

Bateman, Ian; Kahneman, Daniel; Munro, Alistair; Starmer, Chris; Sugden, Robert

**Working Paper**

## Is there loss aversion in buying? An adversarial collaboration

CSERGE Working Paper EDM, No. 03-07

**Provided in Cooperation with:**

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

*Suggested Citation:* Bateman, Ian; Kahneman, Daniel; Munro, Alistair; Starmer, Chris; Sugden, Robert (2003) : Is there loss aversion in buying? An adversarial collaboration, CSERGE Working Paper EDM, No. 03-07, University of East Anglia, The Centre for Social and Economic Research on the Global Environment (CSERGE), Norwich

This Version is available at:

<http://hdl.handle.net/10419/80267>

**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.*

**IS THERE LOSS AVERSION IN BUYING?  
AN ADVERSARIAL COLLABORATION**

by

**Ian J. Bateman<sup>1</sup>, Daniel Kahneman<sup>2</sup>  
Alistair Munro<sup>3</sup>, Chris Starmer<sup>4</sup>  
and Robert Sugden<sup>3</sup>**

<sup>1</sup>**The Centre for Social and Economic Research  
on the Global Environment  
and  
School of Environmental Sciences,  
University of East Anglia,  
Norwich NR4 7TJ, UK.**

<sup>2</sup>**Woodrow Wilson School for Public and International Affairs,  
Princeton University.**

<sup>3</sup>**School of Economic and Social Studies,  
University of East Anglia.**

<sup>4</sup>**School of Economics, University of Nottingham**

**Contacts:**

**email: [i.bateman@uea.ac.uk](mailto:i.bateman@uea.ac.uk); tel: (44) (0)1603 593125;  
e-mail: [r.sugden@uea.ac.uk](mailto:r.sugden@uea.ac.uk); tel: (44) (0)1603 593423**

**IS THERE LOSS AVERSION IN BUYING?  
AN ADVERSARIAL COLLABORATION**

by

**Ian J. Bateman, Daniel Kahneman  
Alistair Munro, Chris Starmer  
and Robert Sugden**

**CSERGE Working Paper EDM 03-07**

**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 Centre for Social and Economic Research on the Global Environment (CSERGE).

We are grateful to Colin Camerer, Jack Knetsch and Peter Wakker, and to participants at an Economic Science Association seminar on 'Perspectives on Prospect Theory' and at seminars at the University of Amsterdam, Brunel University, Cardiff Business School and the University of Warwick, for comments on earlier versions of this paper. This research was supported by the Economic and Social Research Council of the UK as part of its Risk and Human Behaviour Programme (award number L MG1 25 2053). Robert Sugden's work was also supported by the Leverhulme Trust.

**ISSN 0967-8875**

## Abstract

This paper reports an exercise in *adversarial collaboration*. An adversarial collaboration is an investigation carried out jointly by two individuals or research groups who, having proposed conflicting hypotheses, seek to resolve the issue in dispute. The experiment reported was designed to reconcile differences between the apparently conflicting results of two previous experiments, one carried out by Kahneman, the other by the other authors. Specifically, it investigates whether, when consumers consider giving up money in exchange for goods, they construe potential money outlays as losses. This issue bears on the explanation of the widely observed disparity between willingness-to-pay and willingness-to-accept valuations of costs and benefits, which has proved so problematic for contingent valuation studies. The results of the experiment do not decisively resolve the question in dispute, but they are broadly consistent with the hypothesis that money outlays *are* perceived as losses.

**Key words:** Adversarial collaboration, loss aversion, willingness to pay, reference-dependent preferences, status quo bias.

## 1. Introduction

This paper reports an exercise in a research methodology that we believe is new to experimental economics, and is unusual in experimental psychology: *adversarial collaboration*. An adversarial collaboration is an investigation carried out jointly by two individuals or research groups who, having proposed conflicting hypotheses, seek to resolve the issue in dispute. The work we describe arises from our attempts to reconcile an apparent difference between the results of two previous experiments, one of which was carried out by Kahneman in association with Jack Knetsch and Richard Thaler (Kahneman *et al.*, 1990) the other by Bateman, Munro, Starmer and Sugden – whom we shall call the ‘British group’ – in association with Bruce Rhodes (Bateman *et al.*, 1997). In each of these experiments, a null hypothesis derived from the received theory of consumer choice is rejected in favour of an alternative hypothesis that individuals’ preferences are conditional on ‘reference states’. However, the alternative hypotheses that are confirmed in the two experiments are not quite the same.

Through adversarial collaboration, we attempt to settle the issue that is left open by these conflicting experimental results. The question we address is this: when consumers consider giving up money in exchange for goods, do they construe potential money outlays as losses? This issue in the theory of reference-dependent preferences has practical significance, since it bears on the explanation of the widely observed disparity between willingness-to-pay and willingness-to-accept valuations of costs and benefits, which has proved so problematic for contingent valuation studies.

This paper can be read on two different levels. At the level of methodology, it pioneers a method of resolving scientific disputes that has wide applicability across experimental economics. At the substantive level, it reports one of the largest experimental investigations to date into the determinants of loss aversion. Although the experiments we report are designed to discriminate between two alternative formulations of the theory of reference-dependent preferences, the null hypotheses for our tests are given by standard consumer theory. Hence, our experiments also constitute tightly-controlled tests of whether, as standard theory predicts, preferences over consumption bundles are invariant with respect to changes in reference states.

In Section 2, we explain the principles of adversarial collaboration and discuss its value in general. In Section 3, we explain the particular issue that our adversarial collaboration was intended to resolve, and state our rival prior hypotheses. In the remainder of the paper, we describe our attempts to resolve this issue. Our experimental design works by eliciting a range of alternative

measures of individuals' valuations of changes in consumption. In Sections 4 and 5, we explain these measures, and show that the competing hypotheses generate different implications about disparities between them. In Section 6, we consider a potentially confounding factor, and how it can be neutralised. In Sections 7 and 8, we explain the experiment itself. In Sections 9 and 10 we report our results and consider how they should be interpreted. Section 11 concludes. Readers who expect a one-line summary of the results at this stage must be disappointed: our findings are not straightforward enough for that to be possible.

## 2. Adversarial Collaboration

In an adversarial collaboration, the two parties agree on the design of an experiment which they will conduct jointly. Before knowing what the experiment will find, they accept its validity as a test – not necessarily a conclusive test – of their respective hypotheses. Each party anticipates its interpretation of possible outcomes of the experiment, particularly those that it does *not* predict. The two parties agree that particular outcomes of the experiment would support one hypothesis, and particular other outcomes would support the other. Both parties commit to publishing the results, whatever they may be.<sup>1</sup> We believe that this methodology has some advantages over the more conventional form of scientific debate, in which each research group designs and runs experiments independently, chooses which of its results to publish, and can challenge the validity of other groups' experimental designs after knowing the results those designs have produced. Adversarial collaboration encourages a more constructive approach to the resolution of disagreements.

Adversarial collaboration, as compared with conventional scientific debate, requires different attitudes on the part of researchers – in particular, more attention to understanding the other side's arguments, and less to rhetorical strategies for defeating them. But it also requires different expectations on the part of the scientific community as a whole. We are all used to reading journal articles which report apparently clear-cut experimental results and which draw strong conclusions from them. But as readers, we learn to apply some discount to such claims. We have to allow for confirmation bias in the design of experiments – the tendency for researchers to look for 'tests' which seem likely to confirm their prior hypotheses. We also have to keep in mind that the experiments that are reported in the journals are not necessarily representative of the larger set of experiments that have been run: we have to allow for the possibility that research groups publish only their most 'successful' experiments, and use conformity with their prior hypotheses as one of their criteria of success. And we expect the authors of journal articles to talk up their conclusions, drawing wide-ranging implications from them and down-playing doubts and ambiguities. Adversarial collaboration must be expected to lead to a different kind of publication.

Because the experimental designs used in adversarial collaboration have to be agreed by both parties, each party has to subject its hypothesis to a genuinely

---

<sup>1</sup> A proposed protocol for adversarial collaboration is included in Mellers, Hertwig and Kahneman (2001). This protocol includes the rule that, after the initially-agreed experiment has been run, and before the principle that binds the parties to publication comes into play, each party may propose one additional experiment 'to exploit the fount of hindsight wisdom which commonly becomes available when disliked results are obtained'.

stringent test. The guarantee of stringency is this: each party's hypothesis is being subjected to a test that the other party expects it to fail. Thus, one of the mechanisms which tends to generate *apparently* decisive experimental results in the existing literature, positive confirmation bias, is neutralised. The commitment to publication, backed up by the two parties' common knowledge of the outcomes of the experiment, neutralises another such mechanism: selection bias at the publication stage.

Adversarial collaboration will not always bring the parties into full agreement about the issue in dispute: they may have different interpretations of what their jointly-conducted experiment has found. Scientific debate is better served if such differences are reported frankly than if they are concealed by bland generalities. From the reader's point of view, a report of this kind may be more useful than the superficially more definite conclusions that are customarily expected of non-adversarial research papers. Ultimately, however, the value of an adversarial collaboration is to be found in the validity and power of the experimental design it has adopted, and in the quality of the data this has generated. Whether, having seen the results, the parties to the collaboration agree on how they should be interpreted is a secondary matter: it is the reader who must draw the conclusions.

### 3. Theoretical Background

The hypothesis that there is an asymmetry between individuals' attitudes to gains and to losses was first brought to the attention of economists by Kahneman and Amos Tversky (1979) and by Richard Thaler (1980). Since then, there has been an accumulation of evidence – from experiments, from survey data, and from the field – which suggests that individuals' choices are more responsive to anticipated losses than to equal and opposite anticipated gains.<sup>2</sup> These findings are inconsistent with the standard (or *Hicksian*) theory of consumer choice, in which preferences over final consumption states are independent of individuals' current endowments.

