of payments in a sequence may have mitigated the influence of the selected scale, the fact that participants could backtrack if they felt they had made a mistake and the general transparency of the procedure could also have induced participants to consider the options simultaneously. Indeed, Loomes (1988) has found differences between CEs elicited through such a sequence of choices and those that are generated by participants independently.

If participants' CEs are affected by the set of choices presented, then it could be countered that when the set is generated using some reasonable and lawful process, in some sense, context has been held constant across the experiment. However, this defense does not seem satisfactory: There is no reason for us to assume that these effects will be stable. For example, Tversky and Kahneman (1992) used a similar sequencing approach to that described for Tversky and Fox (1995). Although Tversky and Kahneman also used seven sure options, spanning the extreme outcomes of the relevant risky prospect, the options were spaced logarithmically. Furthermore, CEs were elicited for a wide variety of outcome ranges, including both gains and losses and differing scales of money. Because all these CEs were then used in combination to fit curves and to test cumulative prospect theory, it seems unlikely that the effects of the different contexts during the elicitation process would have acted in the affine manner necessary to preserve the findings. For other examples of potentially distorted CEs, see M. Cohen, Jaffray, and Said (1987); Hershey and Schoemaker (1985); and Lichtenstein, Slovic, and Zink (1969).

The findings we have described have potentially important implications for practical survey research methods that are used to elicit people's utilities. One implication, which we have already discussed, is that providing sets of response options can substantially, if unintentionally, bias people's responses. Even if people are not presented with options - presenting options is typically avoided in the measurement of utility (see Baron, 2000, for a review) - they may still construct anchors to help make a decision (cf. the construction of preference, Slovic, 1995). Indeed, this is the method by which Mussweiler and Strack (2001a) hypothesized that implausible anchor values affect subsequent judgments (see *Anchoring Effects* above). Constructed anchors may produce effects of the sort seen in our experiments. Presumably, these effects vary randomly between participants.

However, our results may also have a second and more fundamental implication concerning the existence of an underlying scale of utility for risky prospects. The present studies were inspired by Garner's (1954) study showing that people's judgments of what sound is half as loud as another sound can be dramatically manipulated by choosing different response alternatives. Laming (1997) argued that this result is the strongest single piece of evidence against the notion that there is an underlying internal psychophysical scale for loudness, an assumption that has been taken for granted in much psychophysical research (see, e.g., Stevens, 1975). If Laming's reasoning is correct, it would seem that the present data provide equal difficulties for the idea that there is an underlying internal utility scale for risky prospects. If this is the case, then the project of asking people to make judgments about risky prospects may be ill founded, and hence, decision analysis methods that typically involve such judgments (see Baron, 2000, for a review) may also be difficult to interpret. Thus, although methods for measuring utility may not be subject to the kinds of effects shown in the experiments in this article (indeed, care is taken to avoid these sorts of effects when measuring utility), to the extent that these experiments are problematic for the concept of an underlying scale of utility, the enterprise of measuring utility may be challenged. For example, contingent valuations studies (Cummings, Brookshire, & Schulzze, 1986; Mitchell & Carson, 1989), which are widespread in environmental economics, frequently require people to assign a value to some risky outcome (the possibility of a nuclear accident or an oil spill), which they may be unable to do.

If this line of reasoning is accepted, then a natural recommendation is to elicit preferences for simple outcomes rather than risky prospects. Related problems may, of course, arise to the degree that people evaluate simple outcomes by considering prospects over their consequences. Yet it is also possible that the prospect relativity effects that we have found here do not apply only in the context of risk. It is possible that similar effects might be found even where people make direct judgments about simple outcomes (e.g., judging that one injury is twice as bad as another or that a level of disability is a certain fraction along a continuum between normal health and death). If response options can radically affect people's judgments in contexts of these kinds, then concerns must at least be raised over the existence of an underlying subjective utility that can be elicited. The discrepancy between subjective utility measured using different methods (see Baron, 2000, pp. 330-333) is not reassuring. If the concept of a subjective utility that can be elicited is not to be abandoned, then, at minimum, important theoretical developments are required to show how utility-based models can be flexible enough to capture these effects.

There is already some evidence that challenges the notion of an underlying subjective probability scale. Slovic and Monahan (2000) investigated risk perceptions in mental health law and found that probability judgments were quite malleable. Participants read vignettes describing various attributes of a person supposedly interviewed by a psychiatrist because of mental illness. Participants estimated "the probability that this person will harm someone else during the three years following the examination" (p. 362) and whether this person should be described as dangerous. Two sets of probability options were used. (This experiment, then, is analogous to our Experiments 1A-1D, and so, we describe it in some detail.) In the large-probability condition, options ran from 0% to 100% in increments of 10%. In the small-probability condition, there were 13 options, starting with < 1/1000, with increasing increments up to > 40%. Both lay and psychiatric professionals exhibited strong effects of the options available. In the small-probability condition, the six categories that were less than or equal to 10% attracted 67.7% of the vignettes. This compares with 10.8% in the large-probability condition. Likewise, the mean probability of doing harm was judged as 44% in the large-probability condition and 12% in the small-probability condition. Furthermore, although the effect was smaller, there was also a carryover consequence from the probability condition to the participants' judgments of dangerousness. In the small-probabilities condition, 30.5% of the vignettes were described as dangerous versus 37.0% for large-probabilities condition. That similar results were also found for participants' judgments on whether a patient should be hospitalized, even if it involved coercion, dramatically highlights the social importance of context effects in decision making.

Finally, Is Utility Like Perceptual Psychological Dimensions?

The experiments presented here were motivated by the large effects of accompanying and recent stimuli in psychophysical judgment. There are differences, however, between the perceptual and decision domains. The most obvious difference is that here we found only very small between-trials effects. Thus, although range-frequency theory can offer an account of context effects in the domains of both decision under risk and psychophysics, we suggest that this is only because of the relativity of judgment in each. The causes of the reliance on this relativity may be quite different: In decision under risk, absolute attribute values are available as numbers. The problem seems to come in integrating information across the attributes. In perception, there seems to be only partial availability of absolute magnitude information from the senses.

References

Akerlof, G., & Yellen, J. (1985). Can small deviations from rationality make significant differences to economic equilibria? *American Economic Review, 75,* 708-720.

Arrow, K. J. (1971). *Essays in the theory of risk-bearing.* Chicago: Markham.

