

PREFERENCE REVERSAL: AN EXPERIMENTAL INVESTIGATION OF ECONOMIC AND PSYCHOLOGICAL HYPOTHESES*

by

Robin Cubitt**

Alistair Munro**

Chris Starmer***

* Research for this paper was funded by the Economic and Social Research Council, under its award no. L211252053. The project funded by this award formed part of the Council's research programme on Risk and Human Behaviour. We are grateful for their comments on earlier versions of the paper to Graham Loomes, Robert Sugden and participants at the seminars and conferences at which we have presented the paper.

** School of Economic and Social Studies, University of East Anglia, Norwich, NR4 7TJ, United Kingdom.

*** School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom.

Abstract

We present a new experimental investigation into the cause of preference reversal. Although economists and psychologists have suggested a variety of possible accounts, the existing data does not adequately discriminate between them. Our design has two key novelties: a “pure” ordinal payoff scheme and a probabilistic valuation task. Relative to previous studies, this design offers enhanced control for economic explanations and new tests of psychological hypotheses. We find strong and systematic preference reversals that are inconsistent with a broad range of previous hypotheses. We explore and evaluate the explanatory strategies that survive exposure to our data.

Keywords

Preference reversal

JEL classifications

C91; D81; A12.

Section I: Introduction

A preference reversal arises when the relative evaluation of two (or more) items by an individual agent depends systematically on the means with which it is elicited. Although first observed by the psychologists Lichtenstein and Slovic (1971) and Lindman (1971), the potential significance of the phenomenon was soon recognised by economists. Grether and Plott, the first to report preference reversal in the economics literature, famously commented: "Taken at face value the data are simply inconsistent with preference theory and have broad implications about research priorities within economics. The inconsistency suggests that no optimisation principles of any sort lie behind even the simplest of human choices" (Grether and Plott, 1979, p. 623).

Since the 1970s, both economists and psychologists have conducted a considerable amount of research on preference reversal, suggesting several possible explanations of it. However, the two groups of researchers have pursued quite different explanatory strategies. Economists have not always taken the data at face value and have proposed explanations of preference reversal which are consistent with the assumption that behaviour is governed by *context-free preferences*. Psychological theories, on the other hand, deny this basic assumption of standard decision theory and instead assume that preferences are *context-sensitive*. Since each approach has been adapted to explain preference reversal in the experimental designs that have typically been used to observe it, the majority of existing data does not discriminate between them.¹

In this paper, we present a new experiment whose design discriminates between psychological hypotheses, whilst incorporating improved controls for economic explanations. To the extent that a

systematic pattern of preference reversals is observed using this design, this would tell against most economic accounts of earlier findings and allow us to assess rival psychological hypotheses. The crucial features of the design are the use of a pure form of an incentive mechanism known as the ordinal payoff scheme and of a new type of valuation task, which we will call probabilistic valuation, along side more conventional tasks. In addition, the design features relatively strong incentives and uses dominated choice options to proxy the extent of subject error.

The remainder of the paper is organised as follows. Section II characterises the economic and psychological hypotheses tested in our study. Section III sets out our experimental design and explains the implications of various theories within it. Section IV describes the results: we observe strong and systematic preference reversals which cannot be explained by traditional economic theories; our results also allow us to reject three important psychological explanations. Since none of the theories under consideration successfully organizes the main patterns in the data, when taken in isolation, we consider what theoretical avenues remain open for explaining them. Section V concludes.

Section II: Economic and Psychological Explanations of Preference reversal

The best known cases of preference reversal have involved decisions relating to pairs of simple monetary gambles often called P- and \$-bets. The P-bet offers a relatively large chance of a modest prize while the \$-bet offers a smaller chance of a larger prize. In the typical experiment, a given subject makes a straight choice between the P- and \$-bets and also states a monetary valuation for each of them. The classic finding² is a systematic tendency for subjects to choose the P-bet over the \$-bet in the choice task, but to place a strictly higher monetary value on the \$-bet. We will call such a tendency

standard preference reversal. We will also call a tendency in the opposite direction (i.e. \$-bet chosen in the choice task, but P-bet valued more highly) *counter preference reversal*. Although they would be equally puzzling from the point of view of standard decision theory, counter reversals have not been widely observed.

Standard reversals have often been observed in experiments that use one or both of the following experimental procedures: the Becker-De Groot-Marschak mechanism (BDM) and the random lottery incentive system (RLIS). The BDM is a device intended to provide incentives for experimental subjects to reveal true certainty equivalents in valuation tasks (as opposed to, say, inflated selling prices which strategically overstate true valuations)³. The RLIS is often used in experiments, like those on preference reversal, where subjects must undertake multiple tasks⁴. It is intended to encourage subjects to think carefully about each task, while controlling for income effects which might otherwise arise if subjects were paid for each task. Many of the reported preference reversal experiments have used both of these mechanisms. As a convenient short hand, we will say that an experiment which does so has the *canonical design*.

II.i. Economic Explanations

Typically, economic explanations of observed preference reversals assert the existence, for each individual, of a set of context-free preferences over gambles, which govern behaviour in both choice and valuation tasks. How could such a theory account for observations of standard reversals?

One important subset of economic explanations implicates the incentive mechanisms used in the

canonical design. A series of papers published in the late 1980s demonstrated that, unless agents' preferences satisfy the axioms of expected utility theory, preferences recorded using the canonical design may be biased and that, in principle, the bias could take the form of the standard preference reversal. For instance, Holt (1986) has shown that, if the independence axiom does not hold, standard reversal could arise from a bias induced by the RLIS. Alternatively, it could arise from distortion of the valuations elicited by the BDM, in the presence of violations of the independence axiom (Karni and Safra, 1987), or in the presence of violations of the reduction principle for compound lotteries (Segal, 1988).

An alternative explanation is provided by regret theory (Loomes and Sugden, 1983; Loomes, Starmer and Sugden 1989, 1991). It permits a more fundamental tenet of standard decision theory to be violated, namely transitivity. Let P , $\$$ and $V_m(g_i)$ denote, respectively, the P -bet, the $\$$ -bet, and the subject's monetary valuation of gamble g . Taking the valuations and choices elicited at face value, a subject who displays a standard reversal reveals preferences satisfying $V_m(P) \sim P \succsim \$ \sim V_m(\$)$ and valuations satisfying $V_m(\$) > V_m(P)$. Given an assumption that more money is preferred to less, this implies a preference cycle over $V_m(P)$, P , $\$$ and $V_m(\$)$.

Whilst the explanations described in the previous two paragraphs differ by relaxing different axioms of expected utility theory, each asserts the existence of context-free preferences that satisfy *some* properties of coherence. The minimal requirement of coherence, imposed by each of them, is that preferences are describable by a *choice function*, i.e. a function $C(\cdot)$ which picks out, for any non-empty feasible set F of gambles, a non-empty preferred subset of F , and that this choice function

governs behaviour in all tasks. All of the theories which economists have proposed to explain preference reversal are so describable, and we will refer to any such theory as a *traditional economic theory*.

II.ii. Psychological Explanations

The psychological approach departs more radically from standard decision theory than any of those just mentioned. Rather than simply denying particular axioms on preferences, it denies the very existence of coherent, context-free, preferences. Instead, it asserts either that preferences are constructed as a consequence of the performance of particular tasks, with different classes of task potentially leading to the construction of different preferences, or that, while preferences may exist prior to the tasks, there are different preferences for different classes of task. We will call either of these assertions an assertion of the *context-sensitivity* of preferences. If preferences are context-sensitive then, even if the preferences applying within (resp. constructed in response to) particular classes of task are themselves coherent, this provides no reason to expect coherence between responses to tasks in different classes. If choice tasks and monetary valuation tasks elicit different preferences, this would explain preference reversal.

In this paper, we report tests of three variants of this approach, known as the prominence, scale compatibility and task goal hypotheses, respectively. These hypotheses are closely related to each other and have common origins in the work of Paul Slovic, Amos Tversky, their collaborators and subsequent researchers. Despite this, the reader should not be misled into thinking that the different hypotheses are equivalent: as we will show and their proponents have noted, they are logically distinct.