Probably the most fully-developed theoretical explanation of this asymmetry is the theory of *reference-dependent preferences*, proposed by Tversky and Kahneman (1991). In this theory, individuals have preference orderings over bundles of goods, as in Hicksian consumer theory, but these preferences are defined relative to *reference states*. A reference state is a point in goods space which the individual treats as the status quo or normal expectation; gains and losses in the various dimensions of goods space are defined in terms of displacements from the reference state. In notation, reference states are represented by subscripting the preference relation; thus 'x is weakly preferred to y, viewed from the reference state r' is written as  $x \succsim_r y$ .

Tversky and Kahneman propose a hypothesis of *loss aversion* which links changes in (reference-dependent) preferences with changes in the reference state. Let  $x = (x_1, \dots, x_n)$  and  $y = (y_1, \dots, y_n)$  be two bundles of the same n goods, such that for some good i,  $y_i > x_i$ . Let  $r = (r_1, \dots, r_n)$  and  $s = (s_1, \dots, s_n)$  be potential reference states, such that  $r_i = x_i$ ,  $s_i = y_i$ , and  $r_j = s_j$  for all  $j \neq i$ . For any given i, there is *loss aversion in good i* if, for all such x, y, r, s:  $y \succsim_r x \Rightarrow y \succsim_s x$ . Tversky and Kahneman's hypothesis is that there is loss aversion, so defined, in all goods.<sup>3</sup>

<sup>2</sup> Bateman *et al.*, (1997), Sugden (1999) and Starmer (2000) give references to relevant experimental and survey research. See also Myagkov and Plott (1997), who find loss aversion in an experimental market setting, and Samuelson and Zeckhauser (1988) and Benartzi and Thaler (1995), who find evidence of loss aversion in 'real' markets.

<sup>3</sup> In fact, Tversky and Kahneman propose a strictly stronger hypothesis, namely that the implication  $y \succsim_r x \Rightarrow y \succsim_s x$  holds for all x, y, r, s such that  $y_i > x_i$ ,  $r_i = x_i$ ,  $y_i \geq s_i > x_i$ , and  $r_j = s_j$  for all  $j \neq i$ . However, the hypothesis we have stated is sufficient for the purposes of this paper, and the intuition behind it is easier to explain.

In order to apply Tversky and Kahneman's theory to a concrete choice problem, it is necessary to specify the chooser's reference state. The theory itself does not tell us how reference states are determined; in this respect, the formal model is left uninterpreted (Tversky and Kahneman, 1991, pp. 1046-1047). Thus, how the theory should be applied to specific decision-making environments can be a matter of judgement.

One possibility is to interpret an individual's reference state as whatever bundle of goods she currently owns. This *current endowment hypothesis* provides a simple method of deriving more specific testable hypotheses from reference-dependent preference theory. It was used by the British group to set up the hypotheses that were tested and confirmed in what we shall call the *Norwich experiment* (reported by Bateman *et al.*, 1997). In the light of the findings of this experiment, the British group has treated the current endowment hypothesis as descriptively adequate, at least in relation to simple laboratory tasks of buying, selling and choosing.<sup>4</sup> Munro and Sugden (2003) have used the hypothesis more generally in an analysis of how the workings of markets are affected by loss aversion.

In contrast, Tversky and Kahneman (1991) do not fully endorse the current endowment hypothesis, even with respect to simple laboratory tasks. Among the cases in which they think this hypothesis may be inappropriate is the design used in Kahneman, Knetsch and Thaler's (1990) *Vancouver experiment*. In this experiment, subjects bought and sold coffee mugs for money, indicating their willingness to trade by reporting their valuations of mugs as 'buyers', as 'sellers' and as 'choosers'. Noting that chooser's valuations were found to be much closer to buyer's valuations than to sellers', Tversky and Kahneman suggest that buyers do not value the money they give up in a transaction as a loss. Their hypothesis is that a subject who is considering buying a coffee mug construes her reference state as including *neither* the mug *nor* the money she would have to spend to buy it. If she chooses to buy, she gains the mug; if not, she gains whatever she would buy with the money instead. We shall call this hypothesis *no loss in buying* (NLIB).

Explaining the intuition behind NLIB, Tversky and Kahneman give two further examples of cases in which, they suggest, giving up an initial endowment will

---

<sup>4</sup> The British group do not claim that the current endowment hypothesis applies to *all* decision situations. For example, that hypothesis might not apply for regular purchases of consumption goods; in this case, customary expenditure and customary consumption might be integrated into the reference state. Or (a case which has some significance for contingent valuation) the hypothesis might not apply if decision-makers perceive current endowments as unfair or morally wrong.

not induce any pain of loss. The first is that of a person who gives up a \$5 bill and receives five \$1 bills in exchange. The second is that of a routine commercial transaction in which a trader sells goods from stock. The intuition behind these examples is that if a person gives up A to get B, there is a perception of loss if and only if A is valuable to her in some way that B is not. The \$5 bill serves no purpose that is not equally well served by the \$1 bills; to the trader, stocks of goods awaiting sale already have the status of tokens for money.

Kahneman has subsequently firmed up the NLIB hypothesis by proposing a theory of the conditions under which the gains and losses associated with a transaction are mentally *integrated* prior to evaluation, rather than being evaluated separately.<sup>5</sup> When a loss and an equal and opposite gain are integrated in this way, painful perceptions of loss do not arise. Kahneman proposes that consumers normally have *budget reserves*, that is, reserves of money that are available for unanticipated spending. When an individual faces an unanticipated opportunity to buy a good, and is able to finance this spending from her budget reserve, gains and losses are integrated: the money that has to be spent to buy the good is already seen as a token for unspecified goods. In such circumstances, money outlays are not perceived as losses. In contrast, if the individual faces an unanticipated buying opportunity which she can finance only by forgoing some *specific* consumption plan, the act of buying involves a definite loss, separable from the gain; and so the money payment is perceived as a loss. In the converse case of selling, gains and losses are integrated if the proceeds of the sale are earmarked for the purchase of a replacement good; but they are treated separately if those proceeds will be added to the budget reserve (Kahneman and Novemsky, 2002). Notice that NLIB does not assert the absence of loss aversion with respect to money: it asserts that loss aversion in money does not impact on certain kinds of buying transactions. Thus, NLIB is compatible with the hypothesis of prospect theory that, when *uncertain prospects* are evaluated, negative monetary outcomes are perceived as losses (Kahneman and Tversky, 1979).

The objective of our adversarial collaboration was to test the current endowment hypothesis against NLIB. Our experimental design elicits subjects' willingness to engage in transactions involving money and low-value, non-staple goods. If the idea of budget reserves is accepted, it seems reasonable to assume that subjects' purchases of such goods can be financed from budget

---

<sup>5</sup> The development of this theory was an early part of the process of adversarial collaboration. It illustrates one of the positive features of such collaboration: theoretical progress may be stimulated through the process of the parties attempting to understand one another's positions.

reserves, and that the proceeds from sales are not earmarked for replacements. Thus, we take it to be an implication of NLIB that, when experimental subjects confront opportunities to buy low-value non-staple goods, gains and losses are integrated.

Of course, NLIB and the current endowment hypothesis have conflicting implications in a much wider class of environments than those we use for our tests. For example, in contingent valuation surveys, respondents are often asked to state their willingness to pay for public goods. If the current endowment hypothesis is true, these (hypothetical) money outlays will be perceived as losses; if respondents are subject to loss aversion that will impact negatively on their willingness-to-pay valuations. In contrast, NLIB might be interpreted as implying that such valuations are not affected by loss aversion.

#### 4. Implications of Competing Hypotheses: the Unrestricted Model

Consider a model in which there are only two goods; quantities of these goods are represented by  $x_i, x_j$ . Usually, we will interpret this model so that one good is some particular private consumption good (for short, 'the good') and the other is an index of general purchasing power, or 'money'. When we use this interpretation, the good will be denoted by  $G$  and money by  $M$ .

For any given individual, consider how one unit of good  $i$  can be valued in units of good  $j$ . (Notice that there is no loss of generality in speaking of the value of 'one unit' of good  $i$ : as modellers, we are free to choose the units in which good  $i$  is measured so that any given real quantity counts as one unit.) More specifically, for any given  $x_i', x_j'$ , suppose that the individual is endowed with  $x_j'$  of good  $j$ , and consider how we might express in units of good  $j$  the value of consuming  $x_i' + 1$  units of good  $i$  rather than  $x_i'$  units. We define three measures of this value. These measures are standardly used in contingent valuation and cost-benefit studies.<sup>6</sup>

*Willingness to pay (WTP).* Suppose the individual's current endowment is  $(x_i', x_j')$ .  $WTP_{ji}$  is the largest amount of good  $j$  that the individual would be willing to give up in return for a gain of one unit of good  $i$ .

*Willingness to accept (WTA).* Suppose the individual's current endowment is  $(x_i' + 1, x_j')$ .  $WTA_{ji}$  is the smallest amount of good  $j$  that the individual would be willing to accept in return for accepting a loss of one unit of good  $i$ .

*Equivalent gain (EG).* Suppose the individual's current endowment is  $(x_i', x_j')$ .  $EG_{ji}$  is the smallest amount of good  $j$  that the individual would be willing to accept in place of a gain of one unit of good  $i$ .

These definitions do not presuppose that the individual has preferences (whether Hicksian or not) over bundles of the two goods. The measures they define are *not* constructs within a theory of preference (in the sense that the Hicksian concepts of compensating and equivalent variation are). They are observable magnitudes, revealed in the individual's behaviour in some given setting.

---

<sup>6</sup> These measures are discussed in more detail by Bateman *et al.*, (1997), who also define a further measure, *equivalent loss (EL)*. (For an individual whose current endowment is  $(x_i' + 1, x_j')$ ,  $EL_{ji}$  is the largest amount of good  $j$  that the individual would be willing to give up in place of a loss of one unit of good  $i$ .)

We now consider the implications of three alternative theories: Hicksian consumer theory, reference-dependent preference theory in conjunction with the current endowment hypothesis, and reference-dependent theory in conjunction with NLIB. We assume that income effects are weakly positive. It will be convenient to state the implications of theories in terms of predictions about *ratios* of valuations. The relevant predictions are summarised in Table 1.