Baron, J. (2000). *Thinking and deciding* (3rd ed.). Cambridge, England: Cambridge University Press.

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

Bell, D. E., & Fishburn, P. C. (1999). Utility functions for wealth. *Journal of Risk and Uncertainty, 20,* 5-44.

Birnbaum, M. H. (1992). Violations of monotonicity and contextual effects in choice-based certainty equivalents. *Psychological Science, 3,* 310-314.

Birnbaum, M. H., Patton, J. N., & Lott, M. K. (1999). Evidence against rank-dependent utility theories: Tests of cumulative independence, interval independence, stochastic dominance, and transitivity. *Organizational Behavior and Human Decision Processes, 77,* 44-83.

Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect of preference-reversal phenomenon of using choice indifferences. *Journal of Economic Behavior and Organization, 13,* 193-212.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432-459.

Camerer, C. F. (1995). Individual decision making. In J. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 587-703). Princeton, NJ: Princeton University Press.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral*

Decision Making, 7, 223-242.

Christenfeld, N. (1995). Choices from identical options. *Psychological Science, 6,* 50-55.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Cummings, R. G., Brookshire, D. S., & Schulzze, W. D. (Eds.). (1986). *Valuing environmental goods: A state of the arts assessment of the contingent method.* Totowa, NJ: Rowman & Allanheld.

de Canio, S. (1979). Rational expectations and learning from experience. *Quarterly Journal of Economics, 93,* 47-57.

Diecidue, E., & Wakker, P. P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty, 23,* 281-298.

Fishburn, P. C., & Kochenberger, G. A. (1979). Two-piece von Neumann-Morgenstern utility functions. *Decision Sciences, 10,* 503-518.

Friedman, M. (1953). *Essays in positive economics.* Chicago: University of Chicago Press.

Friedman, M., & Savage, L. J. (1948). The utility analysis of choices involving risks. *Journal of Political Economy, 56,* 279-304.

Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology, 46,* 373-380.

Garner, W. R. (1954). Context effects and the validity of loudness scales. *Journal of Experimental Psychology, 48,* 218-224.

González-Vallejo, C. (2002). Making trade-offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review, 109,* 137-154.

Grether, D. M., & Plott, C. R. (1979) Economic theory of choice and the preference reversal phenomena. *American Economic Review, 38,* 129-166.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3: Speech acts* (pp. 41-58). New York: Academic Press.

Hershey, J. C., & Shoemaker, P. J. H. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science, 31,* 1213-1231.

Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences, 24,* 383-451.

Holland, M. K., & Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception and Psychophysics, 3,* 409-414.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin, 125,* 576-590.

Hu, G. (1997). Why is it difficult to learn absolute judgment tasks? *Perceptual and Motor Skills, 84,* 323-335.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9,* 90-98.

Jacowitz, K. E., & Kahneman, D. (1995). Measures of anchoring in estimation tasks. *Personality and Social Psychology Bulletin, 21,* 1161-1166.

Janiszewski, C., & Lichtenstein, D. R. (1999). A range theory account of price perception. *Journal of Consumer Research, 25,* 353-368.

Jesteadt, W., Luce, R. D., & Green, D. M. (1977). Sequential effects of the judgments of loudness. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 92-104.

Kagel, J. H., & Roth, A. E. (Eds.). (1995). *The handbook of experimental economics.*

Princeton, NJ: Princeton University Press.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291.

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values and frames.* New York: Cambrige University Press & the Russell Sage Foundation.

Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: A response-time analysis. *Psychological Research, 60,* 121-133.

Laming, D. R. J. (1984). The relativity of "absolute" judgements. *British Journal of Mathematical and Statistical Psychology, 37,* 152-183.

Laming, D. R. J. (1997). *The measurement of sensation.* London: Oxford University Press.

Levinson, S. C. (1983). *Pragmatics.* Cambridge, England: Cambridge University Press.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89,* 46-55.

Lichtenstein, S., & Slovic, P. (1973). Response induced reversals of preference in gambling: An extended replication in Las Vegas. *Journal of Experimental Psychology, 101,* 16-20.

Lichtenstein, S., Slovic, P., & Zink, D. (1969). Effects of instructions in expected value on optimality of gambling decisions. *Journal of Experimental Psychology, 79,* 236-240.

Lockhead, G. R. (1984). Sequential predictors of choice in psychophysical tasks. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 27-47). Hillsdale, NJ: Erlbaum.

Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 461-473.

Long, L. (1937). A study of the effect of preceding stimuli upon the judgment of auditory intensities. *Archives of Psychology (New York), 30,* 209.

Loomes, G. (1988). Different experimental procedures for obtaining valuations for risky actions: Implications for utility theory. *Theory and Decision, 25,* 1-23.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal, 92,* 805-824.

Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches.* Mahwah, NJ: Erlbaum.

Luce, R. D., Nosofsky, R. M., Green, D. M., & Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception & Psychophysics, 32,* 397-408.

MacCrimmon, K. R., Stanbury, W. T., & Wehrung, D. A. (1980). Real money lotteries: A study of ideal risk, context effects, and simple processes. In T. S. Wallsten (Ed.). *Cognitive process in choice and decision behavior* (pp. 155-177). Hillsdale, NJ: Erlbaum.

Mellers, B. A., Ordóñez, L. D., & Birnbaum, M. H. (1992). A change-of-process theory for contextual effects and preference reversals in risky decision making. *Organizational Behavior and Human Decision Processes, 52,* 311-369.

Mitchell, R. C., & Carson, R. T. (1989). *Using surveys to value public goods: The contingent valuation method.* Washington, DC: Resources for the Future.

Mussweiler, T., & Strack, F. (1999). Hypothesis-consistent testing and semantic priming in the anchoring paradigm: A selective accessibility model. *Journal of Experimental Social Psychology, 35,* 136-164.

Mussweiler, T., & Strack, F. (2000). Numeric judgments under uncertainty: The role of knowledge in anchoring. *Journal of Experimental Social Psychology, 36,* 495-518.

Mussweiler, T., & Strack, F. (2001a). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition, 19,* 145-160.

Mussweiler, T., & Strack, F. (2001b). The semantics of anchoring. *Organizational Behavior and Human Decision Processes, 86,* 234-255.

Nelson, R., & Winter, S. (1982). *An evolutionary theory of economic capabilities and behavior.* Cambridge, MA: Harvard University Press.