As the hypotheses were not formulated specifically to explain preference reversals between monetary gambles, we will present them initially as more general claims about tasks involving multi-attribute items.

For the purposes of this discussion, we imagine that there is a domain of items each of which is describable in terms of a set of attributes. The set of attributes is the same for each item in the domain, but different items may differ in their quality with respect to any given attribute. A maintained hypothesis of the psychological approach is that a subject's performance of a task involving two or more items in the domain will depend in some way on their attributes.

The *prominence hypothesis* (Tversky, Sattath and Slovic, 1988)⁵ is concerned with the relationship between choice and matching tasks. A matching task is a task in which the subject is faced with two multi-attribute items. With one exception, all of the attributes of both items are given. But, there is one (quantifiable) attribute for which the value of one of the items is unspecified. The task is to set this missing value so that the two items are equally attractive. The prominence hypothesis presupposes the existence of an attribute which, in some sense⁶, is the most important. The hypothesis is that the prominent attribute weighs more heavily in choice tasks than in matching tasks.

One explanation for this hypothesis is that a choice task is qualitative, whereas a matching task is quantitative. Thus, it may be that choice tasks cue qualitative reasoning strategies designed to find a decisive reason to choose one way or the other, thereby bringing the prominent attribute to the fore,⁷ whereas matching tasks cue quantitative reasoning strategies involving tradeoffs between attributes. This explanation of the prominence hypothesis was initially suggested by Tversky, Sattath and Slovic (1988);

Fischer and Hawkins (1993) articulate it in more detail and call it the *strategy compatibility hypothesis*⁸. It should not be confused with the scale compatibility hypothesis which we discuss next.

The *scale compatibility hypothesis* (Tversky, Sattath and Slovic, 1988)⁹ focuses on the fact that a task, of any kind, has a response mode. A particular response mode may be compatible with certain types of attribute; for example, the attribute and the response mode might share the same scale. In general, response modes in different tasks may differ in their compatibility with any given attribute. The hypothesis is that the weight of a given attribute in any task is greater the more compatible that attribute is with the response mode of the task.

The *task goal hypothesis* (Fischer *et al*, 1999), like the prominence hypothesis, postulates the existence of a prominent attribute. Its novelty is to suggest that the weight accorded to this attribute in a task depends upon the goal of the task, as perceived by the subject. The hypothesis is that the prominent attribute is weighted more heavily in tasks whose perceived goal is to differentiate between items than in tasks whose perceived goal is to equate them.

The standard pattern of preference reversal between the P-bet and the \$-bet can be explained by the prominence hypothesis if “win-probability” is the prominent attribute. According to this view, choices are more affected by differences in win-probability than are monetary valuations. Thus, there is a tendency for the P-bet to be chosen even when a higher monetary valuation is put on the \$-bet. However, the same evidence can be explained by the scale compatibility hypothesis, given the claim that the attribute “money-prize” is more compatible with a monetary valuation scale than it is with the

response mode of a choice task. On this view, subjects pay more attention to the size of the monetary prizes when putting monetary values on gambles than they do when choosing between them. This induces a tendency to value the \$-bet more highly even when the P-bet would be chosen in a straight choice. Finally, the evidence can also be explained by the task goal hypothesis, provided that subjects see the goal of choice tasks, but not of valuation tasks, as being to differentiate between the P-bet and the \$-bet. Given this, the task goal hypothesis accords with the prominence hypothesis in predicting standard preference reversal because choices are more affected by differences in win-probability than are monetary valuations.

It should not be surprising that these psychological hypotheses yield identical predictions in experiments using the canonical design: they are closely related and each was designed to account for a body of data that includes the results of such experiments. However, the hypotheses remain logically distinct and can, in principle, be tested against one another, as we will explain below.

Section III: A New Experiment

Before presenting the details of our experiment, it will be helpful to discuss two ingredients of its design in isolation. Each is important for discriminating between different explanations of preference reversal; and each is an important departure from the canonical design.

III.i The Ordinal Payoff Scheme

The ordinal payoff scheme was proposed by Tversky, Slovic and Kahneman (1990) (henceforth, TSK), though both our use and our analysis of it differ from theirs in important respects. Suppose there

are n *parameter sets*, each of which consists of a P-bet and a \$-bet. The subject is told that she will face three tasks in relation to each parameter set: she will be asked to choose between the P-bet and the \$-bet; and she will be asked to put a value on each of them separately. After she has performed these tasks for all n parameter sets, a random device will select one of the parameter sets, each with probability $1/n$. Then, another random device will select either CHOICE or VALUATION, each with probability $1/2$. If it selects CHOICE, she will receive whichever gamble she chose in the choice task for the selected parameter set; if it selects VALUATION, she will receive whichever of the two gambles she put a higher value on (with a coin-toss as a tiebreaker, in the event of equal valuations). Whichever gamble she receives is then played out for real¹⁰.

The ordinal payoff scheme offers considerable improvement in experimental control relative to the canonical design. The latter uses the BDM device and the RLIS as incentive mechanisms. While the combination of these devices is incentive compatible for the elicitation of *absolute* valuations, provided expected utility theory holds, economic explanations have cast doubt on the reliability of the absolute valuations so elicited, by invoking violations of expected utility theory. However, if we observe any systematic pattern of preference reversal in an experiment which uses the ordinal payoff scheme, this cannot be explained by any traditional economic theory¹¹, even one that relaxes expected utility theory.

To understand this property of the ordinal payoff scheme, consider first the special case in which $n = 1$. Considering all possible responses to choice and valuation tasks in this case, the subject faces the feasible set $F = \{P, \$, L_1, L_2, L_3\}$ where $L_1 = (P, 0.5; \$, 0.5)$, $L_2 = (P, 0.5; (P, 0.5; \$, 0.5), 0.5)$ and $L_3 = (\$, 0.5; (P, 0.5; \$, 0.5), 0.5)$. P (resp. \$) can be obtained by choosing P (resp. \$) in the choice

task and giving it a strictly higher value; L_1 can be obtained by choosing one gamble in the choice task and giving the other gamble a strictly higher value; L_2 (resp. L_3) can be obtained by choosing P (resp. \$) in the choice task and giving equal valuations to the two gambles. To give responses that display a preference reversal between choice and valuation corresponds to selecting L_1 from this feasible set.

Notice that, within this design, the observation of an *instance* of preference reversal is not inconsistent with a traditional economic theory. For example, an expected utility maximiser who was indifferent between P and \$ might select lottery L_1 ; and a non-expected utility maximiser whose preferences do not satisfy betweenness¹² might have a strict preference for L_1 over both P and \$. There are, however, two distinct ways in which an agent may select lottery L_1 : either by choosing P and valuing \$ more highly (the standard reversal); or by choosing \$ and valuing P more highly (the counter reversal). Since these two ways of responding constitute selecting exactly the same element of F, traditional economic theories applied to this feasible set of gambles provide no reason for expecting preference reversals in one direction to occur more frequently than their mirror image.

A similar argument extends to the ordinal payoff scheme with $n > 1$. Let P_i and $\$i$ represent the P-bet and the \$-bet in the i th parameter set ($i \in \{1, \dots, n\}$). Consider the position of a subject trying to decide what responses to give to the tasks involving the j th parameter set, given her responses to the other $3(n-1)$ tasks. Let L^* represent the compound gamble which the subject will face if the first randomisation of the ordinal payoff scheme selects a parameter set other than j ; L^* is determined by the subject's responses to the tasks involving parameter sets other than j . The subject's complete set of responses in the experiment will determine a compound gamble of the form $(\cdot, 1/n; L^*, (n-1)/n)$: the missing term

depends on the responses to the parameter set j tasks.

Notice that, unless the independence axiom of expected utility theory holds, it is *not* true in general that the ordinal payoff scheme reliably elicits choices and (even relative) valuations in the j th parameter set.