**Table 1: Predictions of unrestricted theories**

ratio	value of ratio predicted by theory:		
	Hicksian preferences	reference-dependent preferences and current endowment hypothesis	reference-dependent preferences and NLIB
$WTA_{MG}/EG_{MG}$	1	$> 1$	$> 1$
$EG_{MG}/WTP_{MG}$	$\geq 1$	$> 1$	$\geq 1$
$WTA_{GM}/EG_{GM}$	1	$> 1$	1
$EG_{GM}/WTP_{GM}$	$\geq 1$	$> 1$	$> 1$

With respect to money valuations of the good, Hicksian theory predicts  $WTA_{MG}/EG_{MG} = 1$  and  $EG_{MG}/WTP_{MG} \geq 1$ . From the viewpoint of that theory,  $WTA_{MG}$  and  $EG_{MG}$  are *identical* to one another. (The only respect in which these measures differ is the location of the status quo point, and the concept of a status quo point is not recognised by Hicksian theory.)  $EG_{MG}$  and  $WTP_{MG}$  measure different entities within Hicksian theory; differences between their values reflect income and substitution effects. However, if, as in most experimental environments,  $EG_{MG}$  is small relative to the individual's total wealth, credible values of the rate of change of  $WTP_{MG}$  with respect to wealth imply that the value of  $EG_{MG}/WTP_{MG}$  is close to 1.<sup>7</sup> With respect to valuations of a

<sup>7</sup> For example, consider a student subject with money wealth of \$10,000 whose WTA for a coffee mug is \$10. If her WTP for a mug is proportional to her wealth, the Hicksian prediction is that, given her current wealth, her WTP for the mug is \$9.995. For more discussion of this point, see Sugden (1999).

given sum of money in units of the good, Hicksian theory has the symmetrical predictions  $WTA_{GM}/EG_{GM} = 1$  and  $EG_{GM}/WTP_{GM} \geq 1$ .

The theory of reference-dependent preferences, in conjunction with the current endowment hypothesis, implies  $WTA_{MG}/EG_{MG} > 1$ ,  $EG_{MG}/WTP_{MG} > 1$ ,  $WTA_{GM}/EG_{GM} > 1$ , and  $EG_{GM}/WTP_{GM} > 1$ . The prediction  $WTA_{MG}/EG_{MG} > 1$  is an implication of loss aversion in the good. Intuitively,  $WTA_{MG}$  is based on a comparison between gains of money and *losses* of the good, while  $EG_{MG}$  is based on a comparison between gains of money and *gains* of the good. If preferences are more responsive to losses than to gains, as the hypothesis of loss aversion implies,  $WTA_{MG}$  will be greater than  $EG_{MG}$ . Symmetrically,  $WTA_{GM}/EG_{GM} > 1$  is an implication of loss aversion in money. The other two inequalities result from the conjunction of loss aversion and income and substitution effects; loss aversion in money contributes to  $EG_{MG}/WTP_{MG} > 1$ , while loss aversion in the good contributes to  $EG_{GM}/WTP_{GM} > 1$ .<sup>8</sup>

If instead the theory of reference-dependent preference is combined with NLIB, outlays of money are treated as foregone gains rather than as losses. Thus, as in Hicksian theory,  $EG_{MG}/WTP_{MG} \geq 1$  and  $WTA_{GM}/EG_{GM} = 1$ . If income effects can be assumed to be small, the value of  $EG_{MG}/WTP_{MG}$  will be close to 1. For the other ratios, the implications are just as in the case of the current endowment hypothesis:  $WTA_{MG}/EG_{MG} > 1$  (the result of loss aversion in the good) and  $EG_{GM}/WTP_{GM} > 1$  (the result of loss aversion in the good combined with income and substitution effects).

These predictions can be tested in an experimental design which elicits WTA, EG and WTP valuations. However, such a design has a significant limitation: it cannot measure the extent of loss aversion in money, independently of assumptions about the truth or falsity of NLIB. This induces some lack of sharpness in the interpretation of results.

For example, suppose that an experiment finds that the value of  $WTA_{MG}/EG_{MG}$  is significantly greater than 1, while the value of  $WTA_{GM}/EG_{GM}$  is not. One possible interpretation is that NLIB is true: money is subject to loss aversion, but because money outlays are treated as foregone gains, this effect is not picked up in buying tasks. But there is a rival interpretation: the current endowment hypothesis might be true, but loss aversion in money might be much weaker than loss aversion in the good. Alternatively, suppose we find that both ratios are significantly greater than 1, but  $WTA_{GM}/EG_{GM}$  is close to 1 and much less than  $WTA_{MG}/EG_{MG}$ . This might be interpreted as evidence that the current

<sup>8</sup> The reasoning that leads to the predictions stated in this and the preceding paragraph is presented in more detail in Bateman *et al.*, (1997).



endowment hypothesis is true, and that loss aversion in money is weaker than loss aversion in the good. But again there is a rival interpretation: the underlying propensity to feel loss aversion might be similar for money and for the good, but *in buying tasks*, the full extent of loss aversion in money is not expressed. In other words, even though the NLIB hypothesis is rejected, the truth might still be *closer* to NLIB than to the current endowment hypothesis. In order to discriminate between such rival interpretations, we need an *independent* measure of loss aversion in money.

## 5. Implications of Competing Hypotheses: a Restricted Model

By eliciting valuations *under risk*, it is possible to infer the extent of loss aversion in money without assuming either the current endowment hypothesis or NLIB. The first step is to define two further measures of the value of one unit of good  $i$ . Using the same notation as in the definitions of WTP, WTA and EL, these measures are:

*Risky willingness to pay (RWTP)*. Suppose the individual's current endowment is  $(x_i', x_j')$ . Consider a gamble with two mutually exclusive outcomes, each with probability 0.5. One outcome is that the individual gains one unit of good  $i$ , with no change in good  $j$ . The other is that she loses some amount of good  $j$ , with no change in good  $i$ .  $RWTP_{ji}$  is the largest such loss of good  $j$  consistent with her being willing to accept the gamble.

*Risky willingness to accept (RWTA)*. Suppose the individual's current endowment is  $(x_i' + 1, x_j')$ . Consider a gamble with two mutually exclusive outcomes, each with probability 0.5. One outcome is that the individual loses one unit of good  $i$ , with no change in good  $j$ . The other is that she gains some amount of good  $j$ , with no change in good  $i$ .  $RWTA_{ji}$  is the smallest such gain of good  $j$  consistent with her being willing to accept the gamble.

In order to make use of these measures, some additional assumptions are necessary. These assumptions specify a restricted model of reference-dependent preferences under risk, incorporating elements of Kahneman and Tversky's (1979) prospect theory. Assume, as in prospect theory, that reference-dependent preferences over consumption bundles depend only on the displacement of each bundle from the reference state; thus, there are no income effects.<sup>9</sup> Assume that these preferences can be represented by an additively separable *value function*. Thus, in the case of two goods, the value function can be written as  $v(\Delta x) = v_1(\Delta x_1) + v_2(\Delta x_2)$ , where  $\Delta x = (\Delta x_1, \Delta x_2)$  is a *displacement vector* of changes in consumption relative to the reference state. Set  $\Delta x_1 > 0$  and  $\Delta x_2 < 0$ , and define the *gain component* of  $\Delta x$  as  $\Delta x^+ = (\Delta x_1, 0)$  and the *loss component* as  $\Delta x^- = (0, \Delta x_2)$ . Applying prospect theory,<sup>10</sup> whether the individual prefers this lottery to

<sup>9</sup> In the framework of reference-dependent preference theory, this is a simplifying assumption. In that theory, preferences depend both on the absolute levels of consumption and on the reference state.

<sup>10</sup> This analysis uses prospect theory as presented by Kahneman and Tversky (1979), or (equivalently in this case) the rank-dependent formulation proposed by Starmer and Sugden (1989). In Tversky and Kahneman's (1992) rank-dependent formulation, there are separate probability weighting functions for gains and losses.

the reference state depends on the sign of  $\pi(0.5)v(\Delta x^+) + \pi(0.5)v(\Delta x^-)$ , where  $\pi(\cdot)$  is the *probability weighting function*. Equivalently (cancelling out the  $\pi(0.5)$  terms and using additive separability), this preference depends on the sign of  $v(\Delta x)$ . Thus, if the individual is indifferent between the reference state and the displacement vector  $\Delta x$ , she is also indifferent between the reference state and a *balanced lottery* which gives her a 0.5 chance of the gain component of  $\Delta x$  and a 0.5 chance of its loss component. This property of *balanced-lottery risk neutrality* allows balanced lotteries to be used to elicit loss aversion.

As a simplification, assume the following functional form for the value function: for each good  $i$ ,  $v_i(\Delta x_i) = a_i(\Delta x_i)^\beta$  if  $\Delta x_i \geq 0$  and  $v_i(\Delta x_i) = -b_i(\Delta x_i)^\beta$  if  $\Delta x_i \leq 0$ , where  $a_i$ ,  $b_i$  and  $\beta$  are constants satisfying  $a_i > 0$ ,  $b_i > 0$ ,  $1 \geq \beta > 0$ . Given this functional form, the value of  $b_i/a_i$  is a natural measure of loss aversion in good  $i$ ;  $\beta$  is a parameter which represents the extent of ‘diminishing sensitivity’ for gains and losses, lower values of  $\beta$  corresponding with stronger effects of diminishing sensitivity. This model implies the following relationships for  $i = M, G$ :

$$EG_{ji} = [a_i / a_j]^{1/\beta} \quad (1)$$

$$RWTP_{ji} = [a_i / b_j]^{1/\beta} \quad (2)$$

$$RWTA_{ji} = [b_i / a_j]^{1/\beta} \quad (3)$$

If the current endowment hypothesis holds, the model also implies:

$$WTP_{MG} = [a_G / b_M]^{1/\beta} \quad (4a)$$

$$WTP_{GM} = [a_M / b_G]^{1/\beta} \quad (5)$$

$$WTA_{MG} = [b_G / a_M]^{1/\beta} \quad (6)$$

$$WTA_{GM} = [b_M / a_G]^{1/\beta} \quad (7a)$$

However, if NLIB holds, outlays of money in  $WTP_{MG}$  and  $WTA_{GM}$  valuations are construed as foregone gains rather than as losses. Thus, NLIB implies (5) and (6), but in place of (4a) and (7a), it implies:

$$WTP_{MG} = [a_G / a_M]^{1/\beta} = EG_{MG} \quad (4b)$$

$$WTA_{GM} = [a_M / a_G]^{1/\beta} = EG_{GM} \quad (7b)$$

To simplify notation, we define  $A_i = [b_i / a_i]^{1/\beta}$  for  $i = M, G$ . Notice that the assumptions of the model imply  $A_i > 1$ . Notice also that, for any given value of  $\beta$ ,  $A_i$  is an increasing function of  $b_i/a_i$ ; thus, if the extent of diminishing sensitivity is assumed constant across goods,  $A_i$  can be interpreted as a measure of loss aversion in good  $i$ . Using equations (1) to (7b), it is straightforward to derive the predictions shown in Table 2. These can be used to discriminate between NLIB and current endowment hypothesis.