Parducci, A. (1965). Category judgment: A range-frequency theory. *Psychological Review, 72,* 407-418.

Parducci, A. (1974). Contextual effects: A range-frequency analysis. In L. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 2, pp. 127-141). New York: Academic Press.

Purks, S. R., Callahan, D. J., Braida, L. D., & Durlach, N. I. (1980). Intensity perception X. Effect of preceding stimulus on identification performance. *Journal of the Acoustical Society of America, 67,* 634-637.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organisation, 3,* 323-343.

Quiggin, J. (1993). *Generalized expected utility theory: The rank-dependent model.* Boston: Kluwer Academic.

Quiggin, J. (1994). Regret theory with general choice sets. *Journal of Risk and Uncertainty, 8,* 153-165.

Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychological Review, 108,* 370-392.

Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence, and limitations. *Journal of Economic Literature, 20,* 529-563.

Schwarz, N. (1994). A judgment in a social context: Biases, shortcomings, and the logic of conversation. *Advances in Experimental Social Psychology, 26,* 123-162.

Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales: Effects of category range on reported behavior and subsequent judgments. *Public Opinion Quarterly, 49,* 388-395.

Shafir, E., & LeBoeuf, R. A. (2002). Rationality. *Annual Review of Psychology, 53,* 491-517.

Simon, H. A. (1959). Theories of decision-making in economics and behavioral science. *American Economic Review, 49,* 253-283.

Simon, H. A. (1992). *Economics, bounded rationality and the cognitive revolution.* Aldershot, England: Elgar.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16,* 158-174.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29,* 281-295.

Slovic, P. (1995). The construction of preference. *American Psychologist, 50,* 364-371.

Slovic, P., & Monahan, J. (2000). Probability, danger, and coercion. In P. Slovic (Ed.), *The perception of risk* (pp. 347-363). London: Earthscan.

Staddon, J. E. R., King, M., & Lockhead, G. R. (1980). On sequential effects in absolute judgment experiments. *Journal of Experimental Psychology: Human Perception & Performance, 6,* 290-301.

Stevens, S. S. (1975). *Psychophysics.* New York: Wiley.

Stewart, N. (2001). *Perceptual categorization.* Unpublished doctoral dissertation, University of Warwick, Coventry, England.

Stewart, N. Brown, G. D. A., & Chater, N. (2002). *Identification of simple perceptual stimuli: A new model of absolute identification.* Manuscript submitted for publication.

Stewart, N., Brown, G. D. A., & Chater, N. (2002). Sequence effects in categorization of simple perceptual stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 3-11.

Strack, F., & Mussweiler, T. (1997). Explaining the enigmatic anchoring effect: Mechanisms of selective accessibility. *Journal of Personality and Social Psychology, 73,* 437-446.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79,* 281-229.

Tversky, A., & Fox, C. R. (1995). Weighting risk and uncertainty. *Psychological Review, 102*, 269-283.

Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124-1131.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 204-217.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review, 95,* 371-384.

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science, 39,* 1179-1189.

Tversky, A., Slovic, P., & Kahneman, D. (1990). The cause of preference reversal. *American Economic Review, 80,* 204-217.

von Neumann, M., & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd ed.). Princeton, NJ: Princeton University Press.

Ward, L. M., & Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology, 84,* 27-34.

Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception & Psychophysics, 9,* 73-78.

Wernerfelt, T. (1995). A rational reconstruction of the compromise effect: Using market data to infer utilities. *Journal of Consumer Research, 21,* 627-633.

Wong, K. F. E., & Kwong, J. Y. Y. (2000). Is 7300 m equal to 7.3 km? Same semantics but different anchoring effects. *Organizational Behavior and Human Decision Processes, 82,* 314-333.

Appendix A

Free-Choice Certainty Equivalents and Options for Experiment 1A.

Prospect		Free choice		Option							
p	*x*	Mean	*SD*	L_1	L_2	L_3	L_4	H_1	H_2	H_3	H_4
.2	200	35.0	36.8	15	20	25	30	40	45	50	55
.4	200	48.2	35.1	25	30	35	40	55	60	65	70
.6	200	90.7	94.4	50	60	70	80	100	110	120	130
.8	200	100.0	64.0	60	70	80	90	110	120	130	140
.2	400	70.4	71.5	50	55	60	65	75	80	85	90
.4	400	92.9	65.5	20	40	60	80	100	120	140	160
.6	400	142.9	91.3	60	80	100	120	160	180	200	220
.8	400	148.6	104.2	70	90	110	130	170	190	210	230
.2	600	102.5	102.5	20	40	60	80	120	140	160	180
.4	600	138.6	105.1	60	80	100	120	160	180	200	220
.6	600	225.0	176.8	125	150	175	200	250	275	300	325
.8	600	269.3	176.4	175	200	225	250	300	325	350	375
.2	800	95.0	104.5	20	40	60	80	110	130	150	170
.4	800	161.4	109.7	80	100	120	140	180	200	220	240
.6	800	212.9	171.0	125	150	175	200	250	275	300	325
.8	800	324.3	192.7	225	250	275	300	350	375	400	425
.2	1,000	140.0	136.7	100	110	120	130	150	160	170	180
.4	1,000	207.9	136.2	120	140	160	180	220	240	260	280
.6	1,000	298.6	190.6	200	225	250	275	325	350	375	400
.8	1,000	382.9	267.5	200	250	300	350	450	500	550	600

Note. In selecting the options, participants whose free-choice values followed an inconsistent pattern (e.g., increasing their certainty equivalent as the amount to win was reduced) were excluded.

Appendix B

Free-Choice Certainty Equivalents and Options for Experiment 1B and the Narrow Condition of Experiment 1C.