Consider an agent whose true relative valuations accord with her choices and who prefers P_j to $\$j$. By choosing P_j and giving it a strictly higher value she could achieve $(P_j, 1/n; L^*, (n-1)/n)$ whereas, by choosing $\$j$ and giving it a strictly higher value, she could achieve $(\$j, 1/n; L^*, (n-1)/n)$. However, it does not follow from the fact that she prefers P_j to $\$j$ that she also prefers $(P_j, 1/n; L^*, (n-1)/n)$ to $(\$j, 1/n; L^*, (n-1)/n)$. The claim that it did so follow, for all $P_j, \$j, L^*$ and n , would be a statement of the independence axiom of expected utility theory. Thus, any claim that the ordinal payoff scheme reliably elicits relative valuations (or choices) within a given parameter set, when $n > 1$, rests on independence.

Nevertheless, we may conduct a test for the presence of a systematic preference reversal using the ordinal payoff scheme *without* invoking independence. To see how, note that the subject's responses to the j th parameter set reveal which element of the set F_j of compound gambles she will receive, where:

$$F_j = \{(P_j, 1/n; L^*, (n-1)/n), (\$j, 1/n; L^*, (n-1)/n), (L_{j1}, 1/n; L^*, (n-1)/n), (L_{j2}, 1/n; L^*, (n-1)/n), (L_{j3}, 1/n; L^*, (n-1)/n)\}$$

and

$$L_{j1} = (P_j, 1/2; \$j, 1/2), L_{j2} = (P_j, 1/2; (P_j, 1/2; \$j, 1/2), 1/2), L_{j3} = (\$j, 1/2; (P_j, 1/2; \$j, 1/2), 1/2).$$

The first (resp. second) listed element of F_j can be obtained by choosing P_j (resp. $\$j$) in the choice task

for parameter set j and giving it a strictly higher value than $\$j$ (resp. P_j) in the valuation tasks; the third listed element of F_j can be obtained by choosing one gamble in the choice task for parameter set j and giving the other gamble a strictly higher value; the fourth (resp. fifth) listed element of F_j can be obtained by choosing P_j (resp. $\$j$) in the choice task for parameter set j and giving equal valuations.

Giving responses that display a preference reversal in parameter set j corresponds to selecting $(L_{j1}, 1/n; L^*, (n-1)/n)$ from F_j . If subjects are indifferent between P and $\$$ (or have a strict preference for randomization between them), preference reversal is possible; but, a traditional economic theory provides no reason to expect a *systematic* tendency for it to occur in one direction rather than the other. Thus, if one observes a systematic tendency towards preference reversal in a particular direction, in any parameter set of an experiment that uses the ordinal payoff scheme, this observation cannot be explained consistently with the claim that subject's behaviour is governed by context-free preferences that are describable by a choice function which applies to the feasible set of gambles determined by the parameters of the whole experiment.

A second important feature of the ordinal payoff scheme, for our purposes, is that it drives a wedge between the *form* of the valuation tasks and their goal. Each valuation task has the form of a matching task; but the goal of the valuation tasks is to differentiate between the gambles, just as the goal of the choice task is (recall that the gamble received by the subject depends on chance, her choice, and her *relative* valuations only). The implication is that, provided the subject perceives this, she will regard the goals of the choice task and of the pair of valuation tasks as identical. Given these perceptions, the task goal hypothesis would not predict preference reversal between choice and valuations¹³. In

contrast, the prominence hypothesis would predict it, because the valuation tasks are matching tasks.

It is possible that subjects' perceptions of the goals of the valuation tasks might be affected both by their differentiating function and by their matching form. On this view, and to the extent that perceptions of the goal are influenced by the form of the task, the task goal hypothesis would permit some tendency towards preference reversal in the same direction as the prominence hypothesis.

III.ii Monetary and Probabilistic Valuation

Throughout our discussion of the ordinal payoff scheme in the previous sub-section, we have referred to choice and valuation tasks. Although TSK used monetary valuation tasks, nothing in our discussion required valuation tasks to use a monetary scale, rather than some other scale of value. The second distinctive ingredient of our design is the use, in different groups of subjects, of two types of valuation task which differ in terms of the valuation scale.

All of the P and \$-bets used in our design are what we will call *basic gambles*, i.e. each gives a single strictly positive monetary prize with some probability and zero otherwise. Any basic gamble g_i can be represented by an ordered pair (m_i, p_i) of attributes, whose first and second elements are, respectively, the money-prize and win-probability of g_i . In all of our valuation tasks, the subject is presented with a *yardstick*, which is a basic gamble with one of its attributes unspecified. In a monetary valuation task, the missing attribute of the yardstick is its money-prize while the win-probability of the yardstick is fixed at unity. The task of giving a monetary valuation of some gamble g_i is to set the money-prize in the yardstick so that it and g_i are equally attractive. The subject's response is her *monetary valuation* of

g_i . (Thus, straightforwardly, the monetary valuation of g is its certainty equivalent.) But, in a probabilistic valuation task, the missing attribute of the yardstick is its win-probability and the money-prize of the yardstick is fixed at a value T , where T is constant across all probabilistic valuation tasks.

The task of giving a probabilistic valuation of some gamble g is to set the win-probability in the yardstick so that it and g are equally attractive. The subject's response is her *probabilistic valuation* of g_i . Regardless of whether they are monetary or probabilistic, the valuations of any two gambles elicited by the same type of valuation task can be directly compared to elicit a binary preference between the gambles.

Our experimental design uses the ordinal payoff scheme in each of two groups. (Subjects were assigned to groups by a random device.) Each group faced the same parameter sets but, in one group, the valuation tasks were monetary whereas, in the other group, they were probabilistic. To our knowledge, this experiment is the first to examine the incidence of preference reversal comparing probabilistic and monetary valuations while exploiting the controls offered by the ordinal payoff scheme¹⁴. As we now explain, this combination of design features allows particularly well-controlled tests among various groups of hypotheses.

From the perspective of the traditional economic approach, each group in our design faces an *identical* overall task. It is to respond to n pairwise choice questions and to assign a number to each of $2n$ gambles. The absolute magnitudes of the numbers are irrelevant; all that matters is their ordering within each parameter set. The scale in which valuations are given should make no difference. Thus, no account of preference reversal which takes the traditional economic approach, and which treats the

experiment as a single whole, would predict a systematic difference between the behaviour of our two groups, in terms either of the responses to the choice tasks or of the pairwise preferences implied by the valuation tasks.

From the perspective of the psychological hypotheses, it is important to note that in our design both probabilistic and monetary valuation tasks are matching tasks¹⁵. Thus, the prominence hypothesis implies that the prominent attribute should weigh more heavily in choice tasks than in *either* type of valuation task. If prominence is the dominant source of preference reversal, any systematic tendency towards preference reversal in a given parameter set should be in the *same* direction in both groups. Moreover, if the prominence hypothesis is extended to imply that the relative weight of attributes is the same in any matching task, then there should be no systematic difference between the two groups, in terms of the binary preferences implied by their valuations.

On the assumption that the perceived goal of the valuation tasks is independent of the valuation response scale, the task goal hypothesis also implies that there should be no systematic difference between groups. Note that this holds regardless of whether the goal of the valuation tasks is seen as differentiation or equation. Thus, this prediction of the task goal hypothesis does not depend on subjects perceiving the goal of the valuation tasks as being to differentiate. It merely requires that there is no systematic difference between groups in the perception of the goal.

Alone among the hypotheses which we have discussed, the scale compatibility hypothesis asserts that the response scale *itself* will affect behaviour. On the assumptions that money-prize is more compatible

with monetary valuation than with the choice response mode and that win-probability is more compatible with probabilistic valuation than with choice, the scale compatibility hypothesis predicts a systematic tendency for preference reversals to be in the *standard* direction (P chosen, \$ valued more highly) in the group doing monetary valuations and to be in the *opposite* direction (\$ chosen, P valued more highly) in the group doing probabilistic valuation.

The above discussion has emphasised the logical differences between the psychological hypotheses. However, it is possible that psychological influences might interact. We will consider this possibility in Section IV.ii below.