**Table 2: Predictions of restricted form of reference-dependent theory**

Ratio	Value of ratio predicted by restricted form of reference-dependent theory	
	With current endowment hypothesis	With NLIB
$RWTA_{MG}/WTA_{MG}$	1	1
$WTA_{MG}/EG_{MG}$	$A_G$	$A_G$
$EG_{MG}/WTP_{MG}$	$A_M$	1
$WTP_{MG}/RWTP_{MG}$	1	$A_M$
$RWTA_{GM}/WTA_{GM}$	1	$A_M$
$WTA_{GM}/EG_{GM}$	$A_M$	1
$EG_{GM}/WTP_{GM}$	$A_G$	$A_G$
$WTP_{GM}/RWTP_{GM}$	1	1

$A_G, A_M > 1$ .

It can be seen from Table 2 that the extent of loss aversion in money – insofar as this is measured by the value of the parameter  $A_M$  – can be observed independently of assumptions about the truth or falsity of NLIB. Whether it is combined with the current endowment hypothesis or with NLIB, the restricted form of reference-dependent theory predicts  $EG_{MG}/RWTP_{MG} = RWTA_{GM}/EG_{GM} = A_M$ . This method of measuring loss aversion is valid only if the assumption of balanced-lottery risk neutrality holds. However, Table 2 shows how this assumption can be tested. The predictions  $RWTA_{MG}/WTA_{MG} = 1$  and  $WTP_{GM}/RWTP_{GM} = 1$  are direct implications of balanced-lottery risk aversion; and these predictions are independent of whether NLIB holds or not.

As we have explained, prospect theory offers reasons for expecting balanced-lottery risk neutrality to hold. One might ask whether this prior expectation is supported by other theories of choice under uncertainty. The answer is that *some* other theories provide such support. In particular, expected utility theory implies balanced-lottery risk neutrality if the utility function is separable in the two goods. However, other theories predict *balanced-lottery risk aversion*. That is, if an individual is indifferent between the reference state and some displacement vector  $\Delta x$ , she prefers the reference state to a lottery which gives her a 0.5 chance of the gain component of  $\Delta x$  and a 0.5 chance of its loss component. For example, rank-dependent expected utility theory (Quiggin, 1993) predicts balanced-lottery risk-aversion if the probability weighting function has the standard property  $\pi(0.5) < 0.5$ .

## 6. Tactical Heuristics

In experiments which elicit valuations, there is a potentially confounding factor: the possibility that subjects might follow *tactical* (or cautious) heuristics. Here it is useful to distinguish between two forms of valuation. An *outgoing* valuation records the largest amount of some good that an individual is willing to transfer *to* someone else; an *incoming* valuation records the smallest amount of some good that an individual is willing to accept as a transfer *from* someone else.  $WTP_{MG}$ ,  $WTP_{GM}$ ,  $RWTP_{MG}$  and  $RWTP_{GM}$  are outgoing valuations, while  $WTA_{MG}$ ,  $WTA_{GM}$ ,  $EG_{MG}$ ,  $EG_{GM}$ ,  $RWTA_{MG}$  and  $RWTA_{GM}$  are incoming valuations. It is possible that subjects may follow heuristics that lead them to understate their true outgoing valuations and to overstate their true incoming valuations. Although such heuristics do not in fact serve a subject's interests in incentive-compatible experiments such as those we discuss in this paper, they may be well-adapted to many real-world situations in which terms of trade are determined through bargaining. More generally, the principle of erring on the side of caution in appraising other people's trading proposals may be useful in a world in which not all economic interactions are positive-sum games.

Recall that the predictions displayed in Tables 1 and 2 concern *ratios* of valuations. Tactical heuristics are capable of producing confounding effects when the predicted ratio is between an outgoing and an ingoing valuation. Predictions are made about two such ratios:  $EG_{MG}/WTP_{MG}$  and  $EG_{GM}/WTP_{GM}$ . Tactical heuristics, if operative, will tend to increase the values of these ratios. In the case of  $EG_{MG}/WTP_{MG}$ , such an effect could provide spurious support for the current endowment hypothesis (which predicts a ratio greater than one) relative to NLIB (which predicts a ratio of one).

One obvious implication is that elicitation methods should be designed to minimise the salience of tactical considerations. Such considerations seem more likely to come into play if a subject perceives her task as reporting a *valuation* than if she perceives it as making a *dichotomous choice*. (Reporting a valuation may evoke comparisons with problems in which terms of trade are negotiated, in which it may pay to misrepresent one's true preferences; a dichotomous choice is more likely to evoke comparisons with decisions about whether or not to trade at exogenously determined prices, in which there is nothing to gain by misrepresentation.) This line of reasoning counts against the use of *open-ended* elicitation tasks (for example: 'What is the largest amount of money you would be willing to pay for ...?'). In our experiment, we instead use *multiple dichotomous choice* tasks: a subject responds to a series of

questions, each of which presents a choice between two options (for example: 'Would you be willing to pay £1 for ...?').<sup>11</sup>

A more radical solution to the problem of tactical heuristics is not to use  $EG_{MG}/WTP_{MG}$  or  $EG_{GM}/WTP_{GM}$  ratios in testing the rival hypotheses. As Tables 1 and 2 show, the need to use such ratios can be avoided if an experiment elicits, not only valuations of the good in units of money, but also valuations of money in units of the good. It is possible to discriminate between the current endowment hypothesis and NLIB by using only the ratios  $WTA_{MG}/EG_{MG}$ ,  $WTA_{GM}/EG_{GM}$ ,  $RWTA_{MG}/WTAP_{MG}$  and  $RWTA_{GM}/WTA_{GM}$ , all of which are ratios between incoming valuations.

---

<sup>11</sup> The Vancouver experiment used multiple dichotomous choice tasks, while the Norwich experiment used open-ended tasks. At the outset of our adversarial collaboration, Kahneman suggested that this design difference might have contributed to the differences in the results. In the light of experience gained after designing the Norwich experiment, the British group had independently come to the view that dichotomous choice tasks were more effective than open-ended valuation tasks in screening out tactical heuristics.

## 7. Experimental Design and Prospective Interpretations of Results

In the discussions which preceded the design of the experiment, we considered the range of possible tests described in Sections 3 to 5.

Seeking to maximise experimental control, the British group favoured a design which investigated the ratios  $WTA_{MG}/EG_{MG}$  and  $WTA_{GM}/EG_{GM}$ . Each of Hicksian theory, reference-dependent theory with the current endowment hypothesis, and reference-dependent theory with NLIB makes different predictions about these ratios (see Table 1); and the design controls for tactical heuristics by using only incoming valuations. The main disadvantage of this design is that it cannot measure the extent of loss aversion in money independently of assumptions about the truth or falsity of NLIB. In addition, Kahneman had reservations about eliciting valuations of fixed amounts of money in units of a consumption good: he was concerned that subjects might have difficulty in understanding tasks of this kind, or not construe them as buying tasks.

Kahneman favoured a design which investigated the ratios  $RWTA_{MG}/WTA_{MG}$ ,  $WTA_{MG}/EG_{MG}$ ,  $EG_{MG}/WTP_{MG}$ , and  $WTP_{MG}/RWTP_{MG}$ . Given the background assumptions of the restricted form of reference-dependent theory, these ratios can be used to discriminate between the current endowment hypothesis and NLIB (see Table 2). They also allow the relative magnitude of loss aversion in the good and in money to be measured, independently of assumptions about the truth or falsity of NLIB. Kahneman argued that, in proposing a model which included both NLIB and these background assumptions, he was making a bold set of predictions. Indeed, going beyond the qualitative predictions set out in Table 2, he predicted that the ratios  $WTP_{MG}/RWTP_{MG}$  and  $WTA_{MG}/EG_{MG}$  would take values close to 2.<sup>12</sup> The disadvantage of this design is that it depends on auxiliary assumptions about balanced-lottery risk neutrality, the absence of tactical heuristics, and the absence of income effects. Thus, if Kahneman's predictions were not confirmed, the evidence generated by the experiment would be open to a range of different interpretations.

Each of the parties to the adversarial collaboration recognised the advantages and disadvantages of the two designs. Combining the two proposals, we agreed

<sup>12</sup> Using data from an experiment in which subjects reported certainty equivalents for lotteries with money consequences, Tversky and Kahneman (1992) fit a model similar to that presented in Section 5, separately for each subject. The median value of  $b_M/a_M$  is 2.25. The median value of  $\beta$  is 0.88. Thus, one might expect typical values of  $[b_M/a_M]^{1/\beta}$ , i.e. of  $A_M$ , to be in the region of  $2.25^{0.88} = 2.04$ . On the basis of the results of the Vancouver experiment, Kahneman expected typical values of  $WTA_{MG}/EG_{MG}$ , i.e. of  $A_G$ , to be close to 2 also.

on a design that would elicit all of the valuations  $EG_{MG}$ ,  $WTP_{MG}$ ,  $WTA_{MG}$ ,  $RWTP_{MG}$ ,  $RWTA_{MG}$ ,  $EG_{GM}$ ,  $WTA_{GM}$  and  $RWTA_{GM}$ . This design would provide data for all of the tests proposed by the two parties. In addition, it would allow us to test the predictions shown in the fifth and sixth rows of Table 2, concerning the ratios  $RWTA_{GM}/WTA_{GM}$  and  $WTA_{GM}/EG_{GM}$ . These tests are of interest because they allow us to assess the extent of loss aversion in money, independently of the truth or falsity of NLIB, while using only incoming valuations.

