

Bateman, Ian J.; Dent, Sam; Slovic, Paul; Starmer, Chris

Working Paper

Exploring the determinants of affect: Examining rating scale assessments of gambles

CSERGE Working Paper EDM, No. 06-07

Provided in Cooperation with:

The Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia

Suggested Citation: Bateman, Ian J.; Dent, Sam; Slovic, Paul; Starmer, Chris (2006) : Exploring the determinants of affect: Examining rating scale assessments of gambles, CSERGE Working Paper EDM, No. 06-07, University of East Anglia, The Centre for Social and Economic Research on the Global Environment (CSERGE), Norwich

This Version is available at:

<https://hdl.handle.net/10419/80246>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**Exploring the Determinants of Affect:
Examining Rating Scale Assessments of Gambles**

by

**Ian J. Bateman, Sam Dent, Paul Slovic
and Chris Starmer**

CSERGE Working Paper EDM 06-07

**Exploring the Determinants of Affect:
Examining Rating Scale Assessments of Gambles**

by

Ian J. Bateman^{1,2}, Sam Dent¹, Paul Slovic³ and Chris Starmer⁴

¹Centre for Social and Economic Research on the Global Environment (CSERGE),
University of East Anglia, UK

²Centre for Economic and Behavioural
Analysis of Risk and Decision (CEBARD), University of East Anglia, UK

³Decision Research, 1201 Oak Street, Eugene, OR 97601, USA

⁴Centre for Decision Research and Experimental Economics,
School of Economics, University of Nottingham, UK

Author contact details

Email: i.bateman@uea.ac.uk

Tel: +44 (0) 1603 593125

Fax: +44 (0) 1603 593739

Acknowledgements

The support of the Economic and Social Research Council (ESRC) is gratefully acknowledged. This work was part of the interdisciplinary research programme of the ESRC Research Centre for Social and Economic Research on the Global Environment (CSERGE).

ISSN 0967-8875

Abstract:

We undertake a series of tests examining the extent to which the affect heuristic (Slovic et al., 2002) is or is not triggered by changes in the framing and hence context of assessments of hypothetical gamble tasks using a rating scale response mode. Our initial investigations examine the replicability of findings using the same parameters used by Slovic et al. We then examine the robustness of the affect phenomena across a number of previously unexplored dimensions. These latter tests can be subdivided into two types. First we consider the persistence of affective responses to changes in the parameters of a given gamble. Second, we consider whether or not certain individuals are relatively more or less 'immune' to the affect heuristic as a decision making rule. We find that the affect heuristic is a commonly applied decision rule in a variety of contexts.

Keywords:

Affect heuristic, evaluability, gambles, rating scale, personal characteristics, gender, education

1. BACKGROUND: AFFECT AND THE EVALUABILITY PRINCIPLE

In this paper, following Slovic et al (2002), we use the term *affect* to refer to experienced feeling states associated with positive or negative qualities of a stimulus. Slovic et al. (2002) present a wide range of evidence supporting the notion that images, marked by positive and negative affective feelings, guide judgment and decision making. In light of this they propose that people use an *affect heuristic* to make judgments. That is, in the process of making a judgment or decision, people consult or refer to an “affect pool” containing all the positive and negative tags consciously or unconsciously associated with the mental representations of the task. Then, just as imaginability, memorability, and similarity serve as cues for probability judgments (e.g., the availability and representativeness heuristics), affect may serve as a cue for many important judgments. *Affective responses* occur rapidly and automatically – note how quickly you sense the positive and negative feelings associated with the stimulus words *treasure* or *hate*. As such, using an overall, readily available affective impression can be quicker and easier – and so sometimes more efficient – than weighing the pros and cons or retrieving from memory many relevant examples, especially when the required judgment or decision is complex or mental resources are limited.

Perhaps the most striking example of affect given by Slovic et al., is given by a study of individuals assessments of simple gambles. Here three samples are taken, each being asked to assess the attractiveness of playing a single gamble, either g^1 , g^2 or g^3 as detailed below:

Gamble g^1 (7/36, +\$9.00) (29/36, nothing)
Gamble g^2 (7/36, +\$9.00) (29/36, -\$0.05)
Gamble g^3 (7/36, +\$9.00) (29/36, -\$0.25)

From an expected utility perspective gamble g^1 offers the highest expected value and therefore should be the most attractive of the three gambles. However, from an affective perspective gamble g is difficult to evaluate as the \$9.00 gain in the P element of the bet does not have a clear comparator – i.e. the ‘nothing’ outcome in the P element has no clear affective message. Conversely gambles g^1 and g^2 both have well defined P element comparators. Furthermore, these are both small losses relative to the potential \$9.00 gain, i.e. while gamble g is difficult to assess, the affective message of gambles g^1 and g^2 is that for a modest potential loss (a ‘stake’) the subject can potentially obtain a relatively large gain.

In considering these gambles we define the concept of ‘comparability’ as arising where one given outcome can be readily compared to all other potential outcomes within the choice set of which subjects are aware¹. We can define two forms of comparability: (i) intra-gamble comparability, where all possible outcomes are clearly defined and (ii) inter-gamble comparability, which arises when subjects are permitted to compare across different gambles. This is not permitted within the Slovic et al design.

Slovic et al present the possible outcomes of each gamble to subjects via the modified roulette wheel illustrated in Figure 1 (which depicts gamble g^1). The figure also reproduces the response vehicle used by Slovic et al to elicit respondents assessments of these gambles. This was a 20 point rating scale in which a rating score of 0 was labelled as “Not at all an attractive bet”, a score of 10 was labelled as a “Moderately attractive bet” and a score of 20 was labelled an “Extremely attractive bet”.

¹ Bateman et al (2004) refer to this as the ‘visible choice set’.

Figure 1: Description and rating task for gamble g^2

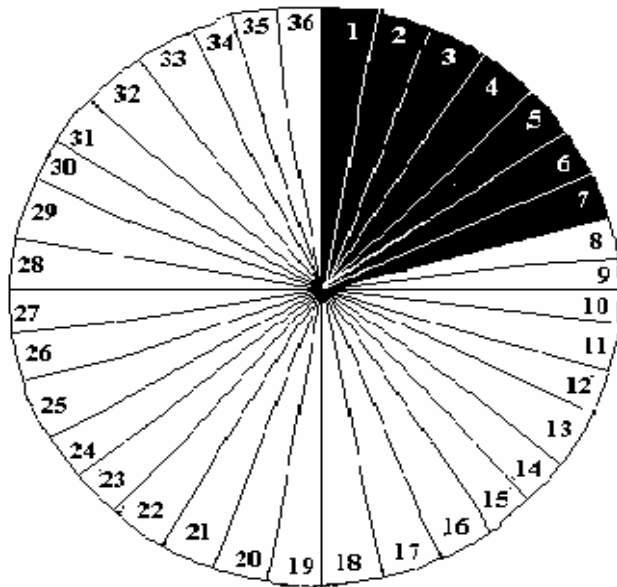
Evaluating the Attractiveness of a Bet

We would like you to rate how attractive the prospect of playing the following bet is to you

7/36 to win **\$9.00**
 29/36 to lose **\$0.05**

This means that there are 7 chances out of 36 that you will win the bet and receive **\$9.00** and 29 chances out of 36 that you will lose **\$0.05**

Visualise the roulette wheel on the left with 36 numbers along the circumference. If a ball lands on any of the 7 numbers between 1 and 7 inclusive, you win **\$9.00**. If it lands on numbers 8-36, you lose **\$0.05**



Indicate your opinion of this bet's attractiveness by **circling** one number on the rating scale below.

There is no right or wrong answer, we are interested only in your opinion about the **attractiveness** of playing this bet.

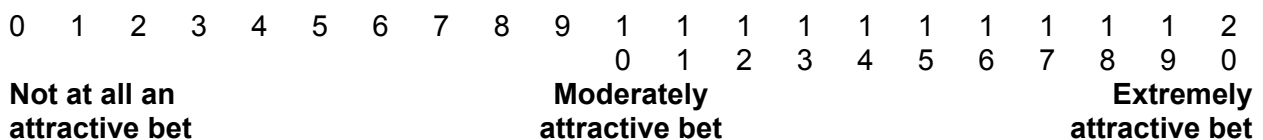


Table 1 presents results from this experiment showing that gamble g^1 is rated significantly worse than either g^2 or g^3 . While this result runs contrary to the predictions of expected value, it conforms to expectations based upon the affect heuristic.