Prospect		Free choice		Option							
p	*x*	Mean	*SD*	L_1	L_2	L_3	L_4	H_1	H_2	H_3	H_4
.2	200	40.7	19.0	16	22	28	34	46	52	58	64
.4	200	70.2	23.0	30	40	50	60	80	90	100	110
.6	200	92.1	30.2	50	60	70	80	100	110	120	130
.8	200	122.1	41.6	60	75	90	105	135	150	165	180
.2	400	71.0	32.9	30	40	50	60	80	90	100	110
.4	400	121.7	49.8	60	75	90	105	135	150	165	180
.6	400	197.1	64.9	120	140	160	180	220	240	260	280
.8	400	279.5	69.1	200	220	240	260	300	320	340	360
.2	600	80.5	47.5	40	50	60	70	90	100	110	120
.4	600	199.8	93.1	120	140	160	180	220	240	260	280
.6	600	292.9	92.0	210	230	250	270	310	330	350	370
.8	600	390.7	105.5	290	315	340	365	415	440	465	490
.2	800	133.8	73.4	55	75	95	115	155	175	195	215
.4	800	261.7	99.5	160	185	210	235	285	310	335	360
.6	800	324.3	146.0	185	220	255	290	360	395	430	465
.8	800	511.0	174.1	350	390	430	470	550	590	630	670
.2	1,000	149.5	85.6	70	90	110	130	170	190	210	230
.4	1,000	311.9	92.5	210	235	260	285	335	360	385	410
.6	1,000	444.1	165.8	305	340	375	410	480	515	550	585
.8	1,000	666.7	160.6	490	535	580	625	715	760	805	850

Appendix C

Free-Choice Splits and Options for Experiment 2

		Free-choice *y*		*y* option							
z	*p*	Mean	SD	L_1	L_2	L_3	L_4	H_1	H_2	H_3	H_4
250	.2	59.7	15.5	52	54	56	58	62	64	66	68
250	.4	82.3	20.2	66	70	74	78	86	90	94	98
250	.6	96.1	17.6	84	87	90	93	99	102	105	108
250	.8	102.9	13.5	91	94	97	100	106	109	112	115
500	.2	109.4	36.0	86	92	98	104	116	122	128	134
500	.4	170.0	43.2	142	149	156	163	177	184	191	198
500	.6	192.8	18.4	170	175	180	185	195	200	205	210
500	.8	204.7	30.2	185	190	195	200	210	215	220	225
750	.2	174.2	54.2	140	150	160	170	190	200	210	220
750	.4	240.7	66.4	200	210	220	230	250	260	270	280
750	.6	285.8	46.4	245	255	265	275	295	305	315	325
750	.8	304.1	49.0	276	283	290	297	311	318	325	332
1,000	.2	232.2	101.3	170	185	200	215	245	260	275	290
1,000	.4	323.8	49.4	274	286	298	310	334	346	358	370
1,000	.6	370.7	87.1	322	334	346	358	382	394	406	418
1,000	.8	416.9	28.4	375	385	395	405	425	435	445	455
1,250	.2	299.4	121.5	220	240	260	280	320	340	360	380
1,250	.4	404.7	99.4	345	360	375	390	420	435	450	465
1,250	.6	479.7	99.2	420	435	450	465	495	510	525	540
1,250	.8	514.7	102.9	466	478	490	502	526	538	550	562
1,500	.2	363.8	173.7	265	290	315	340	390	415	440	465
1,500	.4	499.1	121.5	440	455	470	485	515	530	545	560
1,500	.6	590.0	112.5	510	530	550	570	610	630	650	670
1,500	.8	645.7	79.8	585	600	615	630	660	675	690	705
1,750	.2	436.7	208.3	310	340	370	400	470	500	530	560
1,750	.4	585.6	121.7	505	525	545	565	605	625	645	665
1,750	.6	678.3	131.3	600	620	640	660	700	720	740	760
1,750	.8	770.7	80.6	710	725	740	755	785	800	815	830
2,000	.2	445.3	166.8	305	340	375	410	480	515	550	585
2,000	.4	659.4	123.1	580	600	620	640	680	700	720	740
2,000	.6	771.6	128.5	670	695	720	745	795	820	845	870

2,000	.8	870.6	111.6	790	810	830	850	890	910	930	950

Author Note

Neil Stewart, Nick Chater, Henry P. Stott, and Stian Reimers, Department of Psychology, University of Warwick, Coventry, England.

This work was supported by a grant from Oliver, Wyman, & Company to the Institute of Applied Cognitive Science, Department of Psychology, University of Warwick. Neil Stewart was supported by Economic and Social Research Council Grant R000239351. Nick Chater was supported by European Commission Grant RTN-HPRN-CT-1999-00065. We thank Jonathon Baron, Gordon D. A. Brown, Klaus Fiedler, and D. Stephen Lindsay for their comments and Lewis Bott, Elisabeth Pennington, and Rod Freeman for their help in running experiments.

Correspondence concerning this article should be addressed to Neil Stewart, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom. E-mail: neil.stewart@warwick.ac.uk

Footnotes

[1]This experiment differed from Birnbaum's (1992) experiment. Birnbaum manipulated the distribution of potential CE options while holding the maximum and minimum constant. Here, the spacing of potential CE options was held constant, while the maximum and minimum were manipulated.

[2]In this analysis and similar analyses in Experiments 1B-1D and 2, an alpha of .05 is particularly conservative. Both the test on L_4 and the test on H_1 must show proportions significantly below .5 to reject the hypothesis that there is no effect of context. Thus, the probability of a type I error is $.05^2$ .Furthermore, the tests we ran were two-tailed, although rejecting the null hypothesis requires showing that both proportions are below .5.

[3]We thank Jonathon Baron for suggesting this analysis.

[4]We think this reduction in variation may be due to participants being more motivated in Experiment 1B as they were paid, rather than participating for course credit as participants in Experiment 1A did.

[5]Consider the options £401, £402, £403, and £404 as CEs for the prospect "50% chance of £1,000." Much of the time, the true CE lies outside this narrow range, and thus, the extreme options should be selected much of the time.

[6]There was not time to run both free-choice and context conditions within the same population of conference guests.

[7]The choice of a power function is a reasonably standard assumption. Fishburn and Kochenberger (1979) fitted power and exponential functions and found that although there was little difference between these two functions, they both fitted the data better than a linear function. See Bell and Fishburn (1999) for a consideration of alternative functions, and Luce (2000, pp. 80-84).

[8]Roe et al. (2001) did not consider multialternative decision field theory for probabilistic outcomes.

[9]We thank Klaus Fiedler for pointing out the importance of this issue.

Table 1

Number of Participants Who Selected Each Prospect in Experiment 4

		Condition		
p	*x*	Free choice	More risky	Less risky
.50	50	8	10	
.55	45	0	3	
.60	40	0	9	
.65	35	3	4	
.70	30	3	5	
.75	25	6		8
.80	20	3		2
.85	15	1		8
.90	10	5		9
.95	5	1		3
Total		30	31	30

Note. Blank cells indicate that the prospect was not available for selection in that condition.