III.iii Details of design

As explained above, our experiment used an ordinal payoff scheme in each of two groups. The parameter sets were the same in each case; and we set $n = 6$ and $T = £10$. Thus, each group of subjects faced a total of 18 tasks involving the six parameter sets; the choice tasks were common to the two groups, as were the gambles to be valued. The *only* difference between the two treatments was the form of the valuation tasks. One group (the MV-group) performed monetary valuations; the other group (the PV-group) performed probabilistic valuations. All tasks were presented, one at a time, on a computer screen. The order was randomised separately for each subject. The six parameter sets are described in Table 1.

TABLE 1 ABOUT HERE

Each row of Table 1 describes a pair of gambles. Each gamble is basic; each win-probability is expressed as a number of chances, out of 100, of winning the unique strictly positive prize (which is

expressed in pounds sterling). Parameter sets 1 to 5 each contain a P-bet and a \$-bet. In the interests of comparability, these five parameter sets were based on parameters used by TSK. We replaced US dollars with pounds sterling and doubled the nominal amounts. This preserves the ratios of prizes, whilst substantially increasing their average real value¹⁶. Win-probabilities are unchanged from TSK, except for rounding. Parameter set 6 is different: notice that its two gambles offer the same prize, but one of them (the dominating option) gives more chances of winning the prize than the other. We will comment below on the significance of this parameter set.

For each group, six of the tasks were choices between the pairs of bets described in the rows of Table

1. An example screen display for a choice task is:

FIGURE 1 EXACTLY HERE

At the start of the experiment, subjects were told that they would play one gamble for real at the end. They were also shown a bag containing a set of discs numbered from 1 to 100. Thus, in this example, to play option A for real, the subject would draw a disc from the bag: a number from 1 - 31 would win £32, a number above 31 would win nothing.

Each subject also faced twelve valuation tasks. To record a value for a particular gamble g , the subject was required to fill in the missing attribute in an incompletely specified yardstick lottery in such a way that the yardstick and g_i were equally attractive to them. In the MV-group, the missing attribute in the yardstick lottery was its money-prize and its win-probability was fixed at unity; in the PV-group, win-probability was the missing attribute in the yardstick lottery and its prize was fixed at £10. An example screen display for a valuation task in the MV-group is:

FIGURE 2 EXACTLY HERE

An example screen display for a valuation task in the PV-group is:

FIGURE 3 EXACTLY HERE

Thus, for one group valuations are amounts of money, while for the other they are probabilities, expressed as chances out of 100.

In each group, the ordinal payoff scheme was implemented as follows. Before completing the tasks, subjects were told that, at the end of the experiment, they would each individually role a die to select one of the six parameter sets, that they would play for real one of the two bets in this parameter set, and that which one it was would be determined by a further throw of the die. If this second throw produced 1-3, they would play whichever of the selected parameter set they had chosen when faced with the pairwise choice between them; if it produced 4-6, they would play whichever of the pair they had valued more highly.

As noted above (see footnote 10), TSK used a variant of the ordinal payoff scheme in which, for a subject in one group, all gambles - including the ordinal payoff scheme itself - were hypothetical and, for a subject in the other group, a real gamble would be faced with a 15% probability. We used the scheme in its pure form, in which *every* subject plays for real some gamble, selected by the ordinal payoff scheme in the light of their task-responses. We also substantially increased the average expected value of gambles. While we would not claim that the use of weaker incentives necessarily invalidates TSK's conclusions, our implementation of the ordinal payoff scheme in its pure form, together with increased financial incentives, provides a test of the robustness of TSK's results. This is significant,

given the important role played by the scheme in controlling for explanations of preference reversal offered by traditional economic theories.

A total of 230 subjects - predominantly undergraduate students from a wide range of disciplines at the University of East Anglia - took part across 22 separate sessions involving up to 13 subjects in each. Adjacent sessions were paired and a coin was tossed, in advance, to decide which of the pair would be in the MV-group and which in the PV-group. As a result of this randomization, the MV-group finally contained 114 subjects and the PV-group 116 subjects.

Before discussing the results, we summarize the implications of the various hypotheses in Table 2. The table identifies whether a *systematic* tendency towards preference reversal in a particular direction is allowed, by each of the hypotheses, in the group whose valuation tasks were monetary and in the group whose valuation tasks were probabilistic.

TABLE 2 ABOUT HERE

Section IV: Results and Interpretation

IV.i Tests of individual economic and psychological hypotheses

The main results from our experiment are summarised in Table 3: each column of the table presents the results for a particular parameter set and each column is subdivided for the MV- and PV-groups. The first two rows of data report subjects' responses to choice questions, recording the percentage choosing the P- and \$- bets. Note that, strictly speaking, this P/\$ terminology does not apply for set 6. To begin with we focus on the data for parameter sets 1-5 only. Row three reports the sample size.

TABLE 3 ABOUT HERE

The last four rows of Table 3 combine data on choice and valuation. Let $V(\cdot)$ be the valuation that an individual places on a gamble and let \succsim_c be the relation “is chosen over”. For a given parameter set, any individual’s responses can then be classified into one of four categories:

- | | | |
|------|---|--------------------------------|
| I: | $V(P) \geq V(\$)$ and $P \succsim_c \$$ | (consistent preference for P) |
| II: | $V(\$) \geq V(P)$ and $\$ \succsim_c P$ | (consistent preference for \$) |
| III: | $V(\$) > V(P)$ and $P \succsim_c \$$ | (standard reversal) |
| IV: | $V(P) > V(\$)$ and $\$ \succsim_c P$ | (counter reversal) |

The table gives the number of responses in each category, for each parameter set and each group. A subject in category I (resp. II) reveals a consistent preference for the P-bet (resp. \$-bet) across the choice and valuation tasks. A subject in category III selects the P-bet in choice, but values the \$-bet more highly: this is the standard reversal which has been frequently observed in other studies. A subject in category IV selects \$ in choice but values P more highly: this is the counter reversal. Notice that in each parameter set, a significant proportion of subjects choose each of the options (i.e., a minimum of around 25%): this is a necessary feature of the data if we are to have the possibility of observing reversals in each direction.

The first question we address is whether standard reversal is observed in the MV-group. In this group, standard reversals account for around 35% of all patterns; they also occur much more frequently than the counter reversal (as a proportion of total reversals, the standard reversal rate varies between 85 and

96 per cent).

Table 4 presents the results of some more formal statistical tests, for parameter sets 1-5. Tests 1 and 2 in Table 4 examine whether the pattern of behaviour under choice and monetary valuation is the same. More exactly, for the MV group, test 1 examines the hypothesis that the proportion of subjects who value the \$-bet more highly is the same under choice and valuation. The null hypothesis is rejected at the 1% level in all cases, in favour of the alternative, that the proportion who value the \$-bet more highly is greater than the proportion who choose the \$ bet. Test 2 is a between-groups test of the same hypothesis, comparing the proportion of subjects who value the \$-bet more highly in the MV group with the proportion who choose the \$-bet in the PV group. Again, in all cases the null hypothesis is rejected at the 1% level. These results provide very strong evidence of systematic preference reversal and are broadly consistent with the findings of TSK (recall that the conditions of the MV group closely replicate TSK). In short, the results are inconsistent with the implications of any traditional economic theory.

TABLE 4 ABOUT HERE

We now consider the evidence for the PV-group. Counter reversals are frequent in this group, but only slightly more common in aggregate than standard reversals (126 in total against 113). For some parameter sets, counter reversals are more common than standard reversals, but for other parameter sets the pattern is reversed.

Tests 3 and 4 in Table 4 examine whether the pattern of behaviour is the same under choice and probabilistic valuation. For the PV group, test 3 examines the hypothesis that the proportion of subjects

who value the \$-bet more highly is the same under choice and valuation, against the alternative that the proportion who value the \$-bet more highly is lower under valuation. Test 4 is a between -groups test of the same hypothesis, comparing the proportion of subjects who value the \$-bet more highly in the PV group, with the proportion who choose the \$-bet in the MV group. With both tests, the null hypothesis is rejected at the 1% level for two parameter sets, but for two other parameter sets the test results have a positive sign. Overall, therefore, there is no clear pattern in the difference between choice and probabilistic valuation behaviour.

Our data is inconsistent with the implications of traditional economic theory. Can it be explained by the three psychological hypotheses set out above? To answer this question, we begin by considering whether the broad patterns observed in the data can be organized by each of the individual hypotheses.