Table 1: Split sample ratings of the attractiveness of three gambles

Gamble	Attractiveness Rating (0 to 20 scale)	Significance of difference from gamble g
g ¹ (7/36, +\$9.00), (29/36, 'nothing')	9.4	-
g ² : (7/36, +\$9.00), (29/36, -\$0.05)	14.9	p<0.001
g ³ : (7/36, +\$9.00), (29/36, -\$0.25)	11.7	p<0.050

Source: Adapted from Slovic et al. (2002)

Affective impressions vary not only in their valence, positive or negative, but in the precision with which they are held. There is growing evidence that the precision of an affective impression substantially impacts judgments. In particular, more precise affective impressions may carry more weight (Mellers et al., 1992). Developing this idea, Hsee (1996a, 1996b, 1998) has proposed the notion of evaluability to describe the interplay between the precision of an affective impression and its meaning or importance for judgment and decision making. Evaluability is illustrated by an experiment in which Hsee asked different groups of people to assume they were music majors looking for a used music dictionary. Hsee presented different subsamples with either one or other or both of the dictionaries described as follows:

- Dictionary A: Year of publication = 1993; Number of entries = 10,000; Defects = None, it's like new
- Dictionary B: Year of publication = 1993; Number of entries = 20,000; Defects = The cover is torn; otherwise it's like new

In a joint-evaluation condition, participants were shown both dictionaries, A and B, and asked how much they would be willing to pay for each. Willingness-to-pay was far higher for Dictionary B, presumably because of its greater number of entries. However, when one group of participants evaluated only A and another group evaluated only B, the mean willingness to pay was much higher for Dictionary A. Hsee explains this reversal by means of the *evaluability principle* (we will refer to this as EP for short). He argues that, without a direct comparison, the number of entries is hard to evaluate, because the evaluator does not have a precise notion of how good or how bad 10,000 (or 20,000) entries is. However, the defects attribute is evaluable in the sense that it translates easily into a precise good/bad response and thus it carries more weight in the independent evaluation. Most people find a defective dictionary unattractive and a like-new one attractive. Under joint evaluation, the buyer can see that B is far superior on the more important attribute, number of entries. Thus the number of entries becomes *evaluable* through the comparison process.

According to the EP, the weight of a stimulus attribute in an evaluative judgment or choice is proportional to the ease or precision with which the value of that attribute (or a comparison on the attribute across alternatives) can be mapped into an affective impression. In other words, affect bestows meaning on information (cf., Osgood, Suci, & Tannenbaum, 1957; Mowrer, 1960a, 1960b) and the precision of the affective meaning influences our ability to use information in judgment and decision making. EP can thus be seen as an extension of the general relationship between the variance of an impression and its weight in an impression-formation task (Mellers et al., 1992).

Hsee's work on evaluability is noteworthy because it shows that even very important attributes may not be used by a judge or decision maker unless they can be translated precisely into an affective frame of reference. Moreover, while Hsee (1998) finds evaluability effects with familiar attributes such as the amount of ice cream in a cup, Slovic et al. (2002) demonstrate similar effects relating to concepts such as numbers of human lives.

The present paper extends the Slovic et al., gamble rating work across a variety of further questions frames. The objective being to examine the robustness of tendency to rely upon the affect heuristic across different contexts. This is further extended by an analysis of the extent to which the personal characteristics of subjects may make them more or less likely to rely upon the affect heuristic.

2. FORMULATION OF HYPOTHESES AND EMPIRICAL METHODOLOGY

Overview

The present study sets out to address a number of interlinked questions concerning the circumstances under which reliance upon the affect heuristic may be observed and the responsiveness of such behaviour to changes in the payoff parameters of gambles². This research was initially prompted by the observation that the results reported by Slovic et al. directly conflict with expectations derived from standard economic theory. An economic view is that, provided these gamble assessment tasks are considered to constitute consequential questions³, then those individuals presented with gamble *g* would generally be expected to accord it at least an equal rating to that given by individuals facing gamble *g'*, with gamble *g* rated strictly higher if the potential loss in the *P* element of gamble *g'* is considered significant. Similarly, economic-theoretic expectations are that gamble *g'* would be rated as equal or better than gamble *g*". This expected preference for gamble *g* over these other gambles highlights the concept of 'opportunity cost', a fundamental principle of economics which is the cornerstone of any basic economics course. Here the opportunity cost is the loss associated with not obtaining the possible gain. In the case of gamble *g* there is a 7/36 chance of gaining \$9.00 and a 29/36 probability of a no-consequence or (as described) 'nothing' outcome. Here then there is a zero opportunity cost attached to gamble *g*. Therefore, while the affect account suggests that gamble *g* is difficult to assess, from an economic perspective its assessment is a trivial matter; the potential for a costless gain should be accorded the highest possible assessment, i.e. a score of 20 on the rating scale.

While the parameters of certain gambles are varied, the methodology used by Slovic et al is generally retained. Therefore, except where this conflicts with a subsequent test of inter-gamble comparability, each of the gambles used in the present experiments is presented to a separate sample. However, whereas the samples used by Slovic et al., were all drawn from a reasonably homogenous pool of students, the samples used in this research were deliberately chosen to contain students from differing backgrounds. This allows us to consider both whether certain individuals are relatively more or less likely to resort to the affect heuristic and see if this any such predisposition can be linked to certain personal characteristics. In particular, given the theoretical conflict between affective and economic expectations outlined above, we examine whether a prior training in economics is likely to make an individual less likely to exhibit affective responses. We also examine possible gender effects noting that a number of risk perception studies have observed relatively greater risk aversion amongst females than males. Furthermore, noting that the Slovic et al results are based solely upon responses from US subjects, by repeating certain examinations in both the US and UK we can test for a possible cultural dimension to affective behaviour⁴.

² Note that for all trials we hold the probability levels of the *S* and *P* elements of all bets constant at the levels used by Slovic et al. While an investigation of the impact of changes to the probability of a given outcome would be an interesting subject for future research, findings reported in Bateman et al., (2006) suggest that, while changes in payoff amounts have considerable impacts upon the attractiveness of bets, this is less true for changes in probability levels.

³ A strict economic interpretation of the gamble assessment tasks set by Slovic et al., is that they are inconsequential questions concerning hypothetical gambles which do not have binding real-world payoffs or losses. Such a view makes no predictions concerning the likely assessments observed and would not expect the significant differences and pattern observed by Slovic et al.

⁴ Obviously the cultural difference between US and UK populations may be less extreme than other potential pairings. However, for precisely that logic, if a difference in affective behaviour were observed between the US and UK we might suppose that such differences might be magnified across more diverse cultures.

Varying gamble parameters: Hypothesis formulation

Our initial investigations examine the replicability of findings reported by Slovic et al. Gambles g, g' and g'' were repeated in the UK using three separate samples of subjects. We denote these subject groups and the gambles to which they responded UK0, UK1 and UK2 respectively (Table 2, reported subsequently, collates all of the gambles assessed in this research). To permit assessment of any cultural influence upon affective behaviour, gamble g and g' were also re-run using separate samples in the US, these gambles and responses being denoted US0 and US1 respectively (it was not felt necessary to re-run gamble g'' in the US as the cross cultural comparison of gambles g and g' was felt sufficient to address this aspect of the research).

Ignoring for the moment the difference between US dollars and pounds sterling, we can then clarify these various gambles as follows:

gamble g = gamble UK0 = gamble US0 = (7/36, +9.00) (29/36, nothing)
gamble g' = gamble UK1 = gamble US1 = (7/36, +9.00) (29/36, -0.05)
gamble g'' = gamble UK2 = (7/36, +9.00) (29/36, -0.25)

then the Slovic et al affective behaviour result is that:

$$g' < g \quad \text{and} \quad g'' < g$$

were < indicates the direction of a statistically significant rating dominance. Note that this conflicts directly with economic theoretic-expectations which, assuming that potential losses are considered significant, are that:

$$g < g' < g''$$

Therefore, we have a readily testable hypothesis regarding conformity with either affective or economic expectations. Affective expectations would be satisfied if we obtained the following relationships:

$$UK1 < UK0 \quad \text{and} \quad UK2 < UK0 \quad [1]$$

and

$$US1 < US0 \quad [2]$$

Conversely, economic expectations are either that there will be no significant difference between any of these gamble ratings (suggesting either that gamble assessment tasks are considered inconsequential and/or that differences are considered trivial) or that ratings will follow expected values, i.e.:

$$UK0 < UK1 < UK2 \quad [1a]$$

and

$$US0 < US1 \quad [2a]$$

The clear difference between affective and economic expectations allows us to formulate testable hypotheses: affective expectations being reflected in Equations [1] and [2] while economic expectations are given in Equations [1a] and [2a].