In the spirit of adversarial collaboration, we sought to agree in advance on the inferences to be drawn from the various possible outcomes of the experiment.

With respect to comparisons between  $WTA_{MG}$  and  $EG_{MG}$  and between  $WTA_{GM}$  and  $EG_{GM}$ , the main problem in interpreting results is the lack of independent evidence about the relative magnitudes of underlying loss aversion in money and in the good. However, both parties agreed that, on the basis of theory and of currently available evidence, there was no specific reason to expect underlying loss aversion to be greater in one case than in the other.<sup>13</sup> Thus, we were able to agree on two points of interpretation. First, if  $WTA_{GM}$  valuations were found to be significantly greater than  $EG_{GM}$  valuations, that would be evidence against NLIB. In itself, however, that observation would not be evidence that the truth was closer to the current endowment hypothesis than to NLIB. Second, there was a presumption that, if the current endowment hypothesis was true,  $WTA_{GM}/EG_{GM}$  would be approximately equal to  $WTA_{MG}/EG_{MG}$ . Thus, conditional on our observing  $WTA_{MG}/EG_{MG} \geq 1$  and  $WTA_{GM}/EG_{GM} \geq 1$ , the *relative* magnitudes of  $(WTA_{GM}/EG_{GM}) - 1$  and  $(WTA_{MG}/EG_{MG}) - 1$  would be an indication of the relative success of the two hypotheses.

With respect to those tests that are based on the restricted form of reference-dependent theory, we agreed that our interpretation of these tests would be conditional on our finding no significant difference between  $RWTA_{MG}$  and  $WTA_{MG}$ . The equality of  $RWTA_{MG}$  and  $WTA_{MG}$  is predicted by the restricted form of reference-dependent theory, independently of whether NLIB is true or false. It represents the property of balanced-lottery risk neutrality, which is crucial for our interpretations of the other  $RWTA$  and  $RWTP$  valuations.

Assuming this condition to be satisfied, we agreed on the following prospective interpretations. If  $WTP_{MG}$  valuations were significantly greater than  $RWTP_{MG}$ , or if  $RWTA_{GM}$  valuations were significantly greater than  $WTA_{GM}$ , that would be

<sup>13</sup> Interpreting existing evidence in the light of the restricted form of reference-dependent theory, Kahneman argued that there was some positive reason to expect the two kinds of loss aversion to be approximately equal in magnitude: see footnote 13.

evidence against the current endowment hypothesis (since it would indicate loss aversion in money in a situation in which, according to that hypothesis, loss aversion should not be observed). If  $WTA_{GM}$  valuations were significantly greater than  $EG_{GM}$ , that would be evidence against NLIB (since it would indicate loss aversion in money in a situation in which, according to NLIB, it should not be observed). If  $EG_{MG}$  valuations were significantly greater than  $WTP_{MG}$ , that would be evidence *either* against NLIB (for the same reasons as apply to the comparison between  $WTA_{GM}$  and  $EG_{GM}$ ) *or* of tactical heuristics (since  $EG_{MG}$  is an incoming valuation while  $WTP_{MG}$  is an outgoing one).<sup>14</sup> The relative success of the current endowment hypothesis and of NLIB would be indicated by the relative degrees of loss aversion in money, as observed in those comparisons in which only the current endowment hypotheses predicts its occurrence and in those comparisons in which only NLIB predicts it. Thus, for example, if the value of  $EG_{MG}/WTP_{MG}$  was close to 1, but the value of  $WTP_{MG}/RWTP_{MG}$  was markedly greater than 1, that would suggest that NLIB was closer to the truth than the current endowment hypothesis.

<sup>14</sup> In principle,  $EG_{MG} > WTP_{MG}$  might also be evidence of the income and substitution effects predicted by Hicksian theory. But in the context of our planned experiment, such effects would be very small.

## 8. The Experiment

The experiment was carried out at the University of East Anglia. Subjects were recruited from the undergraduate population by means of e-mailed invitations; they were broadly representative of that population in terms of age, gender, and subject of study. Subjects were required to bring cash to the experiment, but were assured that any opportunities to spend money would be optional. Initially, we recruited 320 subjects. Each subject was allocated at random to one of eight treatments, each of which was designed to elicit one of the valuations  $EG_{MG}$ ,  $WTP_{MG}$ ,  $WTA_{MG}$ ,  $RWTP_{MG}$ ,  $RWTA_{MG}$ ,  $EG_{GM}$ ,  $WTA_{GM}$  and  $RWTA_{GM}$ . Thus, each subject confronted just one valuation task.<sup>15</sup>

The specific good took the form of luxury chocolates sold by a specialist shop located in the centre of Norwich, easily accessible from the university campus. These chocolates are sold by weight, at an average price of about £0.30 each. To allow exchanges in units of single chocolates to be carried out conveniently,<sup>16</sup> transactions within the experimental sessions were carried out in vouchers. A voucher entitled its holder to a specified number of chocolates, free of charge, when presented at the shop.

On arrival at a session, subjects were told that the experiment had two separate parts. They were told nothing about Part 2 until they had completed Part 1, except that any payoffs they might receive in Part 2 would be additional to their payoffs from Part 1. In fact, Part 2 was a choice, for real, between two lotteries with money prizes. The main purpose of this part of the experiment was to supplement subjects' final earnings, particularly in two treatments ( $WTP_{MG}$  and  $RWTP_{MG}$ ) in which those earnings would otherwise have been rather low.<sup>17</sup> In this paper, we are concerned only with Part 1 of the experiment.

Depending on which task they had been assigned, subjects were given 'endowments' (which in some cases were 'nothing'). Subjects in the  $WTA_{GM}$  and  $RWTA_{GM}$  groups were given £1.00. Those in the  $WTA_{MG}$  and  $RWTA_{MG}$  groups

<sup>15</sup> We considered using an alternative design in which each subject is presented with a series of separate tasks; which task is for real is determined by a random-lottery mechanism at the end of the experiment. Since different tasks involve different endowments, this design makes subjects' endowments contingent on random events, rather than certain. We decided not to use this design because of the possibility that it might attenuate loss aversion – an effect for which there is some evidence (Loewenstein and Adler, 1995). The Norwich experiment had used a random-lottery design; in the Vancouver experiment, each subject had faced only one valuation task.

<sup>16</sup> Such exchanges are necessary for the  $EG_{GM}$ ,  $WTA_{GM}$  and  $RWTA_{GM}$  tasks, which elicit valuations of money in units of chocolate. The chocolates are not individually wrapped, but are boxed to order at the shop.

<sup>17</sup> The second part of the experiment is reported by Cubitt, Starmer and Sugden (2001).

were given 10 chocolates (in the form of vouchers). All other subjects were given nothing. Endowments (money or vouchers) were physically handed over to subjects. It was explained that subjects' endowments were theirs to keep if they so chose. The conditions for the use of the vouchers were explained, and samples of the chocolates were shown; no information was given about the price of the chocolates.

The concept of a 'lottery' was also explained. Lotteries would be resolved by the subject drawing a disc from a bag containing 100 discs, numbered from 1 to 100. The outcomes of lotteries were described in terms of what the subject would gain or lose, conditional on the number of the disc drawn.

Subjects were given booklets in which their tasks were set out. The rest of the instructions were printed in these booklets. As far as possible, the instructions were common to all treatments; the common elements were also read out by an experimenter, who fielded any questions. Subjects were told that they had to make twenty-five 'choices'. For each subject, one of these choices was for real. On arrival at the experiment, subjects had been shown a box containing twenty-five sealed envelopes, each containing a ticket with one of the numbers 1-25. Each subject picked one envelope from the box, to be opened at the end of the experiment. The number in the envelope was the number of the choice problem that was for real for that subject. This device was used to dramatise the fact that one and only one problem was for real, and that the identity of this problem was independent of the subject's responses. At the end of the experiment, each subject carried out whatever transaction (if any) he had chosen in the problem that was for real.

Each choice problem required the subject to choose one of two 'options', one displayed on the left-hand side of the page and one on the right, by ticking the appropriate box. For any given subject, the right-hand option was the same in all twenty-five problems. The left-hand options differed only in respect of one parameter. In the  $EG_{MG}$ ,  $WTP_{MG}$ ,  $WTA_{MG}$ ,  $RWTP_{MG}$  and  $RWTA_{MG}$  treatments, this parameter was an amount of money from the set  $\{\pounds 0.30, \pounds 0.60, \pounds 0.90, \dots, \pounds 7.50\}$ ; we shall say that in these treatments the *response mode* was money. In the  $EG_{GM}$ ,  $WTA_{GM}$  and  $RWTA_{GM}$  treatments, it was a number of chocolates from the set  $\{1, \dots, 25\}$ ; in these cases, the response mode was chocolate. For each treatment, half of the booklets presented the problems in ascending order (i.e. in Choice 1, the parameter was  $\pounds 0.30$  or 1, in Choice 2 it was  $\pounds 0.60$  or 2, and so on), while the other booklets presented them in descending order. Subjects were allocated randomly between these two presentations, so as to

control for order effects. We required each subject's choices to be mutually consistent in the sense of respecting dominance.<sup>18</sup>

The  $EG_{MG}$ ,  $WTP_{MG}$ ,  $WTA_{MG}$ ,  $RWTP_{MG}$  and  $RWTA_{MG}$  treatments elicit, to within  $\pounds 0.30$  bands, money valuations of 10 chocolates. (In the notation of Section 3, with  $m$  denoting a subject's money wealth before coming to the experiment and on the assumption that no one then owned chocolates: 10 physical chocolates constitute one 'unit' of chocolate,  $x_1' = 0$ , and  $x_2' = m$ .) The  $EG_{GM}$ ,  $WTA_{GM}$  and  $RWTA_{GM}$  treatments elicit chocolate valuations of  $\pounds 1.00$  (i.e.  $\pounds 1.00$  constitutes one 'unit' of money,  $x_1' = 0$ , and  $x_2' = m$ ). Figure 1 shows how a typical problem for each treatment (in each case, the twelfth problem in ascending order) was displayed. The entries in square brackets were not seen by subjects; these have been added for the benefit of the reader, to identify the relevant treatment.