Recall that the existence of standard reversal in the MV-group is inconsistent with the task goal hypothesis, conditional on the assumption that the perceived goal of valuation tasks is differentiation between gambles. While this is in fact the role of valuation tasks in our design, it is of course possible that subjects did not perceive it as such. Comparison between groups, however, affords us a further test of the task goal hypothesis. On the much milder assumption that perceived task-goals are independent of the valuation scale, the task goal hypothesis predicts no significant difference between groups with respect to the pairwise preferences implied by valuations. Test 5 in Table 4 tests the null hypothesis that the proportion who value \$ more highly in the MV-group is equal to the proportion who value \$ more highly in the PV-group. In a one-tailed test, against the alternative that the proportion who

value \$ more highly is greater in the MV-group, the null is decisively rejected in all cases. The prominence hypothesis also predicts that these proportions should be the same, hence prominence *acting alone* can also be rejected as an explanation for the results: prominence is consistent with the standard reversal in the MV-group but it cannot explain any systematic between group effects.

Among the hypotheses that we have considered, the only theory which predicts a systematic difference between our two experimental groups is the scale compatibility hypothesis. Table 5 summarises tests of the hypothesis that subjects' relative valuations in the PV-group are drawn from the same population as those for the MV-group. In all cases, except for the dominance parameter set, the null hypothesis of no significant difference is rejected at the 1% level. Are these observed differences consistent with compatibility?

TABLE 5 ABOUT HERE

Table 3 revealed some very obvious differences between the patterns of response for the two groups. In the PV-group, relative to the MV-group, standard reversals occur less frequently, counter reversals occur much more frequently. The two forms of reversals occur with approximately equal frequency for the PV-group (aggregating across parameter sets, the relative frequency of the standard reversal is 46%). In some cases, tests in Table 4 revealed a statistically significant difference between choice and probabilistic valuation, but overall there is no clear pattern. A natural interpretation of the compatibility hypothesis would be to assume: money-prize is more compatible with monetary valuation than with choice; win-probability is more compatible with probabilistic valuation than with choice. Given this interpretation we should not observe counter reversals in the MV-group and we should not observe

standard reversals in the PV-group. The high frequency of standard reversals in the PV-group, therefore, cannot be explained by this simple version of compatibility acting alone.

IV.ii Some Possible Explanations for Our Findings

The foregoing discussion has demonstrated that none of the economic or psychological hypotheses considered in isolation can explain our data. The existence of standard reversal in the MV-group is prime facie evidence against traditional economic theories, as are the differences observed between groups. The between group differences are also inconsistent with the prominence hypothesis and the task goal hypotheses. Finally, although the difference between groups is in the direction predicted by scale compatibility, that theory does not explain the high frequency of standard reversals among PV-group subjects. Nevertheless, the patterns in our data are broadly consistent with the findings of related research including MacCrimmon and Smith (1986) and Slovic et al (1990). In the light of this, in the rest of this section, we discuss three potential accounts which strike us as theoretical possibilities for organizing these patterns. Each account extends some existing theory and, in each case, we offer some opinion as to the merits of the option, though the reader should keep in mind that, since our experiment was not specifically designed to test these conjectures, our discussion here is much more speculative.

IV.ii.a Conjunctions of psychological effects

Although our data cannot be organized by any of the hypotheses we have presented, considered in isolation, it could be explained by some conjunction of hypotheses. As we see it, there are two particularly striking features of our data in need of explanation: one is the occurrence of standard

reversal in *both* groups; the second is the difference between groups. We can characterise the latter feature as a tendency for the frequency of standard reversal to fall, and counter reversal to rise, moving from the MV-group to the PV-group. This tendency is in the direction predicted by scale compatibility and since that theory is, to our knowledge, the only hypothesis consistent with that observation, a natural question to ask is whether there is some additional hypothesis which, when combined with scale compatibility, would explain those features of the data, otherwise inconsistent with it? The feature of our data in need of additional explanation is the occurrence of standard reversal in the PV-group. The obvious contender for explaining this is the prominence hypothesis. Referring back to the summary table at the end of Section III, recall that scale compatibility predicts the occurrence of standard reversals in the MV-group and counter reversals in the PV-group while prominence predicts standard reversals in both groups. A theory which proposed the simultaneous operation of both a scale compatibility and a prominence effect would predict a relatively high frequency of standard reversal in the MV-group (where both effects operate in the same direction); and, while standard reversals should be less frequent in the PV-group (assuming a non-negligible compatibility effect), they may still occur as a result of the operation of prominence. While there is ongoing dispute concerning the relative importance of scale compatibility effects and prominence effects¹⁷, a growing number of psychologists are accepting that both effects are significant factors in decision processes. Some researchers – we believe Slovic *et al* (1990) were the first - have explicitly suggested the combination of scale compatibility and prominence effects as explanations of preference reversal. Our data are broadly consistent with that conclusion.

On this psychological account of preference reversal, subjects' revealed preferences violate procedure invariance: that is, they vary with the procedure used to elicit them. As explained above, economists

have in the past presented theoretical reasons for resisting this conclusion by pointing to possible alternative explanations consistent with traditional economic theory. Our experiment was explicitly designed to control for the effects postulated in traditional economic explanations. Consequently, our results renew the challenge which preference reversal presents to traditional economic theorists. It thus seems natural to ask, on the supposition that our results are robust, what options remain for explaining the data consistent with context-free preferences? We can think of two possibilities and these are discussed in the next two sub-sections.

IV.ii.b Stochastic preference models

Of late, there has been considerable interest among economists in stochastic theories of decision making (see for example, Loomes and Sugden 1995). One surprising product of this research has been to demonstrate that patterns of behaviour which violate standard theories of preference, like expected utility theory, might nevertheless be consistent with stochastic versions of those preference models. Might some stochastic version of a traditional economic theory explain our data?

Any explanation of our findings must account for the systematic difference between groups. In principle, we can imagine how a stochastic version of a traditional economic theory might achieve this, even though the conditions faced by the two groups were formally equivalent from the point of view of any (non-stochastic) traditional economic theory. For example, suppose that subjects have true preferences which can be represented by some traditional economic theory, but they may make errors in tasks which lead them to mis-record their preferences. This could generate systematic differences between groups if subjects in one group are more prone to mistakes than the other. For example, this could

happen if, notwithstanding the formal equivalence of the tasks faced by the two groups, their perceived difficulty differs between groups, leading to differences in the frequency and/or pattern of errors relative to underlying preferences.

The extent to which subjects choose (or value) a dominated option over a dominating option arguably provides some indicative measure of the relative propensity for error across groups. So consider responses to parameter set 6 tasks. Notice that very few subjects select the dominated option in choice tasks (4.4% in the MV-group, 1.7% in the PV-group) while dominance is violated more frequently in valuation tasks (18% in the MV-group; 13 % in the PV-group). To the extent that violations of dominance reveal ‘error’, we conclude that errors are more likely in valuation, relative to choice, but there is very little difference between the rates of error in the MV- and PV-groups. Moreover, the differences that are observed look to be in the ‘wrong’ direction: in the aggregate, most reversals (standard+counter) are observed in the PV-group. It seems plausible to think that if the between group difference in reversing is to be explained by errors relative to a traditional economic theory, error rates will need to be higher in the PV-group. That is the opposite of what we observe. So while we are generally impressed by the advances being made in modelling stochastic preferences, we conjecture that this is not a promising line of attack for explaining differences between our two experimental groups.