Note that our principle test of cultural difference/similarity is to examine whether the patterns of gamble rating observed within country and replicated across countries. Because we are looking at similarities of difference pattern the fact that gamble payoffs are expressed using different currency units (\$ vs. £) is irrelevant. However, there is reason to suppose that direct comparison between same format gambles assessed in different countries may still be meaningful despite the difference in absolute values (i.e. \$9 is not the same value as £9). Reasons for supposing that such comparisons may have merit can be derived from the findings of (Birnbaum, 1999). This work suggests that, where there are either no clear comparators, or those comparators are of an entirely different magnitude, the subjects may, to some extent (and clearly within some limits) react affectively to the number 9 rather than to the value of goods which it represents. We can conduct a necessary, if not be itself sufficient, test of such an assertion by comparing ratings of common gambles across US and UK subjects, i.e. testing whether ratings of gambles suggest the following:

- $g = US0 = UK0$ [3]
- and
- $g' = US1 = UK1$ [4]
- and
- $g'' = UK2$ [5]

Moving to consider respecifications of gamble parameters not considered in the Slovic et al. work, we note that gambles g to g'' only vary in terms of the scale of financial losses within the P element. Therefore one obvious test is to examine the responsiveness of subject ratings of gambles to increases in the payoff specified in the \$ element of a bet. Here any lack of responsiveness to substantial changes in the size of a payoff would be reminiscent of the scope insensitivity phenomena reported by Kahneman and Knetsch (1992). Note again that the affective interpretation of ratings of gamble g is that, in the absence of any comparator, subjects find the \$ element payoff difficult to assess. Therefore, an interesting set of comparisons is to increase the payoff in the \$ element of gamble g (our gamble US0) and examine to what extent this improves the rating of the new gamble. In particular to examine whether substantial increases in this payoff have anything like the positive affective impact of adding a small comparator as per gamble g' (our gamble US1). To examine this we define a new gamble as follows:

$$\text{gamble US3} = (7/36, +\$90.00) \quad (29/36, \text{nothing})$$

Note that this has ten times the possible payoff of gamble US0 but, as per that gamble, it has no P element comparator. The affective expectation is therefore that this will fail to have the attractiveness of gamble US1 [(7/36, +\$9.00) (29/36, -\$0.05)], i.e.:

$$US3 < US1 \quad [6]$$

whereas economic expectations would be that (i) this inequality would be reversed, i.e.

$$US1 < US3 \quad [6a]$$

and (ii) that, assuming an increase in payoff from \$9.00 to \$90.00 is considered significant, then

$$US0 < US3 \quad [6b]$$

Returning to consider changes in the P element of gambles, the Slovic et al redefinitions of this element across gambles g to g'' move from a no-consequence 'nothing' outcome to a negative payoff (i.e. a loss to the subject). While economic expectations were that this would result in a diminution of the attractiveness of those bets, rating scale responses revealed the

opposite trend with both gamble g' and g'' being rated higher⁵ than gamble g. The affective explanation of this result centres upon the difficulty subjects have in assessing gamble g which has no clear P element comparator and hence no definite affective message. The affective impact of adding the comparator within gambles g' and g'' is highly positive since, although these comparators entail losses, these appear very small relative to the \$ element payoff thus making the overall gamble appear more attractive than gamble g.

The affective line of reasoning set out above, turns on the argument that it is the presence of the P element comparator, rather than (within limits no doubt) its size⁶, which induces the higher rating of gamble g' over gamble g (and hence, from an economic perspective, an anomaly). This leads us to speculate upon the likely impact of adding a positive payoff comparator as the P element of a gamble. In particular we consider a relatively small positive comparator as per gamble US2 as follows:

$$\text{gamble US2} = (7/36, +\$9.00) (29/36, +\$0.05)$$

The economic expectations of gamble US2 are straightforward, that it should be rated above either gamble g (US0) or g' (US1). However, affective expectations are more complex. While we would expect that the introduction in element P of a comparator against which the payoff in element \$ can be assessed would make this new gamble more attractive than gamble g (US0), it is unclear whether it will be as attractive as gamble g' (US2). The reasoning behind this assertion turns on a variant of regret theory (Loomes and Sugden, 1982) combined with conversational norms. Gamble g' is an archetypal bet, wherein a 'stake' is wagered for some return. Here the potential loss (the -\$0.05 comparator) is very small relative to the potential gain (+\$9.00). Subjects see the potential loss as being an excellent 'stake' to risk relative to the potential gain; i.e. this can be described as a potential 'good loss' given the possible positive payoff. Consequently the gamble is both readily evaluated and highly rated. Changing the P element payoff from a small loss to a small gain (+\$0.05) makes this an unusual gamble (i.e. it contravenes conversational norms). While the gamble is still readily evaluable (the \$ element potential payoff is again large compared to the P element outcome; this is clearly a affectively 'good' bet), the gain versus gain makeup of this gamble seems unusual. We hypothesise that subjects may see the potential +\$0.05 gain in the P element as somewhat undesirable as it would be tinged with regret at not winning the \$ element payoff of +\$9.00; i.e. it may be seen as a 'bad gain'. Therefore, while we expect the enhanced evaluability of this gamble to make it more attractive than gamble g, we may not expect it to be as attractive as gamble g'. Such a result would run counter to economic expectations as gamble g' involves a potential loss. Converting these gambles into those for which we gather data we formulate this into the following affective expectation:

$$\text{US0} < \text{US2} < \text{US1} \quad [7]$$

whereas the economic expectation is that:

$$\text{US1} < \text{US0} < \text{US2} \quad [7a]$$

The penultimate aspect of this research attempted to examine the constructed interpretation of the implicit no-consequence or 'nothing' P outcome in gamble g. Here we assess the impact of 'redefining nothing' in ways which, from an economic perspective, still imply a zero opportunity cost, but which may generate different affective responses. Specifically we

⁵ However, note that gamble g'', which had a P element payoff of -\$0.25 was rated below gamble g' as per economic expectations.

⁶ Note that we can provide a further test of this hypothesis through comparison of gamble UK1 and UK2; the expectation being that ratings of these gambles may not differ significantly (although note that such a result would be in accordance with both affective and economic expectations).

examine the impact of describing such as no-consequence outcome as 'winning nothing'. This is then contrasted in a separate trial by defining the no-consequence outcome as 'losing nothing'. These gambles are denoted as follows:

gamble US4 = (7/36, +\$9.00) (29/36, you win nothing)

gamble US5 = (7/36, +\$9.00) (29/36, you lose nothing)

An affective interpretation of these descriptions might note that while 'winning nothing' reminds subjects that they will not gain from the P element of the bet, 'losing nothing' reminds them that they will not suffer a loss. Loss aversion is a commonly observed phenomena in psychological and economic experiments (Kahneman, Knetsch and Thaler, 1990; Tversky and Kahneman, 1991; Bateman et al., 1997) wherein a unit loss is typically judged to be of greater weight than a unit gain. Given this, an affective expectation might be that positive reactions to the no-consequence definition of 'losing nothing' may be stronger than the negative reactions to the re-definition of no-consequence as 'winning nothing'. Therefore, while affective relationships to gamble g (US0) are uncertain we expect that:

$$US4 < US5 \quad [8]$$

while economic expectations are that:

$$US0 = US4 = US5 \quad [8a]$$

Finally, we conduct one further analysis of the concept of comparability. All of the gambles referred to above were presented to different samples of respondents. This means that none of these gambles were subject to any degree of inter-gamble comparability. Therefore, as a final experiment we examine the impact of presenting two of the Slovic et al., gambles (g and g') to the same sample of respondents, ensuring that they study both of these before rating each. These gambles were assessed by a UK sample yielding the response sets denoted as follows:

UK(01)0 = (7/36, +£9.00) (29/36, nothing)

UK(01)1 = (7/36, +£9.00) (29/36, -£0.05)

where the notation (01) indicates that this sample viewed both gamble g and g' before rating each. Here expectations are that the affective presence of the intra-gamble comparator in gamble g' (the -\$0.05 loss) will be offset by the inter-gamble comparability highlighting the zero opportunity cost of gamble g. In short we expect that the ordering of gambles g and g' observed by Slovic et al., will be reversed within this inter-gamble comparability frame. Indeed affective expectations would be that the difference between these gambles would be magnified such that we would observe.

$$UK(01)1 < UK(01)0 \quad [9]$$

The difference between these and economic expectations is less marked than in previous tests being merely a matter of degree rather than direction. Given that the expected value difference between gambles g and g' is roughly just 2%, economic expectations are for the same directional effect as above although the magnitude is anticipated to be small and possibly insignificant as follows:

$$UK(01)1 \leq UK(01)0 \quad [9a]$$

Table 2 summarises the various gambles assessed in our research and, as appropriate, shows where these are equivalent to those tested by Slovic et al. Sample sizes are also reported.

