Journal of Economic Methodology 12:2, 265–276 June 2005

Monetary incentives, what are they good for?

Daniel Read

Abstract This paper is a critical reflection on the use of monetary incentives in economic experiments. The argument is that incentives have their effect through their influence on one or more of three factors: (1) *cognitive exertion*; (2) *motivational focus*; (3) *emotional triggers*. I suggest that these effects can often be achieved without monetary incentives, and incentives are not even guaranteed to achieve those effects. There are also disadvantages to requiring the use of incentives in experiments. The paper concludes by suggesting there is no basis for *requiring* the use of real incentives to do experimental economics.

Keywords: experimental methods, incentives

1 INTRODUCTION

> *Sam Spade:* Ten thousand? We were talking about a lot more money than this.
> *Kasper Gutman:* Yes, sir, we were, but this is genuine coin of the realm. With a dollar of this, you can buy ten dollars of talk. (*The Maltese Falcon*, script by John Huston, from a novel by Dashiell Hammett)

Judgement and decision-making (JDM) research is characterised by a focus on 'anomalies and biases,' meaning systematic deviations from normative standards of rationality. Well-documented deviations include conservatism and overconfidence, the failure to integrate over multiple outcomes, framing effects, and paradoxes like those of Allais and Ellsberg (for a review, see Rabin 1998).

Most experimental demonstrations of these anomalies have been based on tasks with hypothetical consequences. Researchers justify the use of these hypothetical tasks through certain assumptions about the abilities and desires of the respondents: They assume that people are capable of reporting what they believe or what they would do under the specified circumstances, that they want to accurately report this information, and that this desire is strong enough to make them do the task to the best of their ability. In the

Journal of Economic Methodology ISSN 1350-178X print/ISSN 1469-9427 online

DOI: 10.1080/13501780500086180

words of Kahneman and Tversky (1979), who pioneered this method, 'The use of this method relies on the assumption that people often know how they would behave in actual situations of choice, and on the further assumption that the subjects have no special reason to disguise their true preferences' (264).

Not everyone accepts these assumptions. Many believe that unless subjects are offered an incentive compatible payment schedule, their responses will not represent what they would do if given the task 'for real' (Wilcox 1993; Harrison 1994). In a widely discussed recent paper, for instance, Hertwig and Ortmann (2001) suggested that even psychologists should use incentives whenever possible, because 'the benefits of being able to run many studies do not outweigh the costs of generating results of questionable reliability' (394). Since hypothetical tasks form the very foundation of much of the experimental research relevant to economics, this suggestion is potentially very serious.

Because of the importance of this issue, many experiments have been conducted that compare the results from offering real and hypothetical monetary incentives. Collectively, these studies do not give overwhelming support to the claim that non-incentivised studies are unreliable, although they do suggest that using incentives has an effect. Although incentives never eliminate anomalies, for instance, they are more likely to decrease than to increase them (Camerer and Hogarth 1999; Hertwig and Ortmann 2001). Moreover, people are typically willing to pay less for almost anything if the money is real than if it is hypothetical (List and Gallet 2001), and real incentives also reduce response variance (Smith and Walker 1993; Camerer and Hogarth 1999).

The purpose of this paper is not to question whether incentives work, but to consider the more fundamental questions of *how* they work and, relatedly, what are the costs of using them in experiments. My view is that monetary incentives are not an experimental magic bullet. They are one part of the experimentalist's arsenal, to be deployed when it is the best way to achieve a desired effect. Before deciding to use incentives, however, it is necessary to think about why the incentive is likely to have its effect, and what are the alternative non-incentivised methods.

2 HOW INCENTIVES WORK

Although monetary incentives can affect behaviour in diverse ways, these all stem from their influence on one of three factors. These are, 'cognitive exertion', the incentive increases the amount of thought put into the task; 'motivational focus', the incentive changes the agent's goals; or 'emotional triggers', the incentive is a prerequisite for the agent to predict or emit their response.