IV.ii.c Regret theory and bounded rationality

In Section III, we argued that no traditional economic theory can explain systematic preference reversal in any parameter set of our experiment because our design uses the ordinal payoff scheme. We defined traditional economic theories as the class of models in which preferences are describable by a choice

function defined on the feasible set of gambles. We showed that, in an experiment which uses the ordinal payoff scheme, an individual who commits a preference reversal selects the same lottery regardless of the direction of the reversal. Recall, for example, the analysis of the ordinal payoff scheme with a single parameter set (i.e. $n=1$): the individual who records inconsistent rankings between choice and valuation is ultimately selecting the lottery $L_1 = \{P, 0.5; \$, 0.5\}$ regardless of whether their reversal is standard or counter; consequently, no traditional economic theory can explain a systematic pattern of reversal. The conclusion holds, for this case, so long as we interpret traditional economic theories as models to be applied to the feasible set of gambles induced by *both* choice and valuation tasks combined. For $n>1$, the conclusion holds so long as we assume that agents behave as if they consider the choice and valuation tasks in the experiment as a single choice among compound gambles. While this is formally the problem that subjects actually confront, thinking of the whole experiment as a single choice between compound gambles may be too taxing for a real subject, in terms of the computation required, its demands on memory, and so on. Our analysis of traditional economic theories in Section III did not recognize any such constraints and so, implicitly, invoked the assumption of *unbounded rationality*. As such, our analysis is in the spirit of traditional economic theories, but it is not the only possible approach.

An alternative would be to assume bounded rationality. While we offer no formal model of bounded rationality here, we can think of one obvious account of how a boundedly rational agent might approach the tasks in an experiment like ours. Suppose that subjects consider each task in isolation and do their best to give honest and considered responses. Experimental economists often interpret the data from experiments using the random lottery incentive system by making some such background assumption¹⁸.

We will call agents who behave in this way *well-meaning and myopic agents* (WAM for short). The analysis of Section III has established that no traditional economic theory can explain systematic preference reversal in our design, given unbounded rationality, but does a similar conclusion hold if we assume WAM agents?

For a large subset of traditional economic theories the answer is yes. In most traditional economic theories, preferences over lotteries can be represented by a monotonic value function $Q(\cdot)$ which assigns a unique value to each prospect (see Starmer, 2000). Observing preference reversal in our design cannot be explained by any such theory, if agents are WAM. To see why, consider the monetary valuation part of our design. If g_1 and g_2 are any two basic gambles, to observe a preference reversal is to observe $g_1 \succsim_c g_2$ and $V_m(g_1) < V_m(g_2)$. We can infer $Q(g_1) > Q(g_2)$ from the subject's choice and, since WAM agents reveal true valuations, $Q(V_m(g_i)) = Q(g_i)$ ($i = 1, 2$). It follows that $Q(V_m(g_1)) > Q(V_m(g_2))$, which violates monotonicity. A similar argument applies for probabilistic valuation.

There is, however, one traditional economic theory which might explain preference reversal in our design, given WAM agents, and that is regret theory. In the Appendix, we demonstrate that, given WAM agents, regret theory would be consistent with (though would not necessarily predict) the occurrence of standard reversals in the MV-group and the occurrence of counter reversals in the PV-group; it would not allow counter reversals in the MV-group, or standard reversals in the PV-group.

On this interpretation of it, regret theory provides one possible explanation for the asymmetry between standard and counter preference reversal observed among MV-group subjects; it also provides a reason to expect *some* difference between the behaviour of the two groups, though it cannot explain

why standard reversals continue to be observed in the PV-group. Thus, even granted the additional assumption of boundedly rational agents, regret theory only provides a partial fit to our data.

We take this observation to be in line with the results of other empirical studies of regret theory (e.g. Loomes, Starmer and Sugden (1991) and Starmer and Sugden (1998)). These studies have found evidence of novel patterns in choice behaviour predicted by regret theory, yet inconsistent with other traditional economic theories. On the other hand, these experiments also reveal systematic failures of regret theory. We are inclined to agree with the authors of these earlier studies who have concluded that regret might be a real influence on decision under risk but that, if it is, current formalizations have yet to model its operation accurately.

Section V: Conclusion

Taken at face value, the phenomenon of preference reversal is inconsistent with models of behaviour that are based on the assumption of context-free preferences, which is central to most of economics.

Economists provided good theoretical reasons for scepticism about this face value interpretation. For example, some theorists (among them Holt, Segal and Karni and Safra) pointed out that the typical experiments reporting preference reversal do not embody satisfactory controls: in the designs which these authors criticised, preference reversal could arise if agents have context-free, but non-expected utility, preferences. However, Tversky, Slovic and Kahneman (1990) casts some doubt on the practical significance of most of these theoretical defences of the context-free preference assumption.

In this paper, we have reported a new experiment whose design embodies improved controls, a new

valuation task and enhanced incentives. It extends the investigation of whether the economists' standard assumption of context-free preferences can be defended against the alternative view suggested by psychological theories that preferences are context-sensitive. We find strong evidence of systematic preference reversals which are inconsistent with a broad range of economic explanations of previous findings. However, none of the three leading psychological theories which we consider, when taken in isolation, can fully account for the patterns in our data either.

Given these findings, we have reviewed the explanatory strategies which remain open. The behaviour of our subjects would be consistent with a theory which allowed two or more psychological effects to operate simultaneously, as long as those effects include one – for example, a scale-compatibility effect – which can explain the difference between groups. We have also suggested two possible accounts that retain some of the spirit of traditional economic theories, though we have strong reservations about their validity. Although we do not discount these theories, it seems clear that our results renew the challenge to those who seek explanations of preference reversal based on models of context-free preferences.

References

- Cubitt, R.P., Starmer, C.V. and Sugden, R. (1998), 'On the validity of the random lottery incentive system', *Experimental Economics*, 1, 115-131.
- Cubitt, R.P., Starmer, C.V. and Sugden, R. (2000), 'Discovered preferences and the independence axiom: testing for common consequence and common ratio effects using single task designs', *Journal of Economic Methodology*, forthcoming.
- Delquié, P. (1993), 'Inconsistent trade-offs between attributes: new evidence in preference assessment biases', *Management Science*, 39, 1382-95.
- Delquié, P. (1997), "'Bi-matching": a new preference assessment method to reduce compatibility effects', *Management Science*, 43, 640-58.
- Fischer, G.W. and Hawkins, S.A. (1993), 'Strategy compatibility, scale compatibility, and the prominence effect', *Journal of Experimental Psychology: Human Perception and Performance*, 19, 580-597.
- Fischer, G.W., Carmon, Z., Ariely, D. and Zauberan, G. (1999), 'Goal-based construction of preferences: task goals and the prominence effect', *Management Science*, 45, 1057-1075.
- Grether, D. and Plott, C. (1979), 'Economic theory of choice and the preference reversal phenomenon', *American Economic Review*, 69, 623-638.
- Hausman, D. (1992), *The Inexact and Separate Science of Economics*, Cambridge University Press.
- Hershey, J.C. and P.J. Schoemaker (1985), 'Probability versus certainty equivalence methods in utility measurement: are they equivalent?', *Management Science*, 31, 1213-31.
- Holt, C.A. (1986), 'Preference reversals and the independence axiom', *American Economic Review*, 76, 508-515.
- Karni, E., and Safra, Z. (1987), "'Preference reversal" and the observability of preferences by experimental methods', *Econometrica*, 55, 675-685.
- Lichtenstein, S. and Slovic, P. (1971), 'Reversals of preference between bids and choices in gambling decisions'. *Journal of Experimental Psychology*, No. 89, 46-55.
- Lindman, H. (1971), 'Inconsistent preferences among gambles', *Journal of Experimental Psychology*, 80, 590-97.
- Loomes, G., Starmer, C. and Sugden, R. (1989), 'Preference reversal: information processing effect or rational non-transitive choice?', *Economic Journal*, 99, 140-151.