Table 2: Gambles used in affective rating experiments

Slovic et al gambles ¹	Present gamble code	Gamble	Assessed via single or joint sample	Sample size
g	UK0	(7/36, +£9.00)	single	35
g'	UK1	(7/36, +£9.00) (29/36, -£0.05)	single	37
g''	UK2	(7/36, +£9.00) (29/36, -£0.25)	single	33
g	UK(01)0	(7/36, +£9.00)	joint	44
g'	UK(01)1	(7/36, +£9.00) (29/36, -£0.05)		
g	US0	(7/36, +\$9.00)	single	50
g'	US1	(7/36, +\$9.00) (29/36, -\$0.05)	single	50
	US2	(7/36, +\$9.00) (29/36, +\$0.05)	single	50
	US3	(7/36, +\$90.00)	single	50
	US4	(7/36, +\$9.00) (29/36, you win nothing)	single	50
	US5	(7/36, +\$9.00) (29/36, you lose nothing)	single	50

Note: 1. Ignores currency differences between US and UK samples

Empirical methods

Given that we are interested in part in the link between subjects characteristics and their reliance upon the affect heuristic, we require a more varied pool of respondents than that used in the original Slovic et al experiments. Specifically we require varied exposure to training in economics to examine whether those with such training are less susceptible to affective responses to the stimuli embodied in the gambles. Given this requirement, students were recruited⁷ from courses with mixed intakes, such as certain environmental science options, where it was known that both non-economists and those with economics training were registered. Ratings of the gambles prefixed 'UK' were collected from students subjects interviewed at the University of East Anglia between October 1998 and November 2001⁸. Ratings of gambles prefixed 'US' were collected at the State University of New York, Plattsburgh between February and May, 2001.

While subjects rating gambles UK(01)0 and UK(01)1 saw *both* gambles prior to assessing them, all other subjects were presented with just one gamble to rate⁹. Gambles were presented using the same rubric, roulette wheel figure and 0 to 20 rating scale as employed by Slovic et al. Subjects were isolated from each other and prevented from communicating with each other during experimental sessions.

⁷ Subjects were not paid to undertake these experiments and were not compelled to do so.

⁸ Tests confirmed that responses to an identical gamble did not differ significantly across years ($p > 0.80$).

⁹ Procedures were put in place to ensure that no subject could be involved in more than one experimental session.

3. RESULTS

Sample characteristics

A total sample of some 449 subjects was collected, of which roughly two thirds were sampled in the US and the remainder in the UK. Distribution of subjects across treatments is as reported in Table 2. Subjects varied in age from 18 to 52 years old although the majority were young students, the overall sample having a mean and median of 21.4 and 20.0 years respectively. Table 3 details two other characteristics which, for reasons outlined previously, were of particular interest as potential determinants of subjects rating scores. The gender profile of the sample was, roughly speaking, equally divided between the sexes, with 54.0% of subjects being female. Similarly both economists and non-economists were well represented with 56.2% of respondents having at least some training in economics. However, given that it is plausible that a training induced resistance to the affect heuristic might be proportional to the level of that training we disaggregate these statistics noting that only 12.5% have the highest level of training in our sample (those who hold who are obtaining a first degree or above in economics; a characteristic which we encode into the dummy variable 'ECODEGRE' for which the value 1 denotes the presence of such training and the value 0 denotes its absence).

Table 3: Sample characteristics

Variable	Category	Count	Percent of total sample ¹
Country in which sampled	US	300	66.8
	UK	149	33.2
Gender ²	Female	242	54.0
	Male	206	46.0
Level of training in economics	No training in economics	201	44.8
	Some pre-University level training only	158	35.2
	Some University level training in economics but not taken as a major subject	34	7.8
	Economics taken as major subject for first or higher University degree	56	12.5

Note: 1. Rounding may make column subtotals appear marginally different from 100.
2. One subject failed to record their gender.

Rating results are presented as follows. We initially consider the impact of moving between gambles and test the various hypotheses formally defined. We then consider the impact of respondents characteristics upon ratings, in particular considering the effects which culture, gender and a training in economics may have upon reliance upon the affect heuristic.

3.1 Testing Hypotheses: Rating Differences Across Gambles

Table 4 details summary statistics for the various gambles assessed in this research. cursory inspection of these findings show that all relationships are in the direction of the affective expectations set out previously and, where these conflict with economic expectations, accordingly diverge from the latter. We now consider the significance of these differences for each of the hypotheses embodied in Equations [1] to [9]. In so doing we adopt the usual $\alpha = 5\%$ threshold for significance but also note instances where an $\alpha = 10\%$ rule would indicate significance. Appendix 1 to this paper reports both parametric and non-parametric tests for all possible comparisons of gambles. These two types of test imply similar conclusions in all cases. However, in the discussions set out below we report the more conservative non-parametric tests throughout.

Table 4: Descriptive statistics for gamble ratings by group (0-20 response scale)

Slovic et al gambles ¹	Present gamble code	Gamble	Assessed via single or joint sample	Sample size	Mean	Median	StDev
g	UK0	(7/36, +£9.00)	single	35	9.28	8.0	6.03
g'	UK1	(7/36, +£9.00) (29/36, -£0.05)	single	37	13.24	16.0	6.15
g''	UK2	(7/36, +£9.00) (29/36, -£0.25)	single	33	12.61	14.0	5.64
g	UK(01)0	(7/36, +£9.00)	joint ²	44	13.09	14.0	5.57
g'	UK(01)1	(7/36, +£9.00) (29/36, -£0.05)			9.82	10.0	6.64
g	US0	(7/36, +\$9.00)	single	50	8.66	7.5	4.86
g'	US1	(7/36, +\$9.00) (29/36, -\$0.05)	single	50	12.20	13.5	5.69
	US2	(7/36, +\$9.00) (29/36, +\$0.05)	single	50	11.40	11.0	4.79
	US3	(7/36, +\$90.00)	single	50	10.48	10.0	4.93
	US4	(7/36, +\$9.00) (29/36, you win nothing)	single	50	9.02	9.5	5.08
	US5	(7/36, +\$9.00) (29/36, you lose nothing)	single	50	13.16	14.5	5.63

Note: 1. Ignores currency differences between US and UK samples

2. Subjects in sample group UK(01) saw both of their gambles prior to responding to both; their responses being labelled UK(01)0 and UK(01)1 respectively.

Re-testing the Slovic et al Gambles

The hypotheses embodied in Equations [1] to [5] all refer to re-examinations of the original Slovic et al. gambles. The central result from the Slovic et al. tests is the affective finding that $g' < g$ and $g'' < g$, differences which are significant at $p < 0.001$ and $p < 0.050$ respectively. This is re-tested in full within the UK samples as per Equation [1]. This confirms affective expectations, showing that $UK1 < UK0$ and $UK2 < UK0$, with differences significant at $p = 0.008$ and $p = 0.038$ respectively. Within our US samples the first Slovic et al inequality was re-tested and again found to hold, i.e. Equation [2] is satisfied with $US1 < US0$ at $p = 0.001$.

The central Slovic et al. result is therefore replicated in our re-testing and its message reconfirmed; the addition of a small comparator, even though it is a negative amount makes the affective message of gambles g' and g'' clear and positive relative to the difficult to evaluate gamble g . Economic expectations, as embodied in Equations [1a] and [2a] are rejected, indeed these fail to correctly predict the ordering of these gambles.

Closer inspection of the absolute magnitude of the absolute rating scores of these gambles shows remarkable consistency both across US and UK samples and in comparison with the scores reported by Slovic et al. This suggests a cultural similarity of the magnitude of affective impacts as hypothesised in Equations [3] to [5]. Formal testing of these shows that ratings of gamble US0 are insignificantly different from those of gamble UK0 ($p = 0.734$) both of which are very similar to ratings reported by Slovic et al for gamble g . Similarly ratings of gamble US1 are insignificantly different from those of gamble UK1 ($p = 0.256$), both of which have similar ratings to Slovic et al's gamble g' . While we do not formally test the similarity between gamble g'' and gamble UK2, summary statistic suggest that these are also insignificantly different.