Figure Captions

Figure 1. The curve represents a hypothetical free-choice distribution of certainty equivalents. The leftmost set of labels (L_1 - L_4) represents the options available in the low condition, and the rightmost set (H_1 - H_4) represents those available in the high condition.

Figure 2. The proportion of times each option was chosen in Experiment 1A. (Error bars are standard error of the mean.)

Figure 3. The proportion of times each option was chosen in Experiment 1B. (Error bars are standard error of the mean.)

Figure 4. The proportion of times each option was chosen in Experiment 1C. The spacing of the options on the abscissa is to scale across the three spacing conditions. (Error bars are standard error of the mean.)

Figure 5. The proportion of times each option was chosen in Experiment 1D. The spacing of the options on the abscissa is to scale across the two spacing conditions. (Error bars are standard error of the mean.)

Figure 6. Example questions from Experiment 2.

Figure 7. The proportion of times each option was chosen in Experiment 2. (Error bars are standard error of the mean.)

Figure 8. Prospects used in Experiment 3. Curves represent contours of equal expected value.

Figure 9. Certainty equivalents for the prospect on the current trial as function of the expected value of the previous prospect.

Figure 10. The utility of simple prospects of the form "p chance of x." x and p are linearly related so increasing the probability of winning p reduces the amount won x. The different curves represent different degrees of risk aversion, from risk neutral (γ = 1.0) to very risk averse (γ = 0.2). To force the curves to lie in the same range, utility has been normalized for each curve so that maximum utility over the entire prospect set is 1 for each value of γ.

Figure 1

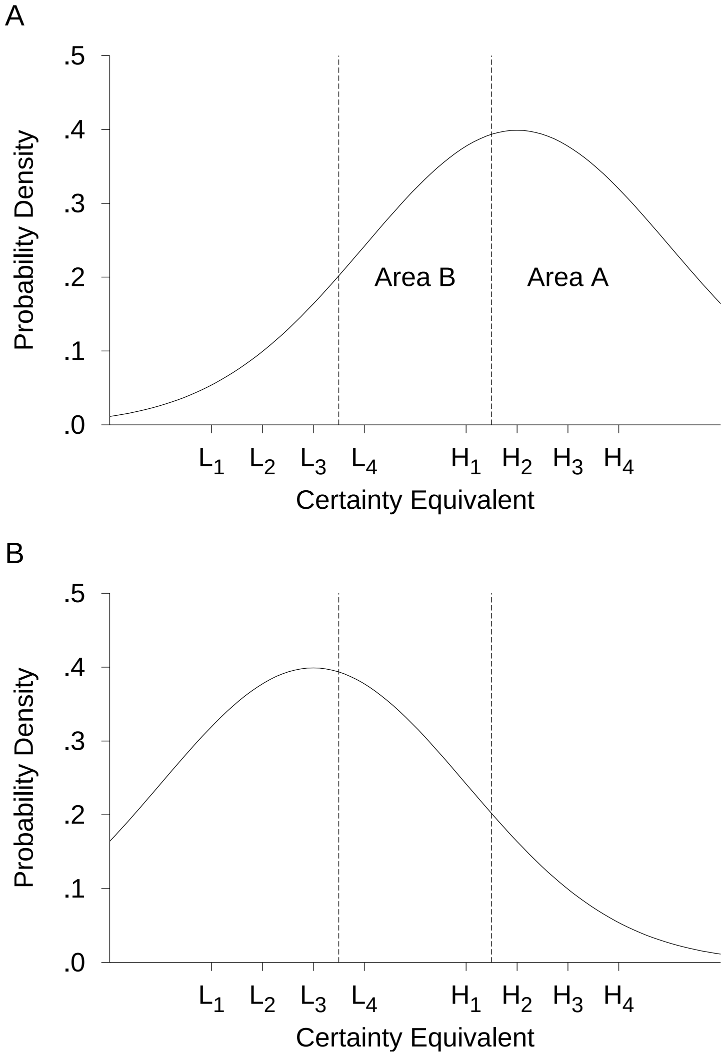

A
.5
.4
.3
.2
.1
.0
Probability Density
Area B
Area A
L1 L2 L3 L4 H1 H2 H3 H4
Certainty Equivalent
B
.5
.4
.3
.2
.1
.0
Probability Density
L1 L2 L3 L4 H1 H2 H3 H4
Certainty Equivalent