Two additional treatments were run in a follow-up experiment. The responses to the treatments described above turned out to indicate surprisingly low levels of loss aversion for chocolate, as measured by the ratio  $WTA_{MG}/EG_{MG}$ . Kahneman conjectured that this was the result of our having used vouchers rather than actual chocolates in the experimental sessions. The use of vouchers, he suggested, might attenuate loss aversion by mentally distancing subjects from the consumption experiences associated with the chocolates, and thus weakening the sense of ownership associated with chocolate endowments. In addition, since vouchers have some of the properties of money, the psychological mechanisms which (on his account) give rise to NLIB might also affect tasks in which vouchers are given up in trade. To test this conjecture, we ran a follow-up experiment which repeated the  $EG_{MG}$  and  $WTA_{MG}$  treatments, exactly as before except for one detail: the '10 chocolates' took the form of a pre-packed box of 10 chocolates (the same kind as we had used before) rather than vouchers to be redeemed at the supplier's shop. At the start of the follow-up experiment, each subject in the  $WTA_{MG}$  treatment was handed such a box as his endowment; subjects who retained or gained chocolates in the course of the experiment took a box of chocolates away with them. Using the same procedures as before, we recruited an additional 107 subjects and divided them at random between the two additional treatments. These treatments are denoted by  $EG_{MG}^*$  and  $WTA_{MG}^*$ ; we shall say that they involved *immediate chocolate* as contrasted with chocolate vouchers.

<sup>18</sup> For example, if an EG valuation was being elicited in ascending order, a subject who chose the left-hand option in any given choice problem was not allowed to choose the right-hand option in a subsequent problem. If a subject's responses were mutually inconsistent, the nature of the inconsistency was explained to him, and he was asked to revise those responses. In fact, all but three of the 427 subjects responded consistently at the first attempt. Since all the theories we consider (indeed, all the theories we know) imply this form of consistency, requiring responses to satisfy it does not bias our tests. We think it reasonable to interpret deviations from this consistency condition as errors.

**Figure 1: Examples of tasks used in the experiment**

<b>Choice 12</b> [EG <sub>MG</sub> ]	We give you £3.60	9	<b>or</b>	We give you 10 chocolates	9
<b>Choice 12</b> [WTA <sub>MG</sub> ]	You give us your 10 chocolates and take £3.60 in exchange	9	<b>or</b>	You keep your 10 chocolates	9
<b>Choice 12</b> [RWT <sub>MG</sub> ]	You enter the lottery shown below: discs 1-50 You lose your 10 chocolates	9	<b>or</b>	You keep your 10 chocolates	9
				discs 51-100 You keep your 10 chocolates and win £3.60 in addition	
<b>Choice 12</b> [WTP <sub>MG</sub> ]	You give us £3.60 and take 10 chocolates in exchange	9	<b>or</b>	You do not trade	9
<b>Choice 12</b> [RWTP <sub>MG</sub> ]	You enter the lottery shown below: discs 1-50 You lose £3.60	9	<b>or</b>	You do not enter it	9
				discs 51-100 You win 10 chocolates	
<b>Choice 12</b> [EG <sub>GM</sub> ]	We give you 10 chocolates	9	<b>or</b>	We give you £1.00	9
<b>Choice 12</b> [WTA <sub>GM</sub> ]	You give us your £1.00 and take 10 chocolates in exchange	9	<b>or</b>	You keep your £1.00	9
<b>Choice 12</b> [RWT <sub>GM</sub> ]	You enter the lottery shown below: discs 1-50 You lose your £1.00	9	<b>or</b>	You keep your £1.00	9
				discs 51-100 You keep your £1.00 and win 10 chocolates in addition	

## 9. Results

The responses to the ten treatments are summarised in Tables 3 and 4. In presenting the data, we use the following conventions. Recall that for any given subject in any given treatment, there are 25 choice problems. Since subjects' choices are required to respect dominance, there are 26 alternative permissible ways of answering any such set of problems: *either* the left-hand option is chosen in every problem, *or* there is a switch from left to right (or right to left, depending on the task) after exactly one of the choices 1, ..., 24, *or* the right-hand option is chosen in every problem. Each response indicates a different *valuation* of the right-hand option, expressed in terms of one of the valuation measures defined in Section 3.

We assign these responses the values 1, ..., 26, in ascending order of the valuation of 10 chocolates (for tasks in which the response mode is money) or in ascending order of the valuation of £1.00 (for tasks in which the response mode is chocolate). Thus, for tasks in which the response mode is money, the valuation 1 corresponds with the range of money values of 10 chocolates from zero to £0.30; the valuation 2 corresponds with values from £0.30 to £0.60, and so on up to the valuation 26 which corresponds with values from £7.50 upwards. For tasks in which the response mode is chocolate, the valuation 1 corresponds with the range of chocolate values of £1.00 from 0 to 1 chocolate; the valuation 2 corresponds with values from 1 to 2 chocolates, and so on up to the valuation 26 which corresponds with values from 25 chocolates upwards.

We also report subjects' *implicit preferences* between 10 chocolates and £1.00. In treatments in which the right-hand option is 10 chocolates, a subject whose valuation is 1, 2 or 3 has chosen to have £0.90 rather than 10 chocolates, and so can be presumed to prefer £1.00 to 10 chocolates. Conversely, a subject whose valuation is 5 or more has chosen to have 10 chocolates rather than £1.20, and so can be presumed to prefer 10 chocolates to £1.00. (A valuation of 4 does not reveal the subject's preference either way.) In treatments in which the right-hand option is £1.00, the valuations 1, ..., 10 reveal an implicit preference for £1.00 over 10 chocolates, while the valuations 11, ..., 26 reveal the opposite preference. Implicit preferences are of interest because they are comparable across all treatments, irrespective of whether the response mode is money or chocolate. Notice that (10 chocolates, £1.00) is the *only* pair for which our design allows us to identify implicit preferences for all treatments.

Table 3 refers to the seven treatments for which the response mode was money. The upper part of the table reports, for each treatment, the geometric mean, arithmetic mean, median and standard deviation of the distribution of subjects' valuations. We shall give particular attention to the geometric mean. This is

because we are concerned with the values of *ratios* and, as summary statistics for describing ratios, geometric means are more satisfactory than arithmetic means. More specifically, we are concerned with two kinds of ratios. First, it is fundamental to our experimental design that the money and chocolate response modes are symmetrical with one another. Viewed in this perspective, valuations should be interpreted as ratios between quantities of money and (equally preferred) quantities of chocolate; whether these ratios are expressed as '£ per chocolate' or 'chocolates per £' is arbitrary. Suppose we have valuations  $v_1, \dots, v_n$  in units of £/chocolate from  $n$  subjects. Since the geometric mean of  $1/v_i$  is equal to the inverse of the geometric mean of  $v_i$ , the information content of the geometric mean of those valuations is independent of the units in which they are expressed. Second, we shall be making use of ratios of valuations (such as the ratio  $WTA_{MG}/EG_{MG}$ , used as an indicator of the extent of loss aversion in chocolate). Suppose we have valuations  $v_1, \dots, v_n$  from one treatment and valuations  $w_1, \dots, w_n$  from another treatment, with different subjects in the two treatments. Irrespective of how subjects are indexed, the ratio between the geometric mean of  $v_i$  and the geometric mean of  $w_i$  is equal to the geometric mean of  $(v_i/w_i)$ .<sup>19</sup>

The lower part of Table 3 reports, for each treatment, the distribution of responses classified by implicit preferences. Table 4 presents the corresponding data for the three treatments for which the response mode was chocolate.

Table 5 reports summary statistics for those comparisons between treatments that are relevant for our design. The first entry in each row identifies a ratio of two valuations. The next five entries indicate whether particular causal factors, if operating, have a tendency to increase the value of that ratio above 1. (Each of the five factors, if present, works in the same direction.) The second entry states whether loss aversion in chocolate and/or money would tend to increase the value of the ratio, and in the case of loss aversion in money, whether this tendency is conditional on the truth of NLIB or on the truth of the current endowment hypothesis (CEH). The third entry states whether balanced-lottery risk aversion would have the same tendency. The fourth entry states whether subjects' use of tactical heuristics would have that tendency. In this column, 'yes' signifies that the numerator of the ratio is an incoming valuation and that the denominator is an outgoing valuation; 'no' signifies that both valuations are of the same type. The fifth entry states whether Hicksian income and substitution effects would have that tendency. Here, 'yes' signifies that it is a

<sup>19</sup> A possible objection to the use of mean valuations (whether arithmetic or geometric) is that they are sensitive to extreme values. However, our subjects rarely used the extremes of the response scales. Of the 307 subjects using the money response mode, only 13 recorded the lowest valuation 1 and only 9 recorded the highest valuation 26. For the 120 subjects using the chocolate response mode, the corresponding numbers were 6 and 5.



prediction of Hicksian theory that, if income effects are normal, the ratio of valuations is greater than 1. (Notice, however, that the *size* of such effects can be expected to be small: see Section 4.) ‘No’ signifies that Hicksian theory predicts that the two valuations are exactly equal; ‘n.a.’ signifies that Hicksian theory makes no firm predictions (because of the presence of risk). The sixth entry indicates whether a difference between subjects’ attitudes to immediate chocolate and to chocolate vouchers would impact on the value of the ratio; we assume that such an effect, if it existed, would imply higher valuations for immediate chocolates than for chocolate vouchers.