Testing Scope Sensitivity

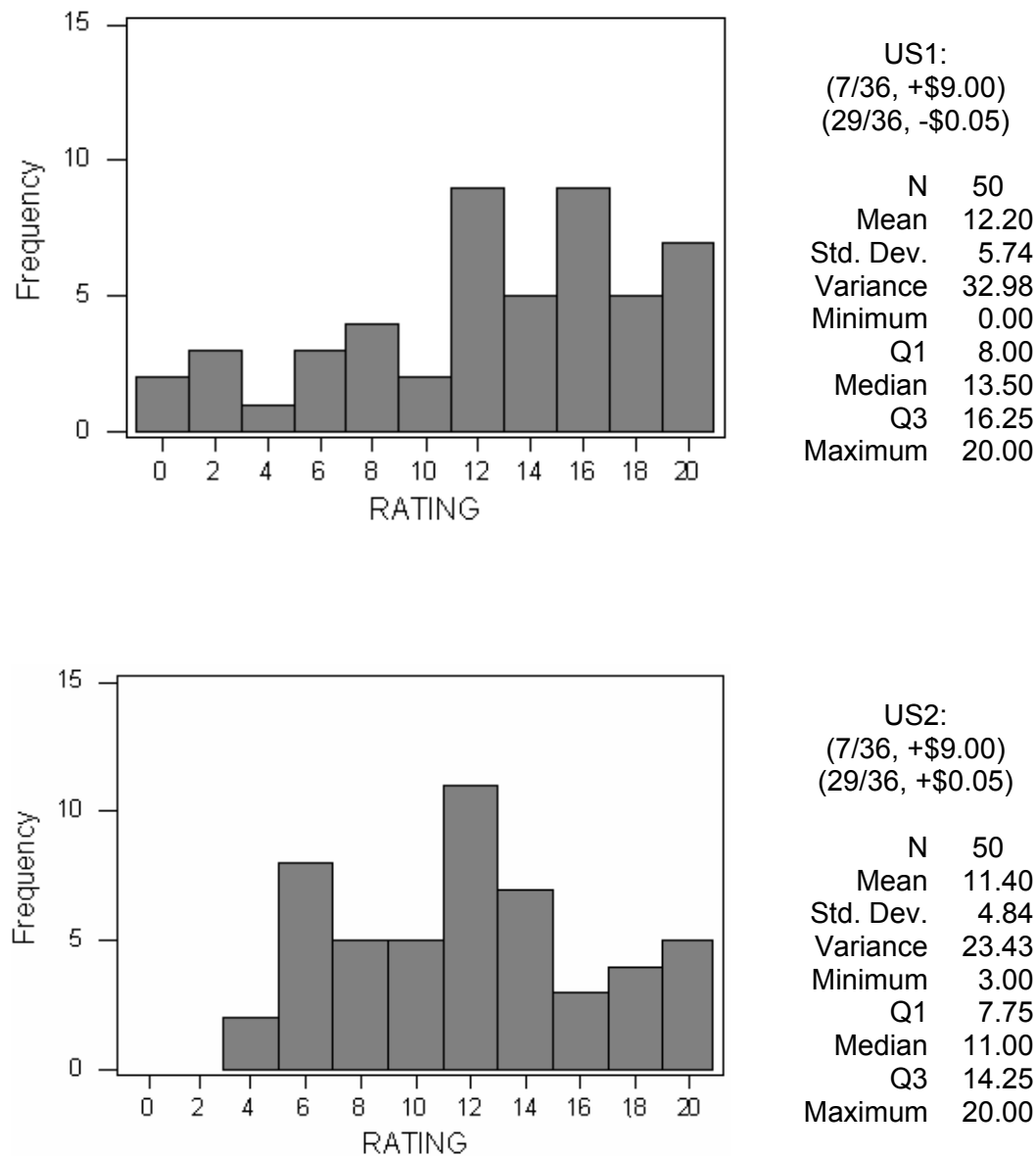
Economic expectations regarding the test of scope sensitivity afforded by gamble US3 are set out in Equations [6a] and [6b]. These are that, provided that the expected value difference of this gamble is considered significantly greater than that of gambles US0 and US1 (an assertion that seems reasonable given that US3 offers roughly ten times the expected payoff of both these other gambles) then we should observe $US0 < US3$ and $US1 < US3$. In the event we find that this latter directional expectation is reversed, with gamble US3 being rated as less attractive than gamble US1. While this conflicts with economic expectations based upon the much higher expected payoff of gamble US3, it conforms to affective expectations (however, the difference between these gambles is not significant at our accepted significance level of $\alpha = 5\%$, although as $p = 0.064$ it is significant at $\alpha = 10\%$). This underlines the affective power of adding the comparator to a gamble which appears stronger than even a ten times increase in the \$ element of a gamble without such a comparator. Turning to the comparison of gambles US3 and US0, while there is no clear affective expectation here the economic expectation that $US0 < US3$ is fulfilled although this difference is only significant at $\alpha = 10\%$ ($p = 0.053$). This suggests that economic considerations do weight upon subjects rating responses in the absence of clear affective messages.

Good Losses and Bad Gains

Gamble US2 permits us to test the 'good losses versus bad gains' hypotheses formalised in the affective expectations of Equation [7] (that $US0 < US2 < US1$) and economic expectations of Equation [7a] (that $US1 < US0 < US2$). As Table 4 shows, it is the affective ordering which is observed. Both economic and affective expectations coincide and are fulfilled regarding the relation of gambles US1 and US2 with the latter significantly higher than those of the former ($p = 0.010$). However, contrary to economic expectations, but in

accord with affective predictions, ratings of gamble US1 are higher than those of gamble US2. Although this difference is not statistically significant ($p=0.301$), Figure 2 shows that the distribution of ratings across these two gambles are somewhat different. Measures of the rating mean, median and both upper and lower quartiles are all lower for US2 than US1. Indeed it seems that the only reason why rating differences are not statistically significant is that while ratings for gamble US1 fall as low as zero, those for US2 do not fall below a value of 3, a result which is unsurprising given that this is a no-loss gamble¹⁰.

Figure 2: Rating response frequency distributions for gambles US1 and US2



¹⁰Further evidence of a difference between gambles US1 and US2 can be found by comparing ratings from these gambles with those for others. For example, comparing both of these with the highest rated gamble (US5) we find that while the difference between US1 and US5 is not significant ($p = 0.432$), that between US2 and US5 does prove significant ($p = 0.051$).

Redefining Nothing

Gambles US4 and US5 provide alternative definitions of the P element 'nothing' payoff of gamble US0. From an economic perspective all are identical and should receive equal rating as per Equation [8a]. However, an affective interpretation, while having no clear expectation about relationships with gamble US0, would note the negation of loss implicit in gamble US5 and therefore (in the light of prevalent loss aversion) expect this gamble to be accorded a higher ratings than either gamble US0 or US4. Results detailed in Table 4 conform directly to these affective predictions and reject economic expectations. Gamble US5 is rated significantly higher than either gamble US0 ($p < 0.001$) or US4 ($p < 0.001$). Indeed gamble US5 is the highest rated of all gambles, with a rating which is significantly higher than that of Gamble US3 ($p = 0.008$) despite the latter having an expected value which is ten times higher.

Investigating Inter-Gamble Comparability

Split sample comparison of gambles UK0 and UK1 found that the addition of the P element comparator in UK1 resulted in a significantly higher rating than that for UK0 despite the former entailing a potential loss and therefore lower expected value ($p = 0.008$). Sample UK(01) viewed both of these gambles prior to rating them, their responses being denoted UK(01)1 and UK(01)0 respectively. The impact of this inter-gamble comparability is dramatic, reversing the line of dominance mentioned above such that we now see gamble UK(01)1 rated as being significantly better than gamble UK(01)0 ($p = 0.015$). This direction of dominance is predicted by both economic and affective expectations although only the affective view predicts this switch of direction when moving from a split sample to joint comparison frame. Furthermore, the degree of difference between ratings of the joint comparison gambles cannot be reconciled with economic expectations. Mean rating for the no-loss gamble UK(01)0 is 13.09 compared to 9.82 for the with-loss gamble UK(01)1, an increase in the former over the latter of one third, for an increase in expected value of roughly 2%. In effect it seems that the joint comparison frame itself generates an affective response, which, although this operates now in the same direction as expectations based upon expected values, results in an over-reaction to the trivial value differences concerned.

Testing hypotheses: Rating differences across gambles - Summary

The above tests yield a persuasive and consistent message that affective expectations are fulfilled in all cases including those where such expectations are at odds with those derived from economic expectations. We now turn to consider the role which the characteristics of individuals may play in determining responses to rating scale evaluations of gambles.

3.2 Individual Determinants of Rating Responses

Our tests of the hypotheses underpinning Equations [3] to [5] show that there is no significant cultural difference between ratings of gambles undertaken in the US and UK. Therefore, in the following set of analyses we apply regression techniques to jointly investigate the influence of both a training in economics and gender upon gamble rating responses.

Initial investigations supported the notion that those with training in economics did indeed rate a subset of gambles significantly differently from how non-economists assessed these gambles. This subset was, as previously hypothesised, that set of gambles for which there was no P element loss, i.e. where there was a zero opportunity cost associated with the gamble. Additional analysis indicated that this difference in rating was positively correlated to the extent of economics training which a respondent had experienced such that those who majored in economics at University level formed the most distinctly different group from those with no prior economics training. Accordingly the dummy variable *ECODEGRE*, previously defined to identify the former group, was used in our regression analyses of the influence of individual characteristics upon gamble ratings. A second dummy variable, *FEMALE*, taking a value of 1 for all female subjects and 0 for males was used to investigate the influence of gender within such regressions models. As outlined previously we expect females to be more risk averse leading us to expect a negative estimated coefficient upon this variable.

We open our analysis by focussing upon the results from our eight single gamble rating tasks (i.e. all responses except for those from our joint assessment of gambles the UK(01)0 and UK(01)1 to which we return subsequently). Considering first the hypothesised link between a training in economics and gamble ratings, our conjecture is that subjects who have experienced the level of training captured in the *ECODEGRE* variable will view those gambles with lack non-zero P element payoffs (gambles UK0, US0, US3, US4 and US5) will see such gambles as costless gains. This perception will derive from familiarity with the concept of opportunity cost which is a fundamental principle of economic analysis and yet one which may not come naturally to the untrained. Therefore, while other find such gambles as lacking a clear affective message and therefore difficult to evaluate, trained economists will be immune to such uncertainty. In effect, while non-economists see these gambles as lacking a comparator, for economists the concept of opportunity cost provides such a comparator and, this being zero, such respondents see these gambles as 'good bets'. This expectation would be reflected in a positive estimated coefficient upon the *ECODEGRE* variable.