- Loomes, G., Starmer, C.V. and Sugden, R. (1991), 'Observing violations of transitivity by experimental methods', *Econometrica*, 59, 425-440.
- Loomes, G. and Sugden, R. (1983), 'A rationale for preference reversal', *American Economic Review*, vol. 73, 428-432.
- Loomes, G. and Sugden, R. (1987), 'Some implications of a more general form of regret theory', *Journal of Economic Theory*, 41, 270-87.
- Loomes, G. and Sugden, R. (1995), 'Incorporating a stochastic element into decision theories', *European Economic Review*, 39, 641-48.
- MacCrimmon, K. R. and M. R. Smith (1986), 'Imprecise equivalences: preference reversals in money and probability', University of Columbia working paper #1211, November.
- Segal U. (1988), 'Does the preference reversal phenomenon necessarily contradict the independence axiom?', *American Economic Review*, 78, 233-235.
- Slovic, P. (1975), 'Choice between equally valued alternatives', *Journal of Experimental Psychology: Human Perception and Performance*, 1, 280-287.
- Slovic, P. (1995), 'The construction of preferences', *American Psychologist*, 50, 364-371.
- Slovic, P., Griffin, D., and Tversky, A. (1990), 'Compatibility effects in judgment and choice', in Hogarth, R., (ed.) *Insights in Decision-Making*, Chicago, University of Chicago Press, pp 5-27.
- Starmer, C.V. (2000), 'Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk', *Journal of Economic Literature*, XXXVIII, 332-382.
- Starmer, C.V. and R. Sugden (1998), 'Testing alternative explanations of cyclical choices', *Economica*, 65, 347-61, 1998.
- Tversky, A., Slovic, P and Kahneman, D. (1990), 'The causes of preference reversal', *American Economic Review*, 80, 204-217.
- Tversky, A., Sattath, S. and Slovic, P. (1988), 'Contingent weighting in judgement and choice', *Psychological Review*, vol. 95, 371-384.
- Tversky, A and R. H. Thaler (1990), 'Preference reversals', *Journal of Economic Perspectives*, 4, 201-11.

Figure 1: The Display for a Choice Task

Option A: You get £32.00 for numbers 1 to 31

Option B: You get £ 8.00 for numbers 1 to 97

Choose A or B:

Figure 2: The Display for a Money Valuation Task

Set the missing amount so that the yardstick
and the option are equally attractive to you

yardstick: You get £ ??? for numbers 1 to 100

option: You get £18.00 for numbers 1 to 19

Pounds:

Pence:

Figure 3: The Display for a Probability Valuation Task

Set the missing number of discs so that the yardstick
and the option are equally attractive to you

yardstick: You get £10.00 for numbers 1 to ???

option: You get £18.00 for numbers 1 to 19

Enter a number in the range 1 to 100:

Table 1: Parameters

\$ - bet			P – bet		
Set	Prize (£)	Winning numbers (out of 100)	Prize	Winning numbers (out of 100)	
1	32	1 – 31	8	1 – 97	
2	18	1-19	4	1-81	
3	13	1-50	6	1-94	
4	17	1-39	5	1-94	
5	10	1-50	4	1-92	
Dominated Option			Dominating Option		
6	7	1-36	7	1-41	

Table 2: Summary of Predictions of Economic and Psychological Hypotheses				
	<i>STANDARD</i> Pref. Rev. Allowed?		<i>COUNTER</i> Pref. Rev. Allowed?	
<i>Maintained Hypothesis</i> ↓	Monet. Valuation (MV)	Prob. Valuation (PV)	Monet. Valuation (MV)	Prob. Valuation (PV)
Traditional Economic Theory	No	No	No	No
Prominence ¹	Yes	Yes	No	No
Scale Compatibility ²	Yes	No	No	Yes
Task Goal ³	No	No	No	No

Notes

¹ Conditional on the assumption that win-probability is the prominent attribute.

² Conditional on the assumption that money-prize is more compatible with monetary valuation than with choice and that win-probability is more compatible with probabilistic valuation than with choice.

³ Conditional on the assumption that the perceived goal of valuation tasks in an ordinal payoff scheme is differentiation.

Table 3: Outline Results

CHOICE	Set 1		Set 2		Set 3		Set 4		Set 5		Set 6 P dominates	
	MV	PV	MV	PV	MV	PV	MV	PV	MV	PV	MV	PV
% choosing P	51.75	51.72	57.89	54.31	40.35	39.66	31.58	42.24	24.56	25.00	95.61	98.28
% choosing \$	48.25	48.28	42.11	45.69	59.65	60.34	68.42	57.76	75.44	75.00	4.39	1.72
Sample Size	114	116	114	116	114	116	114	116	114	116	114	116
CHOICE & VALUATION												
Consistent, prefer P	9	29	16	44	11	27	4	18	3	16	92	100
Consistent, prefer \$	53	34	43	29	62	37	74	47	84	60	4	1
Standard Reversal (P chosen, \$ valued more highly)	50	31	50	19	35	19	32	31	25	13	17	14
Counter Reversal (\$ chosen, P valued more highly)	2	22	5	24	6	33	4	20	2	27	1	1

Note

P6 dominates \$6 $V(\$6) > V(P6)$ for 21 out of 114 subjects (18.4%), $V(\$6) > V(P6)$ for 15 out of 116 subjects (12.9%)

Table 4. Hypothesis Tests

	Set 1		Set 2		Set 3		Set 4		Set 5	
Choice Tasks	MV	PV	MV	PV	MV	PV	MV	PV	MV	PV
P chosen (No. of subjects)	59	60	66	63	46	46	36	49	28	29
\$ chosen (No. of subjects)	55	56	48	53	68	70	78	67	86	87
Sample Size	114	116	114	116	114	116	114	116	114	116
Valuation Tasks										
\$ Valued more highly	102	60	89	41	94	52	107	76	109	73
Test 1 t-stat.	8.20**		6.26**		4.32**		5.44**		4.31**	
Test 2 t-stat.	7.52**		5.34**		3.81**		7.04**		4.08**	
Test 3 t-stat	0.52		-1.55		-2.58**		0.52		-2.93**	
Test 4 t-stat.	0.53		-1.05		-2.26**		1.01		-3.84**	
Test 5 t-stat.	6.89**		7.22**		6.42**		6.29**		8.14**	

Notes

** = significant at 1% level (one-tailed test).

- Test 1. H0: Percentage who value \$ more highly, MV-group = Percentage who choose \$, MV-group
H1: Percentage who value \$ more highly, MV-group > Percentage who choose \$, MV-group
- Test 2: H0: Percentage who value \$ more highly, MV-group = Percentage who choose \$, PV-group
H1: Percentage who value \$ more highly, MV-group > Percentage who choose \$, PV-group
- Test 3: H0: Percentage who value \$ more highly, PV-group = Percentage who choose \$, PV-group
H1: Percentage who value \$ more highly, PV-group < Percentage who choose \$, PV-group
- Test 4: H0: Percentage who value \$ more highly, PV-group = Percentage who choose \$, MV-group
H1: Percentage who value \$ more highly, PV-group < Percentage who choose \$, MV-group

Test 5: H0: Percentage who value \$ more highly, MV-group = Percentage who value \$ more highly, PV-group
H1: Percentage who value \$ more highly, MV-group > Percentage who value \$ more highly, PV-group

Table 5. Do the PV-group relative valuations come from the same population as the MV-group relative valuations?

Parameter set	1	2	3	4	5	6
χ^2_3	258.5**	145**	162.2**	122.9**	381.5**	3.48

** = Significant at 1% level

Appendix: implications of regret theory with boundedly rational (WAM) agents.

In the version of regret theory presented by Loomes and Sugden (1987), preferences are defined over pairs of acts (where an act is a mapping from states of the world to consequences). Let A_i and A_j be two acts which result, respectively, in monetary consequences x_{is} and x_{js} in state of the world s . Preferences between such acts are determined by the expression:

$$A_i \succ A_j \Leftrightarrow \sum_s p_s \Psi(x_{is}, x_{js}) > 0 \quad (1)$$

where p_s is the probability of state s , and $R(x_{is}, x_{js})$ is a measure of the "net advantage" of having consequence x_{is} and missing out on x_{js} . The function $R(., .)$ is skew symmetric by construction, so that, for all consequences x and y , $R(x, y) = -R(y, x)$ and $R(x, x) = 0$. Loomes and Sugden assume that agents are regret averse by imposing the restriction $R(x, z) > R(x, y) + R(y, z)$, for all $x > y > z$.

To examine the implications of regret theory in our design (given WAM agents), consider Table A1 which is an act/event matrix describing three acts \$, P and M defined over five states with probabilities q through u ; x, y and c are money consequences with $x > y > 0, c > 0$.