**Table 3: Responses to tasks with money as the response mode**

task (in = incoming, out = outgoing)							
	RWTP <sub>MG</sub> (out)	WTP <sub>MG</sub> (out)	EG <sub>MG</sub> (in)	WTA <sub>MG</sub> (in)	RWTA <sub>MG</sub> (in)	EG <sub>MG</sub> * (in)	WTA <sub>MG</sub> * (in)
<i>valuations of 10 chocolates</i> (units of £0.30):							
geometric mean	5.38	4.66	8.17	9.95	10.17	8.69	11.30
arithmetic mean	6.75	5.55	10.00	10.80	12.70	10.24	12.46
median	5	6	10	10	12	10	14
standard deviation	4.99	2.82	5.05	4.58	7.35	4.94	5.19
<i>implicit preferences:</i>							
no of subjects who:							
prefer £1	9	9	5	0	3	3	3
not clear <sup>a</sup>	7	4	1	1	1	8	1
prefer 10 chocolates	24	27	34	39	36	44	48
(% who prefer chocolates)	(60.0)	(67.5)	(85.0)	(97.5)	(92.3)	(80.0)	(92.3)
total	40	40	40	40	40	55	52

<sup>a</sup> Subjects whose responses indicated that the valuation of 10 chocolates was at least £0.90 but no more than £1.20

**Table 4: Responses to tasks with chocolate as the response mode**

task (in = incoming, out = outgoing)			
	RWTA <sub>GM</sub> (in)	WTA <sub>GM</sub> (in)	EG <sub>GM</sub> (in)
<i>valuations of £1.00 (units of 1 chocolate):</i>			
geometric mean	9.42	9.62	7.52
arithmetic mean	12.70	10.95	8.85
median	12.5	10	8
standard deviation	7.75	5.70	4.84
<i>implicit preferences:</i>			
no of subjects who:			
prefer £1	24	17	12
prefer 10 chocolates	16	23	28
(% who prefer chocolates)	(40.0)	(57.5)	(70.0)
<b>Total</b>	40	40	40

**Table 5: Comparisons of valuations**

comparison	picks up loss aversion in:	picks up balanced-lottery risk aversion	picks up tactical heuristics	picks up Hicksian effects	picks up voucher effects	ratio of geometric means	Mann-Whitney test: zstatistic
WTA <sub>MG</sub> /WTP <sub>MG</sub>	money (if CEH true) and chocolate	no	yes	yes	no	2.13	5.46##
WTA <sub>MG</sub> /EG <sub>MG</sub>	chocolate	no	no	no	no	1.22	0.37
RWTA <sub>MG</sub> /WTA <sub>MG</sub>	neither	yes	no	n.a.	no	1.02	0.93
EG <sub>MG</sub> /WTP <sub>MG</sub>	money (if CEH true)	no	yes	yes	no	1.75	4.23##
WTP <sub>MG</sub> /RWTP <sub>MG</sub>	money (if NLIB true)	yes	no	n.a.	no	0.87	-0.57
WTA <sub>GM</sub> /EG <sub>GM</sub>	money (if CEH true)	no	no	no	no	1.28	1.49
RWTA <sub>GM</sub> /WTA <sub>GM</sub>	money (if NLIB true)	yes	no	n.a.	no	0.97	1.01
WTA <sub>MG</sub> */EG <sub>MG</sub> *	chocolate	no	no	no	no	1.30	2.03#
EG <sub>MG</sub> */EG <sub>MG</sub>	neither	no	no	no	yes	1.06	0.31
WTA <sub>MG</sub> */WTA <sub>MG</sub>	neither	no	no	no	yes	1.14	1.82#

In final column, # denotes significance at 5 per cent level in a one-tail test; ## denotes significance at 1 per cent level.

The seventh entry in each row is the ratio between the geometric means of the relevant valuations. We use this as our main indicator of the similarity or divergence between responses to different treatments. (If the reader wishes to compare medians or arithmetic means, the relevant information is given in Tables 3 and 4.) The final entry reports the z-statistic for a Mann-Whitney test for differences between the distributions of valuations in the two treatments; # denotes significance at the 5 per cent level in a one-tail test; ## denotes significance at the 1 per cent level. A positive sign indicates that ‘numerator’ valuations are greater than ‘denominator’ valuations.<sup>20</sup>

To provide a benchmark for other comparisons, we begin by comparing the  $WTA_{MG}$  and  $WTP_{MG}$  treatments. Recall that subjects in the  $WTA_{MG}$  treatment are endowed with chocolate and report their willingness to accept money in exchange for giving up their endowment; subjects in the  $WTP_{MG}$  treatment are given no endowments and report their willingness to spend money to buy chocolate. Many experiments and surveys have found willingness to accept to be greater than willingness to pay in comparisons of this kind: we shall call this the *classic*  $WTA/WTP$  comparison. Our experiment replicates the familiar result: the ratio of geometric means is 2.13 and the difference between the distributions of valuations is overwhelmingly significant ( $p < 0.001$ ). This result is not surprising, but it gives some assurance that our experiment is picking up whatever causal factors lie behind commonly-observed differences between willingness to accept and willingness to pay.

As the entries in the first row of Table 5 highlight, the comparison between  $WTA_{MG}$  and  $WTP_{MG}$  is deficient in experimental control. A  $WTA_{MG}/WTP_{MG}$  ratio greater than 1 could be evidence of loss aversion in chocolate, of loss aversion in money (combined with the current endowment hypothesis), of tactical heuristics, of Hicksian income and substitution effects, or of any combination of these factors. Our experiment was designed to allow the two kinds of loss aversion to be disentangled from each other and from these other effects.

We now turn to the relevant tests. In this Section, and in the spirit of adversarial collaboration, we confine ourselves strictly to the data generated by our experiment and to the tests we planned in advance of seeing those data. Discussion, interpretation, and comparisons with the results of other experiments are postponed to Section 9.

<sup>20</sup> We use a non-parametric test because, for several treatments, the hypothesis that the distribution of valuations is normal can be rejected. The alternative Kolmogorov-Smirnoff test for differences between distributions gives very similar results to the Mann-Whitney test; the only case of disagreement is the comparison between  $EG_{MG}^*$  and  $WTA_{MG}^*$  distributions, where the Kolmogorov-Smirnoff test does not find a significant difference ( $z = 0.96$ ).

First, we consider the tests for loss aversion which control for attitudes to risk, tactical heuristics and Hicksian effects. There are two such tests for loss aversion in chocolate, one comparing  $WTA_{MG}$  and  $EG_{MG}$  valuations from treatments in which chocolates were represented by vouchers, the other comparing  $WTA_{MG}^*$  and  $EG_{MG}^*$  valuations from treatments in which subjects were given boxes of chocolates. The first of these comparisons yields a ratio of geometric means of 1.22 and no significant difference between distributions. In the second comparison (with a larger sample size), the ratio of geometric means is 1.30 and the difference between the distributions *is* significant ( $p = 0.02$ ). According to the current endowment hypothesis, loss aversion in money is picked up in the comparison between  $WTA_{GM}$  and  $EG_{GM}$ ; in contrast, NLIB implies  $WTA_{GM}/EG_{GM} = 1$ . In fact, the ratio of geometric means is 1.28; the difference between distributions is not significant at the 5 per cent level ( $p = 0.07$ ).

These results provide some (albeit weak) confirmation for the hypothesis that, after Hicksian effects and tactical heuristics have been controlled for, there is loss aversion in chocolate. Is there loss aversion in money? We cannot reject NLIB with 95 per cent confidence; but there is no positive support for that hypothesis. The observed divergence between  $WTA_{GM}$  and  $EG_{GM}$ , which the current endowment hypothesis would allow us to treat as an indicator of loss aversion in money, is similar in magnitude to the divergences between  $WTA_{MG}$  and  $EG_{MG}$  and between  $WTA_{MG}^*$  and  $EG_{MG}^*$ , which (according to either hypothesis) are indicators of loss aversion in chocolate. This is consistent with our prior expectations about what we would observe, were the current endowment hypothesis true.

We now consider those tests that assume the restricted form of reference-dependent theory. Recall that it is a precondition for the validity of these tests that  $RWTA_{MG}$  is not significantly greater than  $WTA_{MG}$ . That precondition is satisfied. In fact, the ratio of geometric means for this comparison is 1.02, almost exactly as predicted by the hypothesis of balanced-lottery risk neutrality. Given the special assumptions of the restricted form of reference-dependent theory, NLIB implies that loss aversion in money will show up in the comparisons between  $WTP_{MG}$  and  $RWTP_{MG}$ , and between  $RWTA_{GM}$  and  $WTA_{GM}$ . In each case, NLIB predicts that the ratio of geometric means is greater than 1, while the current endowment hypothesis predicts a value of 1. In fact, the ratios are 0.87 and 0.97 respectively; in neither case is there a significant difference between the relevant distributions of valuations. These results give no support to NLIB. They are consistent with the restricted form of reference-dependent theory, combined with the current endowment hypothesis.

Given the same assumptions, the current endowment hypothesis implies that loss aversion in money will show up in the comparisons between  $EG_{MG}$  and

WTP<sub>MG</sub>, and between WTA<sub>GM</sub> and EG<sub>GM</sub>. In each case, the current endowment hypothesis predicts that the ratio of geometric means is greater than 1, while NLIB predicts a value of 1. We have already reported the results of the comparison between WTA<sub>GM</sub> and EG<sub>GM</sub>: the ratio of geometric means is 1.28, and the difference between distributions is not significant at the 5 per cent level. In the case of the comparison between EG<sub>MG</sub> and WTP<sub>MG</sub>, the ratio of geometric means is 1.75, and the two distributions of valuations are significantly different ( $p < 0.001$ ). In this latter case, then, the hypothesis derived from NLIB is rejected. However, it is possible that the comparison between EG<sub>MG</sub> and WTP<sub>MG</sub> has picked up the effects of tactical heuristics and/or Hicksian income and substitution effects.