Table 5 reports a simple cross classification of ratings of these this set of zero opportunity cost gambles according to the two levels of the *ECODEGRE* variable. This shows as expected that subjects with high level training in economics rate these gambles higher than do other subjects, a difference which proves highly significant ($p < 0.001$).

Table 5: Rating scores for zero opportunity cost gambles (gambles UK0, US0, US3, US4 and US5)

Level of economics training	N	Mean	Median	StDev	SE Mean	Min	Max	Q1	Q3
ECODEGRE = 1	35	13.91	14	4.41	0.75	3	20	12	17
ECODEGRE = 0	200	9.52	9	5.46	0.39	0	20	5	14

The simple cross classification test of Table 5 fails to distinguish between gambles within this set and does not control for the influence of gender. To allow for these effects Table 6 reports a regression analysis for ratings of this set of zero opportunity cost gambles (holding gamble UK0 as a base case).

Table 6: Regression model of rating scores for zero opportunity cost gambles (gambles UK0, US0, US3, US4 and US5)

Variable	Unstandardized Coefficients	Std. Error	Standardized Coefficients	t	Sig.
	B		Beta		
(Constant)	9.463	0.906		10.447	.000
<i>ECODEGRE</i>	3.669	0.930	.237	3.944	.000
<i>FEMALE</i>	-2.453	0.673	-.222	-3.643	.000
<i>US0</i>	0.130	1.105	.010	0.118	.906
<i>US3</i>	1.534	1.099	.114	1.396	.164
<i>US4</i>	0.858	1.117	.064	0.768	.443
<i>US5</i>	4.435	1.102	.329	4.022	.000

Dependent variable: gamble rating
Base case = *UK0*

n = 235 $R^2 = 0.212$ $R^2(\text{adj.}) = 0.191$ $F=10.237$ ($p<0.001$)

where:

ECODEGRE = 1 for subjects who have or are engaged in obtaining a first degree or above in economics; = 0 otherwise
FEMALE = 1 if subject is female; = 0 if male.
UK0, US0, US3, US4 and *US5* = dummy (1,0) variables identifying gambles

The findings reported in Table 6 conform directly with our expectations regarding the influence of a training in economics and gender. The economics effect is, as expected positive and proves highly significant ($p < 0.001$). Trained economists rate these zero-loss gambles significantly more highly than do non-economists. They seem able to bring their knowledge of opportunity cost to bear upon these tasks in a way which makes them readily identifiable as good bets. Conversely, non-economists fail to read clear affective messages from such gambles. The lower ratings awarded by non-economists to such gambles suggests that this lack of ready comprehension results in a risk-averse attitude to their rating, reflected in lower scores. Also in line with our expectations is the effect of gender with female subjects rating gambles significantly lower than do males supporting the heightened risk aversion amongst women observed elsewhere. Finally, once we control for these other predictors none of the individual gambles within this set prove significantly different from each other except for our 'negated loss-aversion' gamble *US5* which achieves a significantly higher rating than other gambles, reconfirming the attractive affective message of this gamble.

The obvious extension to the above analysis is to consider economist and gender affects amongst remaining gambles, again reserving the joint evaluation set for subsequent analysis. Table 7 reports ratings for gambles *UK1, UK2, US1* and *US2* (i.e. all those which do have a clear non-zero comparator within the P element) for subjects with and without prior economics training as defined by the *ECODEGRE* variable.

Table 7: Rating scores for the ‘with-comparator’ gambles (gambles UK1, UK2, US1 and US2)

Level of economics training	N	Mean	Median	StDev	SE Mean	Min	Max	Q1	Q3
ECODEGRE = 1	17	10.47	11.00	5.66	1.37	2	20	5	15.5
ECODEGRE = 0	153	12.47	13.00	5.53	0.447	0	20	8	17.0

Studying Table 7 we can see that the direction of difference is now completely reversed from that shown in Table 5. Now economists rate these ‘with-comparator’ gambles significantly below the level accorded by non-economists ($p < 0.001$). Remembering that the higher values accorded to these gambles were attributed to the positive affect induced by the presence of a P element comparator (irrespective of whether that comparator was positive or negative) we are tempted to conclude that economists are significantly resistant to such affective messages. However, the simple analysis of Table 7 ignores other potential predictors such as gender effects. This is rectified in Table 8 which presents a regression analysis of these ‘with-comparator’ gambles.

Table 8: Regression model of rating scores for the ‘with-comparator’ gambles (gambles UK1, UK2, US1 and US2)

Variable	Unstandardized Coefficients	Std. Error	Standardized (Beta) Coefficients	t	Sig.
(Constant)	14.173	1.036		13.684	.000
<i>ECODEGRE</i>	-1.610	1.454	-.087	-1.107	.270
<i>FEMALE</i>	-1.750	.869	-.157	-2.015	.046
<i>UK2</i>	-.776	1.332	-.055	-.582	.561
<i>US1</i>	-.566	1.222	-.047	-.463	.644
<i>US2</i>	-1.594	1.205	-.131	-1.323	.188

Dependent variable: gamble rating
Base case = *UK1*

n = 170 $R^2 = 0.218$ $R^2(\text{adj.}) = 0.018$ $F=1.621$ ($p=0.157$)

where:

UK1, *UK2*, *US1* and *US2* = dummy (1,0) variables identifying gambles

other variables as previously defined

Table 8 suggests that the effect of an economics training upon ratings of the ‘with-comparator’ gambles, while significant in the simple cross classification of Table 7, proves insignificant once we control for other predictors, specifically the gender effect which is the only significant variable in this model and reconfirms the lower rating of gambles accorded by females.

The change in the significance of the economics effect between Tables 7 and 8 suggests some possible correlation between the *ECODEGRE* and *FEMALE* variables and we return to this issue shortly. However, before that report the impact of these variables upon ratings of the jointly assessed gambles *UK(01)0* and *UK(01)1*. Table 9 presents ratings of these gambles disaggregated by the *ECODEGRE* variable.

Table 9: Rating scores for the jointly assessed gambles (UK(01)0 and UK(01)1)

Gamble	Level of economics training	N	Me an	Me d.	s.d.	Min	Ma x	Q1	Q3	Sig. (p)
UK(01)0	ECODEGRE = 0	40	13.18	14.5	5.57	0	20	8.3	18.5	0.756
	ECODEGRE = 1	4	12.25	12.0	6.34	5	20	6.3	18.5	
UK(01)1	ECODEGRE = 0	40	10.00	10.0	6.54	0	20	5.0	15.0	0.572
	ECODEGRE = 1	4	8.00	6.0	8.37	1	19	1.3	16.8	

As can be seen, the low number of trained economists in the joint assessment sample means that differences in rating are clearly insignificant although it is interesting to note that economists are if anything somewhat more susceptible to over-sensitivity to the difference between these gambles with median values halving in response to what is just a 2% change in expected values. Regression analysis of this data, presented in Table 10 reconfirms that this effect is insignificant given the small sub-sample size concerned, although the gender effect is once again significant.

Table 10: Regression model of rating scores for the jointly assessed gambles (UK(01)0 and UK(01)1)

Variable	Unstandardized Coefficients	s.e.	Standardized (Beta) Coefficients	t	Sig.
(Constant)	11.271	1.138		9.908	.000
ECODEGRE	-2.122	2.263	-.097	-.938	.351
FEMALE	-2.640	1.302	-.210	-2.027	.046
UK(01)0	3.273	1.287	.261	2.542	.013

Dependent variable: gamble rating
Base case = UK(01)1

n = 44 $R^2 = 0.340$ $R^2(\text{adj.}) = 0.084$ $F=3.666$ ($p=0.015$)

where:

UK(01)0 and *UK(01)1* = dummy (1,0) variables identifying gambles

other variables as previously defined

Returning to the difference in significance in the economist effect between the simple cross classification of Table 7 (where it was significant) and the regression model of Table 8 (where the addition of the gender effect made the impact of economic training non-significant), this suggests that there may be collinearity between the *ECODEGRE* and *FEMALE* variables. Table 11 examines this issue showing that, of the just over 12% of the overall sample who were trained economists some 57% were male. As Table 11 indicates, while the association between the *ECODEGRE* and *FEMALE* variables is not significant within conventional $\alpha = 5\%$ limits, it is just significant at the $\alpha = 10\%$ level ($p = 0.093$).