Table A1	q	r	s	t	u
\$	x	0	0	0	0
P	y	y	y	0	0
M	c	c	0	c	0

We can represent all of the tasks faced by subjects in our groups MV and PV as special cases of Table A1. To characterise the tasks facing the MV-group, in a given parameter set, set $s = u = 0, x > y > c > 0$ and take q, r and t to be strictly positive constants. The tasks are: choose between \$ and P; select $c = c_s$ such that $M \sim \$$; and select $c = c_p$ such that $M \sim P$. If we are to observe a standard reversal, we need (a) $P \succ M \succ \$$ and (b) $c_s > c_p$. Given (b), there will be some c^* , with $c_s > c^* > c_p$, such that $M \succ P$ and $\$ \succ M$. To characterise the tasks faced by the PV-group, in a given parameter set, set $t = 0, x = c = T > y > 0$ (where T is the money prize of the yardstick) and take q and $(r + s)$ to be strictly

positive constants. In this group, the tasks are: choose between \$ and P; select $r = r_s$ such that $M \sim \$$; and select $r = r_p$ such that $M \sim P$. Probability valuation is effectively setting the relative sizes of r and s , but we can think of it as setting just r (or s) since the value of $r + s$ is fixed, for a given parameter set. If we are to observe a standard reversal in this group, we need $P \succ M \succ \$$ and that there exists some r^* , with $r_s > r^* > r_p$, such that $M \succ P$ and $\$ \succ M$. Hence, we can examine whether regret theory explains preference reversal in our design by examining whether the theory permits cycles in pairwise choice among \$, P and M for some c^* (MV-group) or some r^* (PV-group), given the appropriate restrictions, in each case, on the parameters of Table A1.

Regret theory implies that we may observe such cyclical choices. Equations (2)-(4) below result from the application of regret theory to pairwise comparisons of the acts in Table A1.

$$P \succ M \text{ (–) } \$ \rightarrow q\psi(y,x) + r\psi(y,0) + s\psi(y,0) > (<) 0 \quad (2)$$

$$\$ \succ M \text{ (–) } P \rightarrow q\psi(x,c) + r\psi(0,c) + t\psi(0,c) > (<) 0 \quad (3)$$

$$M \succ P \text{ (–) } \$ \rightarrow q\psi(c,y) + r\psi(c,y) + s\psi(0,y) + t\psi(c,0) > (<) 0 \quad (4)$$

If each of (2)-(4) is positive (resp. negative), regret theory predicts a cycle of choices consistent with the standard (resp. counter) reversal. Summing across the left hand side of these three inequalities generates expression (5):

$$q.[\psi(y,x) + \psi(x,c) + \psi(c,y)] + r.\{\psi(y,0) + \psi(0,c) + \psi(c,y)\} \quad (5)$$

It is a necessary condition for observing the cycle consistent with standard (resp. counter) reversal that (5) is positive (resp. negative). In the MV-group, $x > y > c > 0$. Rearranging (5) gives $q.[\psi(x,c) - \psi(x,y) - \psi(y,c)] + r.\{\psi(y,0) - \psi(c,0) - \psi(y,c)\}$. Since both bracketted terms are positive (by regret aversion), the cycle implied by standard reversal is consistent with (though not necessarily predicted by) regret theory. Notice, however, that the opposite cycle (implied by counter reversal) would not be allowed by regret theory.

In the PV-group, $x = c > y > 0$. Rearranging (5) gives $q.[\psi(x,c) + \psi(c,y) - \psi(x,y)] + r.\{\psi(c,y) +$

$\psi(y,0) - \psi(c,0)$. In this case, regret aversion implies that both bracketed terms are negative (except when $x = c$, as in parameter set 5, then $[\dots] = 0$, and $\{\dots\} < 0$). Hence, the cycle of choices implied by counter reversal is permitted, while that in the opposite direction (standard reversal) is not.

Notes

- ¹ An important exception is the experiment of Tversky, Slovic and Kahneman (1990), who interpret their results as telling in favour of the psychologists' approach. We comment further on their paper below.
- ² Reviews of this evidence are contained in Tversky and Thaler (1990) and Hausman (1992, ch. 13).
- ³ In an experiment that uses the BDM device to elicit a selling price for some item, the subject is first given an entitlement to that item and then asked to state the minimum price at which she would be prepared to sell it back to the experimenter. After this, a random device is used to generate a price. If, and only if, this price exceeds the price stated by the subject, the subject is required to sell the item to the experimenter at the randomly generated price.
- ⁴ In an experiment that uses the random lottery incentive system, each subject performs a number of tasks, knowing from the start that at the end of the experiment exactly one task will be selected at random to be for real, i.e. that it will be played out for real using the response that the subject has given. For discussion, see Cubitt, Starmer and Sugden (1998, 2000).
- ⁵ The formulation of the hypothesis by Tversky, Sattath and Slovic (1988) built upon earlier work by Slovic (1975). For more recent discussions, see Fischer and Hawkins (1993), Slovic (1995) and Fischer *et al* (1999).
- ⁶ The hypothesis does not necessarily require that this attribute is initially recognised by the subject as the most important. It is consistent with the claim that a subject, faced with a choice between items, focusses primarily on one attribute as part of a process of preference construction.
- ⁷ An example of such a reasoning strategy is the dom-DA-lex strategy described by Fischer and Hawkins (1993).
- ⁸ See also Slovic, Griffin and Tversky (1990) and Fischer *et al* (1999).
- ⁹ See also Tversky, Slovic and Kahneman (1990), Slovic, Griffin and Tversky (1990) and Slovic (1995).
- ¹⁰ We take it that this is part of the definition of the ordinal payoff scheme. It is part of the definition presented by TSK in their theoretical discussion of elicitation procedures. However, we note that in their actual experiment, one group were *asked to imagine* that they would play a gamble, selected as in the ordinal payoff scheme. The other group were told that 15% of subjects really would play such a gamble. We comment on this further below.
- ¹¹ In demonstrating this proposition, we will assume that the choice function implied by any traditional economic theory is to be applied to the feasible set of compound gambles determined jointly by the possible responses to all tasks in the experiment. In Section VI.iii we examine the implications of relaxing this assumption.

¹² Betweenness is a weakening of the independence axiom of expected utility theory, satisfied by some but not all generalisations of that theory (see Starmer (2000) for discussion). In the present context it implies that, if P is strictly preferred to \$, then any compound gamble defined only on P and \$, and giving strictly positive probability to each of them, is strictly preferred to \$ and strictly dispreferred to P.

¹³ Fischer *et al* (1999) discuss other designs in which matching tasks are used for differentiation purposes and argue, like us, that the task goal hypothesis implies that, to the extent that the goal of differentiation is transparent to the subject, behaviour in these tasks will accord with choice behaviour.

¹⁴ Related work is reported in Hershey and Schoemaker (1985), MacCrimmon and Smith (1986), Tversky, Sattath and Slovic (1988), Slovic, Griffin and Tversky (1990), Fischer and Hawkins (1993), and Delquié (1993, 1997). We contrast our findings with some of these results later in the paper.

¹⁵ The yardstick and all of the gambles to be valued are basic. Thus, the task of putting a monetary valuation on $g_i = (m_i, p_i)$ is a matching task of the form: select a money-prize m such that the basic gambles (m_i, p_i) and $(m, 1)$ are equally attractive. Similarly, the task of putting a probabilistic valuation on $g_i = (m_i, p_i)$ is a matching task of the form: select a win-probability p such that the basic gambles (m_i, p_i) and (T, p) are equally attractive.

¹⁶ The expected value of gambles in our experiment ranged from approx. UK£2.50 to UK£10 (with £1 \cong \$1.5 at the time of writing). The average expected value was £5.20.

¹⁷ For example, Slovic, Griffin and Tversky (1990) interpret their findings as indicating the primary importance of scale compatibility, whilst accepting that prominence plays a role. On the other hand, Fischer and Hawkins (1993) argue for the primary importance of prominence, according the minor role to scale compatibility. Fischer *et al* (1999) provide evidence in support of the task goal hypothesis over its rivals. See, also Delquié(1993, 1997).

¹⁸ For an empirical investigation of the reliability of this assumption see Cubitt, Starmer and Sugden (1998).