2.1 Cognitive exertion

It is well known that people's performance often fails to capture their competence. One reason might be that they are unwilling to 'do their best' when they are insufficiently compensated for doing so (Wilcox 1993; Harrison 1994). For example, participants in a prisoner's dilemma game might cooperate, not because they want to be nice or believe it will lead others to be nice, but because they do not immediately recognise the dominance structure of the game, and, crucially, they are not motivated to discover that structure. But if they are rewarded for thinking about it, they will discover that structure and then defect. Consistent with this suggestion, there is evidence that people spend more time thinking about what they are doing when there is money on the table (Wilcox 1993).

However, if we are paying people merely because we want them to think harder then incentives should be treated no differently than other ways of achieving this. For instance, incentives are not necessarily better than giving them extra time, writing better instructions, or making them 'accountable' for their choices (Tetlock and Lerner 1999). Moreover, if our goal is to discover whether people conform to normative principles once they 'understand' the task, rather than rely on them to come to that understanding themselves, it might be more efficient to give them the needed information and find out how they behave. Shafir and Tversky (1992), for example, did this when they told participants in a prisoner's dilemma experiment that their opponent had either defected or cooperated – once they knew this, participants were more likely to defect regardless of which decision had been made. This suggests that at least one reason they cooperated was a mistaken belief that this choice could induce their opponent to reciprocate.

Moreover, even if people give different responses when incentives are provided than when they are not, we have not necessarily demonstrated that the incentives have made them more rational. We first have to isolate the locus of the incentive effect – how has the extra thinking changed the problem? As an example of this, consider Harrison's (1994) demonstration that evidence for the Allais paradox is weaker when real money is on the table. The Allais paradox is the following: Most people will take £1M for sure over a gamble offering a 0.89 chance of £1M and a 0.10 chance of £5M; but most people will also take a 0.10 chance of £5M over a 0.11 chance of £1M. This seems to imply that, contrary to expected utility theory, probabilities receive a non-linear decision weight. That is:

$$\pi(.10)u(5M) < (\pi(1.00) - \pi(.89))u(1M)$$

$$\pi(.10)u(5M) > \pi(.11)u(1M),$$

where $\pi(p)$ is the decision weight of probability p, and $u(x)$ is the utility of x. Notice that the decision weight of the difference between probabilities of 1.00 and 0.89 is greater than that of 0.11 (implying a *certainty effect*). When

Harrison introduced (small) monetary incentives, the sign of the top inequality often changed, thus reducing the Allais paradox. One way to interpret this is that when real money is at stake, probability weighting functions become more linear, and thus they more closely conform to the ideal of expected-utility maximisation. Another interpretation, however, is that monetary incentives may have increased the value of the larger amount (FF 5*M* in Allais' example, $20 in Harrison's) relative to that of the smaller amount. If we want to know if people genuinely hold Allais-style preferences, there are alternative methods available, such as asking respondents what they think is normatively correct. This method was advocated by Savage (1954), and used by Slovic and Tversky (1974). They found that even when people were confronted with reasons for believing their Allais choices were irrational, they persisted in their preferences.[1] It is not a straightforward matter to argue, in the face of these results, that a reduction in the frequency of Allais choices shows that people do not hold these preferences.

Of course, incentives might do more than push people to do more of what they are doing already; it can also change *what* they do. The next sections discuss the two ways that they can do this.

2.2 Motivational focus

A commonly raised objection to the use of financial incentives to increase cognitive exertion is that people already have strong *intrinsic* motivations to do their best. By paying people we replace these with extrinsic motives, and may even reduce motivation strength (Lepper *et al.* 1973). While this is often described as a negative outcome, it might be exactly what the experimenter wants. Intrinsic motivation is beyond the experimenter's control and, moreover, actions that maximise intrinsic payoffs may not be the ones that would maximise economic payoffs. To illustrate, when making judgments of probability, people may enjoy stating they are certain, even when they believe there is some chance they are wrong. But if inappropriate certitude will lose money, they will only say they are certain when they actually are. The extrinsic monetary incentive does not necessarily increase the *level* of motivation, but it might change its *focus*, thus making observed behaviour easier to interpret.