Table 11: Cross tabulation of *ECODEGRE* and *FEMALE* variables

	Female (<i>FEMALE</i> = 1)	Male (<i>FEMALE</i> = 0)	Total
Trained economist (<i>ECODEGRE</i> = 1)	26 (43%) [10%]	34 (57%) [15%]	60 (100%) [12%]
Other (<i>ECODEGRE</i> = 0)	237 (55%) [90%]	195 (45%) [85%]	432 (100%) [88%]
All	263 (53%) [100%]	229 (47%) [100%]	492 (100%) (100%)

Chi-Square = 2.814, d.f. = 1, p-value = 0.093

() = row percentages

[] = column percentages

Although the gender difference within the trained economist sub-set is not particularly substantial, the existence of any even marginally significant difference may mean that the effects attributed to the *ECODEGRE* variable are actually reflections in part of the gender effect. To control for this possibility the analyses of the effects of the *ECODEGRE* variable upon the 'zero-opportunity cost' and 'with-comparator' gambles was repeated holding gender constant by omitting all females from our sample. Table 12 reports findings from this analysis.

The results reported in Table 12 show that, even when we remove any possible gender effects, trained economists still rate the 'zero-opportunity cost' gambles significantly higher than do non-economists ($p < 0.001$). Indeed no male economist rated any of these gambles lower than a score of 11, a value which lies above the median score of non-economists. Conversely economists rated the 'with-comparator' gambles lower than did non-economists, although small sub-samples sizes mean that this result remains statistically insignificant.

Table 12: Economist effect when all females are excluded from the sample

Gamble sets	Subject type	Count	Mean	s.e. mean	s.d.	Min.	Q1	Median	Q3	Max.	sig. of rating difference: non- economists vs. economists
'zero-opportunity cost' gambles (UK0, US0, US3, US4, US5)	non-economists	85	10.65	.65	6.00	0	6.00	10.00	15.00	20	p < 0.001
	economists	24	15.54	.59	2.87	11	13.00	15.00	18.00	20	
	Group Total	109	11.72	.56	5.82	0	6.50	13.00	16.00	20	
'with-comparator' gambles (UK1, UK2, US1, US2)	non-economists	70	13.49	.70	5.88	0	9.75	15.00	18.00	20	p = 0.201
	economists	4	10.25	2.66	5.32	4	5.00	10.50	15.25	16	
	Group Total	74	13.31	.68	5.87	0	9.00	15.00	18.00	20	

Finally, returning to our full sample of subjects, we can consider further some of the factors which may underpin the gender effect observed throughout our analyses of the impact of individual characteristics upon gamble ratings. As noted, previous research has suggested that females are more risk averse than males. Furthermore, the central message of the Slovic et al., findings is that subjects rate gambles without comparators lower than gambles with comparators (a result that we replicate but show holds for non-economists rather than those with advanced economics training). One possible interpretation of the Slovic et al result is that gambles without comparators (and hence without clear affective messages) are difficult to assess. Heightened difficulty is liable to lead to uncertainty and plausible greater risk aversion, particularly amongst those disposed to such an attitude. If indeed females are more risk averse then we might expect to observe a greater disparity between males and females in their rating of gambles without comparators than amongst gambles with comparators. We can readily test such an hypothesis.

Table 13 details tests of significance between the ratings given by females and males of gambles without comparators (Appendix 2 reports such tests for all gambles giving further details of rating distributions). As can be seen in every instance there is a significant gender effect within these ratings.

Table 13: Gender effects within gambles without P element comparators (gambles UK0, UK(01)0, US0 and US3)

Gamble code	Gamble	Mean male rating	Mean female rating	Male v. female (p)
UK0	(7/36, +£9.00)	11.14	6.15	0.022**
UK(01)0	(7/36, +£9.00)	14.74	11.29	0.058*
US0	(7/36, +\$9.00)	10.82	6.96	0.008**
US3	(7/36, +\$90.00)	11.96	8.88	0.028**

The hypothesis is well supported by Table 13. However, in order to be distinguished from the expected enhanced risk aversion among females the hypothesis also requires that female scores be more similar to male scores in those gambles which do have P element comparators. Table 14 reports the significance of gender effects amongst this latter set of gambles.

Table 14: Gender effects within gambles with P element comparators (gambles UK1,UK2, UK(01)1, US1, US2, US4 and US5)

Gamble code	Gamble	Mean male rating	Mean female rating	Male v. female (p)
UK1	(7/36, +£9.00) (29/36, -£0.05)	14.06	12.35	0.276
UK2	(7/36, +£9.00) (29/36, -£0.25)	13.37	11.57	0.304
UK(01)1	(7/36, +£9.00) (29/36, -£0.05)	10.52	9.05	0.436
US1	(7/36, +\$9.00) (29/36, -\$0.05)	14.53	10.77	0.016**
US2	(7/36, +\$9.00) (29/36, +\$0.05)	11.50	11.33	0.976
US4	(7/36, +\$9.00) (29/36, you win nothing)	8.88	9.09	0.810
US5	(7/36, +\$9.00) (29/36, you lose nothing)	14.87	11.70	0.054*

Cursory inspection of Table 14 shows that, in line with our hypothesis, levels of significance do appear to be substantially lower than those reported in Table 13. In short it does appear that the gender difference is most marked within gambles which lack P element comparators. The evidence is not absolutely conclusive and the gender difference

associated with gamble US1 is more reminiscent of Table 13. However, the overall pattern does suggest a potentially systematic difference here which may shed light upon the gender effect. A particularly interesting contrast is provided by the joint assessment sub-sample which yields ratings of gamble UK(01)0 and UK(01)1. Here we observe an approximately significant difference ($p=0.058$) between the ratings given by females (mean = 11.29) and males (mean = 14.74) for the 'without-comparator' gamble UK(01)0. In contrast there is clearly no significant difference ($p=0.436$) between the ratings given by females (mean = 9.05) and males (mean = 10.52) for 'with-comparator' gamble UK(01)1. Given that this is a within-subject test, with responses given just moments apart this seems a strong result which supports our hypothesis.

4. CONCLUSIONS

We have conducted a series of experiments involving simple gambles to test a number of hypotheses. Results from tests of the hypotheses underpinning Equations [1] to [9] all support affective expectations. In the absence of evaluability gamble ratings lacked affective meaning and assessments are inconsistent with expected value differences. Furthermore, even when expected value remains unchanged, apparently small changes in presentation, such as those across the 'redefining nothing' treatments, led to some of the largest changes in rating. Further affective cues such as defining outcomes as say "*only*" losing or gaining some small sum might be expected to further enhance these effects.

We have also found clear evidence of the influence of respondents characteristics upon ratings including differences arising from prior training and gender effects.

The dual dimensions of the impact of evaluability in triggering affect and the influence of personal characteristics upon the degree of that impact are, we contend, fertile ground for future research.

References

- Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D. and Poe, G.L., (2004) On visible choice sets and scope sensitivity, *Journal of Environmental Economics and Management*, 47: 71-93.
- Bateman, I.J., Dent, S., Peters, E., Slovic, P. and Starmer, C., (2006) The Affect Heuristic and the Attractiveness of Simple Gambles, *mimeo*, Decision Research, Eugene, Oregon.
- Bateman, I.J., Munro, A., Rhodes, B., Starmer, C. and Sugden, R. (1997) A test of the theory of reference-dependent preferences, *Quarterly Journal of Economics*, 112(2): 479-505.
- Birnbaum M. H. (1999) How to show that $9 > 221$: Collect judgments in a between-subjects design, *Psychological Methods*, 4 (3): 243-249.
- Hsee, C. K. (1996a) The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes*, 67, 247-257.
- Hsee, C. K. (1996b) Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes*, 66, 122-129.
- Hsee, C. K. (1998) Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, 11, 107-121.
- Kahneman, D. and Knetsch, J.L. (1992) Valuing public goods: The purchase of moral satisfaction, *Journal of Environmental Economics and Management*, 22(1): 57-70.
- Kahneman, D., Knetsch, J L., and Thaler, R. H. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy* 98: 1325-1348.
- Loomes, G. and Sugden, R. (1982) Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty, *Economic Journal*, 92, 805-24.
- Mellers, B. A., Ordóñez, L., & 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, 331-369.
- Mowrer, O. H. (1960a). *Learning theory and behavior*. New York: John Wiley & Sons.
- Mowrer, O. H. (1960b). *Learning theory and the symbolic processes*. New York: John Wiley & Sons.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In T. Gilovich & D. Griffin & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 397-420). New York: Cambridge University Press.
- Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: a reference-dependent model. *Quarterly Journal of Economics* 106: 1039-1061.