Finally, we compare the two immediate chocolate treatments with the corresponding chocolate voucher treatments.<sup>21</sup> The distributions of valuations in the EG<sub>MG</sub>\* and EG<sub>MG</sub> treatments are remarkably similar, while WTA<sub>MG</sub>\* valuations are rather greater than WTA<sub>MG</sub> ones: the ratio of geometric means is 1.14, and the difference between distributions is significant ( $p = 0.04$ ). This finding is consistent with the hypothesis that loss aversion is attenuated if what individuals stand to lose is a voucher rather than an immediately consumable good. Notice, however, that the truth or falsity of this hypothesis has no implications for our tests of NLIB: those tests are concerned only with loss aversion *in money*.

<sup>21</sup> In making these comparisons, it should be remembered that the subjects in the WTA<sub>MG</sub>\* and EG<sub>MG</sub>\* treatments were recruited separately from those in the WTA<sub>MG</sub> and EG<sub>MG</sub> treatments. We have no reason to expect any systematic differences between the two subject pools, but the possibility that such differences exist cannot be ruled out completely.

## 10. Discussion

One of the most striking features of our results is the relative weakness of loss aversion effects in *all* those comparisons in which either or both of the parties to the collaboration expected such effects to show up. (The unexpected weakness of these effects is the main reason why our tests of NLIB were less decisive than we had hoped.)

Recall that both parties agreed about the tests to be used to detect loss aversion in chocolate: the comparisons between WTA<sub>MG</sub> and EG<sub>MG</sub> (for treatments using vouchers) and between WTA<sub>MG</sub>\* and EG<sub>MG</sub>\* (for treatments using immediate chocolates). Even in the latter case, WTA<sub>MG</sub>\* valuations were only 30 per cent greater than EG<sub>MG</sub>\* ones. We cannot be as categorical about the degree of loss aversion in money, since the form in which such loss aversion shows up depends on whether NLIB or the current endowment hypothesis is true. If the current endowment hypothesis is true, loss aversion in money is picked up in the comparison between WTA<sub>GM</sub> and EG<sub>GM</sub>, and here we found a divergence between the two valuations of a little less than 30 per cent. If NLIB is true, and given the special assumptions of the restricted form of reference-dependent theory, loss aversion in money is picked up in the comparisons between WTP<sub>MG</sub> and RWTP<sub>MG</sub> and between WTA<sub>GM</sub> and RWTA<sub>GM</sub>.

No evidence of loss aversion was found in either of these comparisons. Nevertheless, in the classic WTA/WTP comparison, we found a divergence of over 100 per cent between WTA<sub>MG</sub> and WTP<sub>MG</sub>. This suggests that our experiment has picked up *some* relatively strong causal factor – or combination of factors – which contributes to the classic WTA/WTP discrepancy that has been found in so many experiments and surveys. So what is that factor? In trying to answer this question, we go beyond the issues on which the parties reached agreement before running the experiment. We can offer two conjectures, one of which is favoured by the British group, the other by Kahneman.

The British group interprets the results as consistent with the current endowment hypothesis. Putting most weight on the most tightly-controlled tests (i.e. on the comparisons between WTA<sub>MG</sub> and EG<sub>MG</sub>, between WTA<sub>MG</sub>\* and EG<sub>MG</sub>\*, and between WTA<sub>GM</sub> and EG<sub>GM</sub>), they conclude that loss aversion in chocolate and loss aversion in money are both relatively weak effects. The implication is that although each of these forms of loss aversion is a contributory cause of the classic WTA/WTP discrepancy, they cannot be the *only* causes. The British group conjectures that a third factor is at work: subjects are using tactical or cautious heuristics which, irrespective of the response mode, generate relatively high incoming valuations and relatively low outgoing ones. On this hypothesis, we should expect relatively chocolate-loving

preferences to be revealed in those money-response tasks that elicit incoming valuations (i.e.  $EG_{MG}$ ,  $WTA_{MG}$ ,  $RWTA_{MG}$ ,  $EG_{MG}^*$  and  $WTA_{MG}^*$ ). Conversely, we should expect relatively money-loving preferences to be revealed both in the chocolate-response tasks (i.e.  $RWTA_{GM}$ ,  $WTA_{GM}$  and  $EG_{GM}$ ), since those tasks elicit incoming valuations in units of chocolate, and in those money-response tasks that elicit outgoing valuations (i.e.  $RWTP_{MG}$  and  $WTP_{MG}$ ). Implicit preferences do in fact show this general pattern (see Tables 3 and 4); the classic discrepancy between  $WTA_{MG}$  and  $WTP_{MG}$  is part of that pattern.

Kahneman, too, interprets the results of the joint experiment as failing to support NLIB. However, he interprets the observed difference between  $EG_{MG}$  and  $WTP_{MG}$  valuations as evidence of strong loss aversion in money (rather than as evidence of tactical heuristics). For Kahneman, the puzzle is to explain why loss aversion in money shows up in this comparison, and not (as NLIB would predict) in the comparison between  $WTP_{MG}$  and  $RWTP_{MG}$ . His tentative interpretation is that, contrary to his prior expectation, subjects treated money given up in return for chocolates as a loss. This account would be compatible with Kahneman's general theory of the conditions under which gains and losses are integrated *if* the Norwich subjects were so financially constrained that they did not perceive themselves as having budget reserves. Kahneman conjectures that this may have been the case, and that this may amount to an unanticipated difference between the Norwich subject pool and the subject pools that he has used previously.

If we look beyond the results of this particular experiment, there is a more general problem: to find a unified explanation for *all* the data that have so far been generated within this experimental paradigm. Specifically, the experiment reported in this paper, the Vancouver and Norwich experiments, and the additional experiments reported by Kahneman and Novemsky (2002) together provide a very large body of data, which one might hope to be able to organise into a consistent pattern. But finding such a pattern is not easy.

The results of our collaborative experiment are remarkably similar to those of the earlier Norwich experiment, despite the differences between the two designs. Specifically: in each experiment, the results are consistent with the current endowment hypothesis; loss aversion in money (as measured by  $WTA_{GM}/EG_{GM}$ ) is similar in magnitude to loss aversion in the good (as measured by  $WTA_{MG}/EG_{MG}$ ); both effects are quite weak; incoming valuations are markedly greater than outgoing valuations; and the classic disparity between  $WTA_{MG}$  and  $WTP_{MG}$  is a strong effect.

However, the results of the Vancouver experiment and of the experiments reported by Kahneman and Novemsky show a different overall pattern. In

common with the experiments we have just summarised, these results show the classic WTA/WTP disparity. But, in addition: they are broadly consistent with the predictions of the restricted form of reference-dependent theory, combined with NLIB; they indicate relatively strong loss aversion in both money and consumption goods; and they show no evidence of a systematic difference between incoming and outgoing valuations.

One possible interpretation of the differences between these two bodies of data is that they reflect systematic differences between subject pools. Our collaborative experiment and the Norwich experiment used British students as subjects, while the other experiments used North American students. However, even within the North American data there is a considerable degree of variation between experiments – variation that cannot be explained merely as the result of random factors, given the hypothesis of a common North American subject pool. It seems that we can identify a range of putative effects – loss aversion in money as mediated by NLIB, loss aversion in money as mediated by the common endowment hypothesis, loss aversion in goods, tactical heuristics – none of which is wholly robust in isolation, but each of which may sometimes contribute to the classic and highly reliable WTA/WTP disparity. Of course, it would be much more satisfying to be able to say of some of these putative effects that they have been shown to be entirely robust, and of others that they have been shown to be non-existent. But the evidence is the evidence.

## 11. Conclusions

The question of whether money spent to buy goods is perceived as a loss is a significant issue in the developing theory of reference-dependent preferences. In the context of cost-benefit analysis and contingent valuation, the corresponding question is whether loss aversion in money, in addition to loss aversion in specific goods, contributes to observed disparities between willingness-to-pay (WTP) and willingness-to-accept (WTA) valuations of goods.

The experiment reported in this paper was conducted as an adversarial collaboration between Kahneman, who predicted that money outlays in buying tasks would not be perceived as losses, and the other authors, who predicted that they would. Both parties agree that the evidence *from this experiment* favours the latter prediction, although not decisively so. However, in the light of conflicting evidence from other related experiments, the role of loss aversion in relation to buying tasks must be considered as still an open question, on which more research is needed.

In terms of scientific method, we believe that our work has demonstrated the value of adversarial collaboration in experimental economics. While we do not fully agree about how best to interpret our findings, we have gone a long way in narrowing down the areas of disagreement. We recommend this method to other experimental researchers as a constructive way of resolving conflicts between rival hypotheses.

## References

- 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: 479-505.
- Cubitt, R., Starmer, C., and Sugden, R. (2001). Discovered preferences and the experimental evidence of violations of expected utility theory. *Journal of Economic Methodology*, 8: 385-414.
- 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.
- Kahneman, D. and Novemsky, N. (2002). Loss aversion in risky and riskless transactions: choices, exchanges and gambles. Mimeo, Princeton University.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47: 263-291.
- Loewenstein, G. and Adler, D. (1995). A bias in the prediction of tastes. *Economic Journal* 105: 929-937.
- Mellers, B., Hertwig, R. and Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science* 12: 269-275.
- Munro, A. and Sugden, R. (2003). On the theory of reference-dependent preferences. *Journal of Economic Behavior and Organization* 50: 407-428.
- Myagkov, M. and Plott, C.R. (1997). Exchange economies and loss exposure: experiments exploring prospect theory and competitive equilibria in market environments. *American Economic Review* 87: 801-828.
- Quiggin, J. (1993). *Generalised Expected Utility Theory: the Rank-Dependent Model*. Kluwer: Boston, Massachusetts.
- Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* 38: 332-382.
- Sugden, R. (1999). Alternatives to the neo-classical theory of choice. In Ian J. Bateman and Kenneth G. Willis (eds), *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries*. Oxford University Press: Oxford.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1: 39-60.
- Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: a reference-dependent model. *Quarterly Journal of Economics* 106: 1039-1061.
- Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5: 297-32.