We can state this in more formal terms. A monetary incentive is one of many sources of utility, amongst which we might number entertainment, altruism, self-esteem, and the satisfaction of primary and secondary drives. The decision-maker's goals can be summarised as follows:

$$
\begin{aligned}
&Max\ u(\$, EE)\\
&s.t.\\
&\$ = f(choice): \quad \text{Money}\\
&EE = f(choice): \quad \text{Everything else}
\end{aligned}
$$

By introducing monetary incentives, the weight put on money is increased. Ideally, the result is a single argument utility function $u(\$)$, so that if people are maximising their earnings they are also maximising their utility. One goal of using financial incentives, therefore, is to 'crowd out' other incentives.

The desirability of such crowding out depends on what we want to discover. Consider, for instance, responses to the dictator game, in which one person divides up a sum of money, taking one part for him or herself, and giving the rest to a second person. Because the second person has no say in the matter, the dictator's profit maximising choice is to keep all the money. The game is extraordinarily simple, and incentives are certainly not needed to induce people to exert sufficient effort to understand the game. Yet participants in dictator games do keep more (although not all) of the money for themselves when the money is real rather than hypothetical (Forsythe *et al*. 1994). This shows that real money reduces the contribution of whatever motivates people to give some of the money away. But how important is this result, and others like them? Does not this finding merely illustrate the fact that if people are paid to do things (such as, in this case, keep more money for themselves), they will be more likely to do them, or to do more of them. But no experiment was needed to show this. I suggest that the interesting result in many studies is not that *money* can motivate decision-makers, but that non-monetary motives, such as altruism, can also motivate them.

Ultimatum games are like dictator games, except that now the recipient can refuse the division, in which case no one gets anything. As in dictator games even when the stakes are pretty large ($100) people offer the same proportion of that stake as when they are small ($10; Hoffman *et al*. 1976), or even hypothetical (Forsythe *et al*. 1974). Does this tell us that the altruism motive grows in proportion to the stakes? Not at all. Again, it is hard to figure out exactly what it teaches us. When there is more money on the table there is a greater financial temptation to make a low offer, but at the same time it costs more to have the offer rejected, and the net effect of these countervailing factors appears similar up to the high $100 stake. We might expect the relationship to break down eventually. If you have £1 million to divide you can be pretty sure the offers will get proportionally smaller, because the probability of rejection will also become small. Varying incentives, along with careful research design, might inform us about the specific values needed to elicit proportionally low offers, but even a thought experiment should convince us there will be *some* such value.

Despite their limitations, however, it is clear that using incentives to increase the motivational weight put on money has an important research role. If we hypothesise that non-economic motives operate *only* when there is no money on the table, then the natural way to test this is by comparing incentive with no-incentive conditions. But once this hypothesis has been

ruled out, there is no obvious further benefit from *partially* crowding out non-monetary motives, especially if those non-monetary motives are what we are interested in studying.

2.3 Emotional trigger

Many phenomena in the behavioural sciences, such as response-time and memory capacity, cannot be studied by asking people to predict what they would do, but only by observing them. Such observation is necessary when (i) the response is not completely under control (meaning, the agent cannot produce the response in the absence of the stimulus), and (ii) the agent cannot accurately predict the response. To illustrate, imagine conducting a study into how people react by having their thumb hit with a hammer. Condition (i) is clearly met, because pain will not be felt merely by imagining the blow. But this does not, by itself, mean that someone has to be hit with a hammer to find out how they would respond. They might be able to state how intense the pain would be without being struck (perhaps from memory or mental simulation), or to state correctly that the first words out of their mouth will be '@#&!' Condition (ii) means they cannot report what their response will be in the absence of the real stimulation. Research into such 'visceral' effects shows that such circumstances do occur and may even be common (Loewenstein *et al.* 2003; Read and Loewenstein 1999).

Real money can play a similar role to the hammer blow, by putting people in a state they cannot imagine without experiencing it. Consider, as an illustration, the common fictional scenario in which intimate friends or family are entrusted with a huge amount of semi-legitimate money, and must decide how to manage their windfall. Disagreements arise, things get out of control, and betrayal and murder ensue.[2] I suspect that had these people been allowed to predict whether, given a large monetary incentive, they would murder their close friends or kin they would say 'no'. The incentive effect can neither be experienced nor predicted based on a hypothetical scenario.