Figure 2

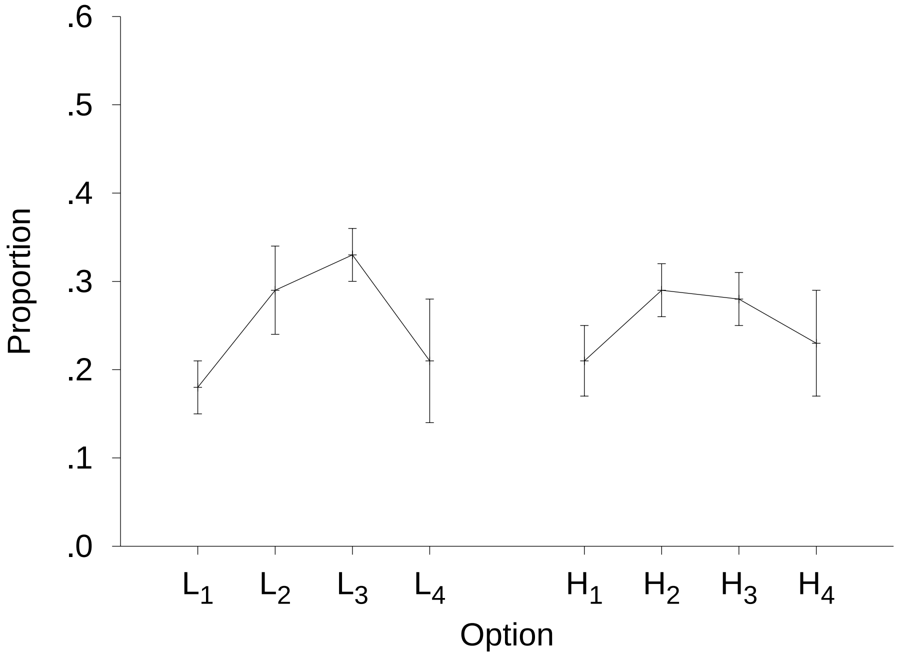


Figure 3

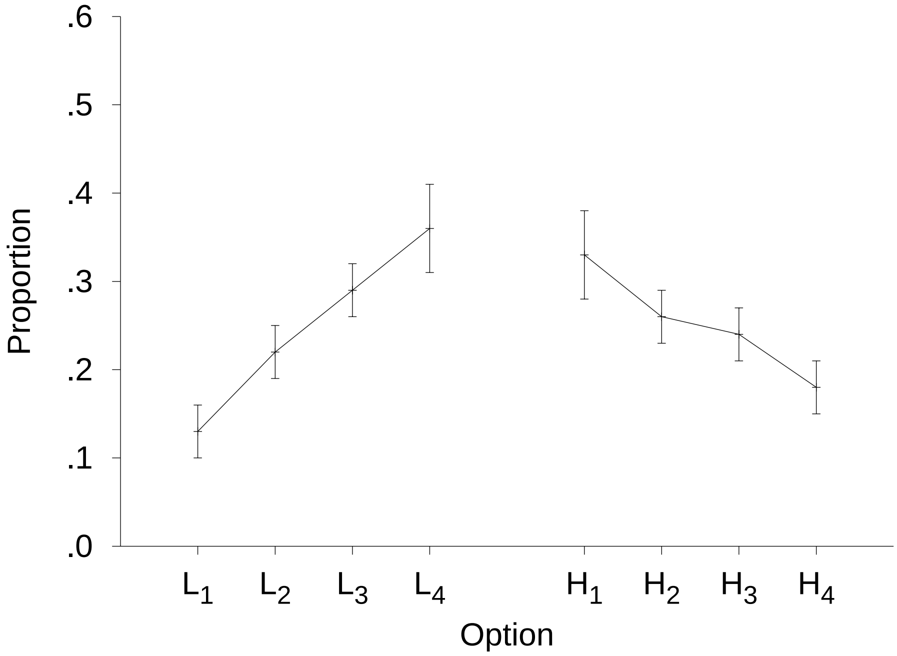


Figure 4

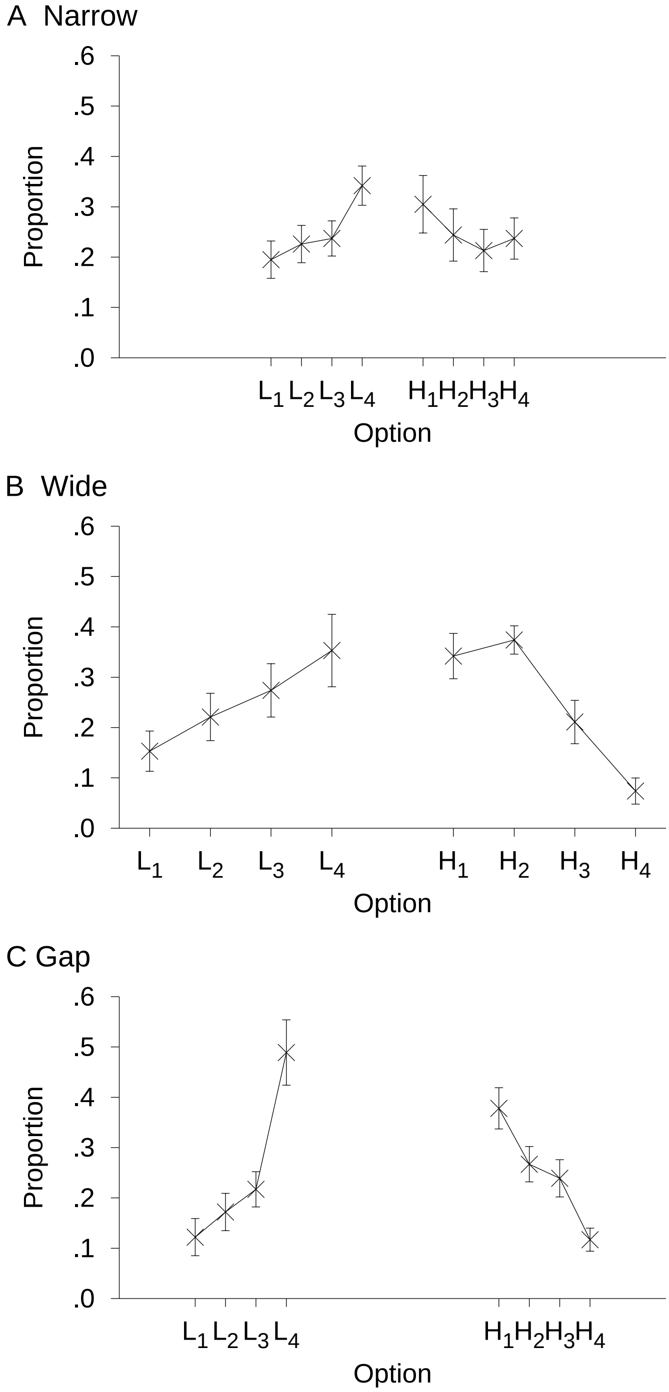

A Narrow
Proportion
.0
.1
.2
.3
.4
.5
.6
L1 L2 L3 L4 H1 H2 H3 H4
Option
B Wide
Proportion
.0
.1
.2
.3
.4
.5
.6
L1 L2 L3 L4 H1 H2 H3 H4
Option
C Gap
Proportion
.0
.1
.2
.3
.4
.5
.6
L1 L2 L3 L4 H1 H2 H3 H4
Option