Appendix 1: Tests of differences between all gamble treatments

Table A1.1: Significance of difference tests (t and Mann Whitney Z) for all groups. Cells show test statistic and corresponding p value)

GROUP	STAT	UK1	UK2	UK(01)0	UK(01)1	US0	US1	US2	US3	US4	US5
UK0	T	-2.756 (0.007)	-2.341 (0.022)	2.909 (0.005)	0.369 (0.713)	-0.526 (0.600)	2.246 (0.027)	1.723 (0.090)*	0.997 (0.321)	-0.219 (0.828)	3.017 (0.003)
	Z	-2.633 (0.008)	-2.073 (0.038)	-2.866 (0.004)	-0.089 (0.929)	-0.340 (0.734)	-2.238 (0.025)	-1.760 (0.078)	-1.061 (0.289)	-0.094 (0.925)	-2.857 (0.004)
UK1	T		0.450 (0.654)	-0.117 (0.907)	-2.392 (0.019)	-3.863 (0.000)	-0.813 (0.419)	-1.510 (0.136)	-2.315 (0.023)	-3.487 (0.001)	-0.065 (0.948)
	Z		-0.779 (0.436)	-0.286 (0.775)	-2.268 (0.023)	-3.604 (0.000)	-1.137 (0.256)	-1.804 (0.071)*	-2.538 (0.011)	-3.536 (0.000)	-0.328 (0.743)
UK2	T			0.376 (0.708)	-1.942 (0.056)	-3.374 (0.001)	-0.317 (0.752)	-1.039 (0.302)	-1.805 (0.075)	-2.993 (0.004)	0.436 (0.664)
	Z			-0.393 (0.695)	-1.693 (0.091)	-3.114 (0.002)	-0.308 (0.758)	-1.264 (0.206)	-1.917 (0.055)	-2.952 (0.003)	-0.491 (0.624)
UK(01) 0	T				-2.505 (0.014)	-4.098 (0.000)	-0.761 (0.448)	-1.575 (0.119)	-2.400 (0.018)	-3.687 (0.000)	0.059 (0.953)
	Z				-2.443 (0.015)	-3.968 (0.000)	-0.635 (0.525)	-1.688 (0.091)	-2.434 (0.015)	-3.366 (0.001)	-0.111 (0.912)
UK(01) 1	T					-0.969 (0.335)	-1.865 (0.065)	1.305 (0.196)*	0.541 (0.590)*	-0.646 (0.520)*	2.630 (0.010)
	Z					-0.543 (0.587)	-1.743 (0.081)	-1.375 (0.169)	-0.657 (0.511)	-0.516 (0.606)	-2.485 (0.013)
US0	T						-3.312 (0.001)	-2.809 (0.006)	-1.840 (0.069)	-0.358 (0.721)	-4.235 (0.000)
	Z						-3.208 (0.001)	-2.567 (0.010)	-1.938 (0.053)	-0.484 (0.629)	-4.012 (0.000)
US1	T							0.753 (0.453)	1.600 (0.113)	2.919 (0.004)	-0.840 (0.403)
	Z							-1.034 (0.301)	-1.851 (0.064)	-2.941 (0.003)	-0.785 (0.432)

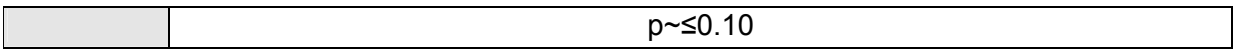
US2	T								0.937 (0.351)	2.385 (0.019)	-1.667 (0.099)
	Z								-0.767 (0.443)	-2.053 (0.040)	-1.955 (0.051)
US3	T									1.444 (0.152)	-2.508 (0.014)
	Z									-1.320 (0.187)	-2.667 (0.008)
US4	T										-3.823 (0.000)
	Z										-3.636 (0.000)

* If the variances of the two groups are significantly different (using the Levene Test; Puri, 1996) the t-test used assumes unequal variances and is indicated by an asterix.

Table A1.2: Significance of difference tests (t and Mann Whitney) for all groups. Cell shading shows level of significance (light grey = 90%, dark grey = 95% confidence)

GROU P	UK1	UK2	UK(01)0	UK(01)1	US0	US1	US2	US3	US4	US5
UK0	Affect	Affect	Affect (v. comp)	NE	Affect	Affect	NE/Affect?	Affect)) Affect	Neg LA
UK1		NE	NE	Affect (v. comp)	Affect	Affect	NE/Affect?	Affect) Affect)	NE
UK2			NE	Affect	Affect	NE	NE/Affect?	Affect	Affect	NE
UK(01)0				EU (OTT)	EU (OTT)	NE	NE	} } Affect	# # Affect v. comp	%Affect v. comp & Neg LA
UK(01)1					NE	NE	NE	}Affect }	# Affect v. comp #	%Affect v. comp & Neg LA
US0						Affect	Affect] Affect] Affect	@ @ Affect	@ @ Affect
US1							*] Affect]	Affect	* NE (affect=Neg LA)
US2								Affect	Affect/EU	* NE
US3									Affect	Affect + Neg LA
US4										Neg LA

p<0.05 - All t and Z stats fall in the same categories within any group v group comparison



Notes for Table A1.2:

Affect = finding is in line with Affect Heuristic expectations / contrary to EU expectations (note that this includes unshaded boxes, i.e. here the Affect hypothesis expects no significant difference)

Affect/EU = in line with congruent Affect and EU expectations

EU = in line with EU expectations – exhibiting reasonable responsiveness to scope

EU (OTT) = in line with EU expectations – but exhibits unreasonable (over-the-top) responsiveness to scope

NE = no Affect or EU expectations

Neg LA = Negated loss aversion (subjects told that they will not lose).

},],), #, @, % = here connecting cells form joint hypothesis

* = see discussion of this result in main text.

Appendix 2: Investigating gender effects within gambles

Table A2.1: Descriptive statistics for gamble ratings (0-20 response scale) by group and gender.

Group	Gamble No.	Gamble	Subset	Count	Mean	s.e. mean	s.d.	Min.	Q1	Median	Q3	Max.	Male v. female (p)
UK0	6	(7/36, +£9.00)	male	22	11.14	1.29	6.06	0	5.75	11.50	16.00	20	0.022**
			female	13	6.15	1.30	4.67	0	2.00	5.00	9.00	15	
			Total	35	9.29	1.02	6.03	0	5.00	8.00	14.00	20	
UK1	7	(7/36, +£9.00) (29/36, -£0.05)	male	16	14.06	1.65	6.62	0	7.25	16.50	18.75	20	0.276
			female	20	12.35	1.31	5.87	0	8.00	13.50	17.00	20	
			Total	36	13.11	1.03	6.18	0	8.00	15.50	18.00	20	
UK2	8	(7/36, +£9.00) (29/36, -£0.25)	male	19	13.37	1.35	5.88	0	10.00	15.00	18.00	20	0.304
			female	14	11.57	1.43	5.35	4	5.00	12.50	16.00	20	
			Total	33	12.61	.98	5.65	0	9.00	14.00	18.00	20	
UK(01)0	9	(7/36, +£9.00)	male	23	14.74	.92	4.39	8	10.00	15.00	20.00	20	0.058*
			female	21	11.29	1.36	6.24	0	7.50	10.00	16.50	20	
			Total	44	13.09	.84	5.57	0	8.25	14.00	18.50	20	
UK(01)1	10	(7/36, +£9.00) (29/36, -£0.05)	male	23	10.52	1.37	6.59	1	5.00	10.00	19.00	20	0.436
			female	21	9.05	1.48	6.76	0	3.50	6.00	15.00	20	
			Total	44	9.82	1.00	6.64	0	4.25	10.00	15.00	20	
US0	0	(7/36, +\$9.00)	male	22	10.82	1.16	5.46	0	6.00	12.50	14.25	20	0.008**
			female	28	6.96	.70	3.73	0	4.25	6.50	9.75	15	
			Total	50	8.66	.69	4.91	0	5.00	7.50	13.00	20	
US1	1	(7/36, +\$9.00) (29/36, -\$0.05)	male	19	14.53	1.22	5.34	0	11.00	16.00	19.00	20	0.016**
			female	31	10.77	1.00	5.59	0	7.00	11.00	16.00	20	
			Total	50	12.20	.81	5.74	0	8.00	13.50	16.25	20	
US2	2	(7/36, +\$9.00) (29/36, +\$0.05)	male	20	11.50	1.28	5.71	3	6.00	11.00	17.50	20	0.976
			female	30	11.33	.78	4.27	3	8.00	11.00	14.00	20	
			Total	50	11.40	.68	4.84	3	7.75	11.00	14.25	20	

US3	3	(7/36, +\$90.00)	male	26	11.96	1.06	5.43	0	7.75	13.00	15.50	20	0.028**
			female	24	8.88	.81	3.95	0	5.25	9.00	12.75	15	
			Total	50	10.48	.70	4.98	0	6.75	10.00	14.00	20	
US4	4	(7/36, +\$9.00) (29/36, you win nothing)	male	16	8.88	1.46	5.83	0	4.25	8.50	13.25	20	0.810
			female	34	9.09	.83	4.86	0	4.00	10.00	13.00	16	
			Total	50	9.02	.73	5.13	0	4.00	9.50	13.00	20	
US5	5	(7/36, +\$9.00) (29/36, you lose nothing)	male	23	14.87	1.12	5.36	1	11.00	15.00	20.00	20	0.054*
			female	27	11.70	1.08	5.64	0	7.00	14.00	15.00	20	
			Total	50	13.16	.80	5.68	0	10.00	14.50	18.00	20	

Differences between male and female ratings assessed using Mann Whitney U test. * = significant at $0.10 > p > 0.05$; ** = significant at $p < 0.05$