Figure 5

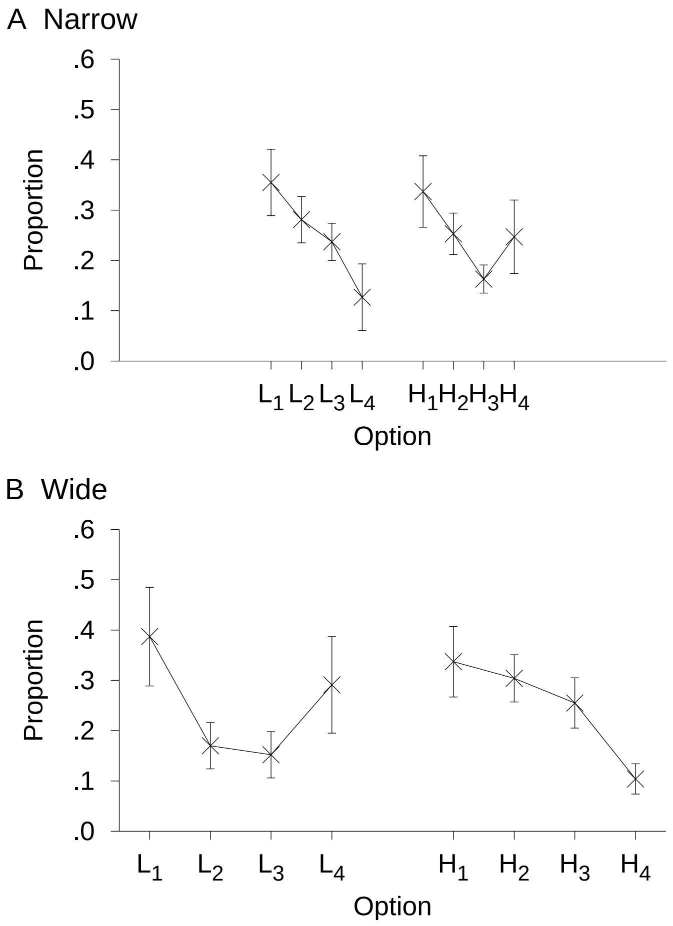

A Narrow
.6
.5
.4
.3
.2
.1
.0
Proportion
L1 L2 L3 L4 H1 H2 H3 H4
Option
B Wide
.6
.5
.4
.3
.2
.1
.0
Proportion
L1 L2 L3 L4 H1 H2 H3 H4
Option

Figure 6

A: An Example of a Free Choice Trial

£1000

£____ or 60% chance of

for certain £____

B: An Example of a Restricted Choice Trial

£1000

£322 for certain	or	60% chance of £678
£334 for certain		60% chance of £666
£346 for certain		60% chance of £654
£358 for certain		60% chance of £642

Figure 7

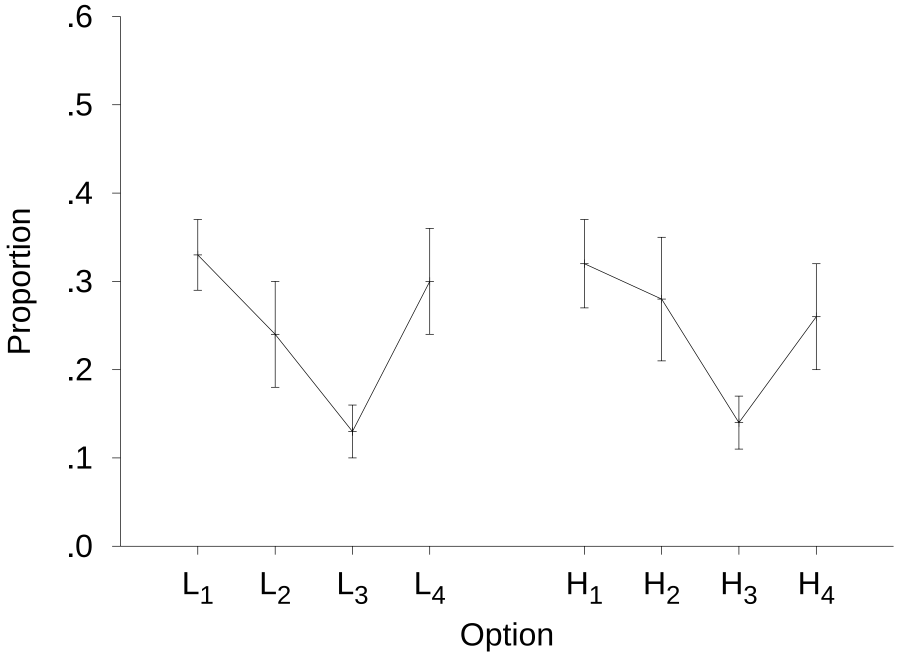


Figure 8

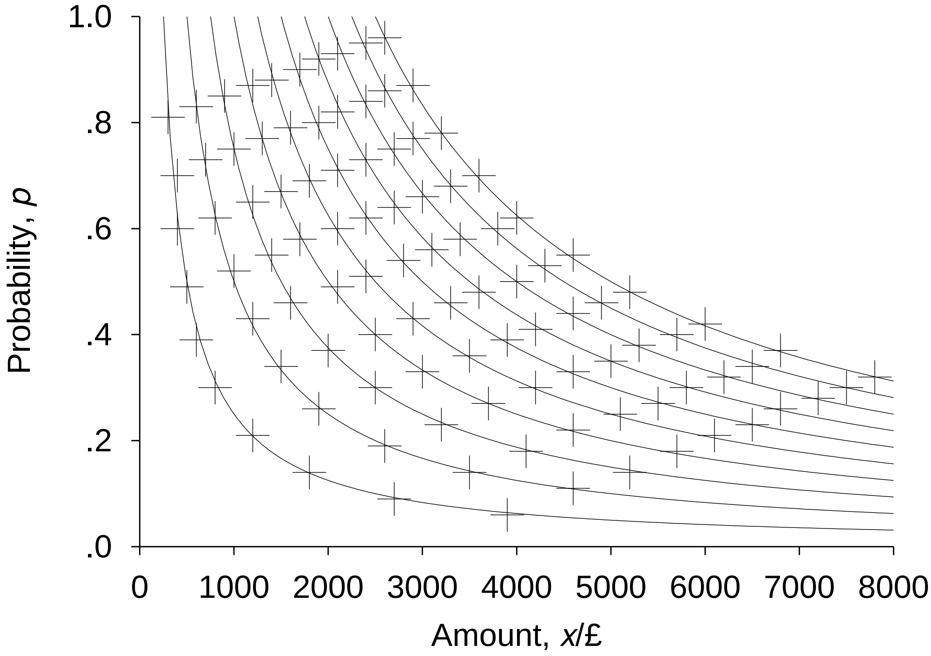

1.0
.8
.6
.4
.2
.0
Probability, *p*
0
1000
2000
3000
4000
5000
6000
7000
8000
Amount, *x*/£

Figure 9

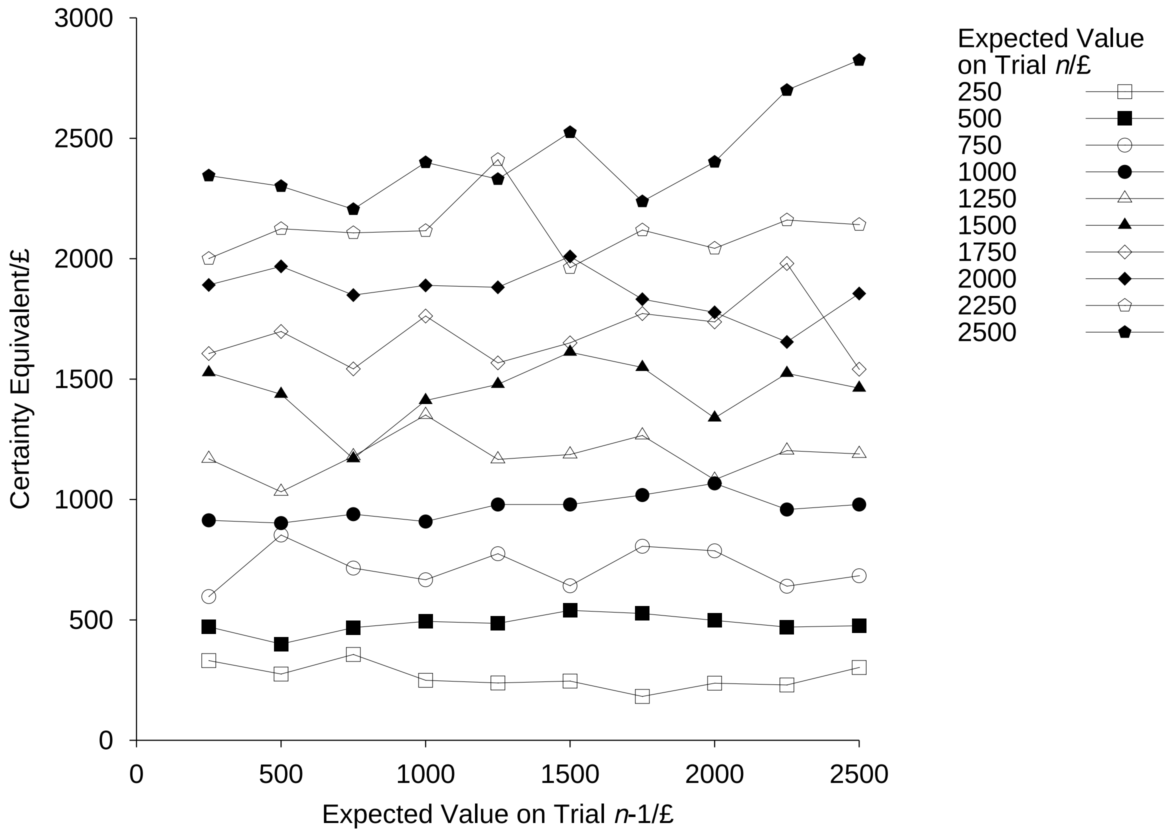
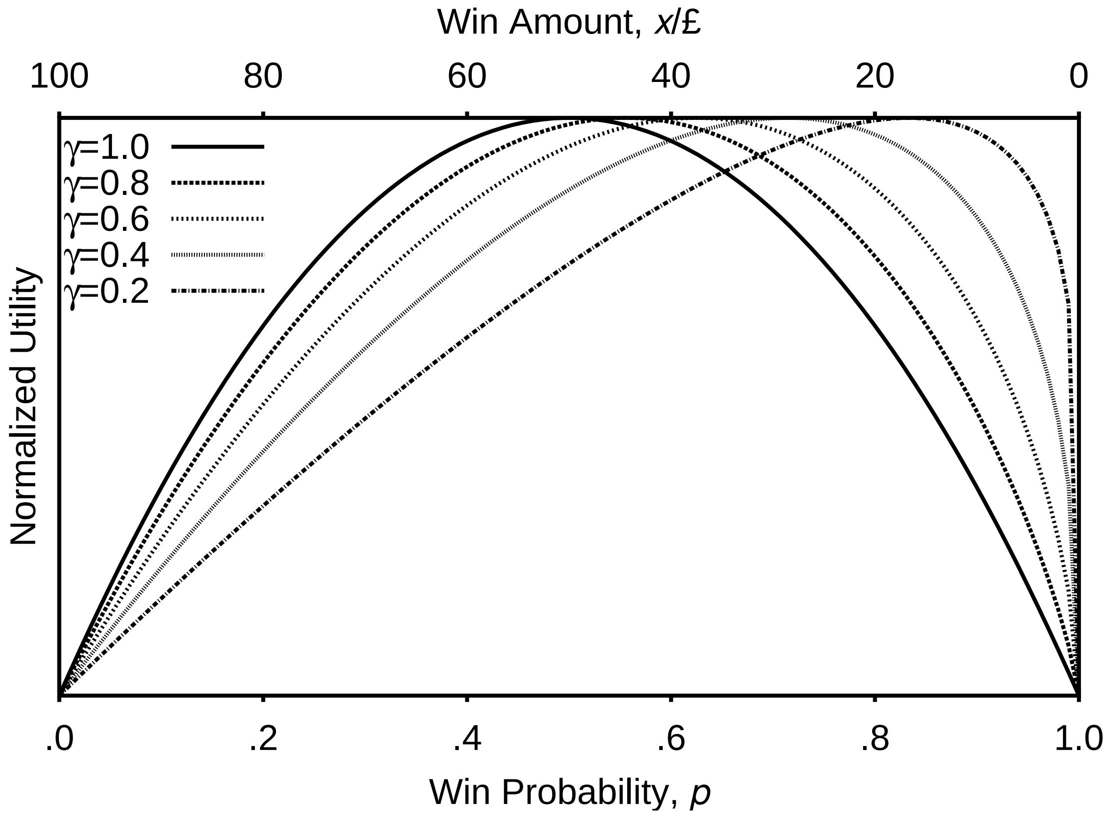

3000
2500
2000
1500
1000
500
0
Certainty Equivalent/£
0
500
1000
1500
2000
2500
Expected Value on Trial n-1/£
Expected Value
on Trial n/£
250
500
750
1000
1250
1500
1750
2000
2250
2500

Figure 10


Win Amount, x/£
100
80
60
40
20
0
γ=1.0
γ=0.8
γ=0.6
γ=0.4
γ=0.2
Normalized Utility
.0
.2
.4
.6
.8
1.0
Win Probability, p