Do the monetary incentives offered in the laboratory also act as emotional triggers? To show that they do, we need to demonstrate not only that there is a difference between responses when faced with real or hypothetical incentives, but also that people are unable to predict their responses to the real incentives. Meeting the second condition is the biggest challenge, and few researchers have, as yet, tried to meet it. One paper that claims to do so is Holt and Laury's (2002) experimental study of the finding that risk aversion increases the larger the amount at stake (Kachelmeier and Shehata, 1992). Holt and Laury looked for this risk-aversion/magnitude effect in both a real and hypothetical payoff condition, and observed it only in their real-payoff condition. They argued that, corresponding to the second condition above, that 'subjects facing hypothetical choices cannot

imagine how they would actually behave under high incentive conditions' (1154). The crucial word in that sentence is 'cannot'. That is, the effect cannot be reasonably attributed to other differences between conditions, such as differences in the cognitive exertion elicited. Yet in this specific study, it is not clear this is true. Perhaps for legal or ethical reasons, Holt and Laury underplayed the hypothetical conditions of their study:

> Even though the earnings from this next choice are larger, they are only hypothetical, and we would like for you to initial the statement at the top of the page indicating that you understand you will not be paid for the choice on that page. All other choices tonight will count towards your earnings. Let me reiterate that your choice in this part has no effect on your earnings and has no effect on what choices will be given to you subsequently; but we are interested in what you would do if you actually faced these choices, so please think about them carefully.

It is clear that these instructions might not induce careful thinking about the hypothetical question. Indeed, consistent with a 'method variance' interpretation of Holt and Laury's results, another study published in the same year (Kühberger *et al*. 2002) replicated the risk-aversion/magnitude effect using both real and hypothetical payoffs. Therefore, the effect of payoff magnitude on risk aversion does not appear to be due to an emotional trigger effect.

Perhaps the best candidate for a laboratory-induced emotional trigger effect is not found when experimenters pay money to subjects, but when subjects pay money to experimenters. Consider, for instance, the distinction between real versus hypothetical willingness-to-pay (reviewed in List and Gallet 2001). In one such study (Cummings *et al*. 1995), participants were first asked in hypothetical terms if they were willing to pay a certain amount for common objects (e.g., juicer, calculator) and then given the chance to really do so. The proportion who agreed to make the purchase plummeted when it became real. Cummings *et al*. (1995) took great care to ensure their participants had every opportunity to think of their hypothetical choices as real, using the following instructions:

> Suppose that you could buy one of these calculators for $3. It would be important that you realise that if you say YES (you are willing to pay $3) you have to pay for it with your own money. You would pay $3 and take a calculator with you. Cash or check would be accepted, but no credit cards.

Despite this instruction, participants were much less likely to agree to pay $3 for a calculator (or anything else) in the real-choice condition. It seems likely that if faced with a hypothetical buying situation, no matter how well we rehearse it, we cannot know precisely what we will do when making the real choice.

The potential loss of real money, therefore, probably can elicit a visceral reaction that can neither be experienced nor predicted in its absence. But even here we can overstate what we can learn from experiments using real choices. If we are interested in *qualitative* differences between conditions, then even if the non-incentivised condition is a diluted version of the incentivised one, it might still reflect the relevant qualitative effects. To illustrate, if we want to know whether people will really pay more for a toaster than for a calculator, a hypothetical WTP may be just as valid as a real one. But if we want to know exactly how much they would pay, we need a real WTP. For each research question we ask, we should first determine whether it is more like the first situation or the second.

3 THE NON-MONETARY COSTS OF MONETARY INCENTIVES

So far, I have argued that 'incentives are not a panacea'. They are usually neither necessary nor sufficient for achieving our research goals. But there are costs to offering monetary incentives, which we must balance against their benefits. The major cost is that the kind of research we can do with monetary incentives is greatly restricted, and their use will influence both what questions we set out to answer, and the way we undertake to answer them. Specifically, the tasks studied and the outcomes involved in incentivised studies can fail to be either realistic or relevant.

One problem is that of incentive size. The economics of paying large numbers of participants means that experimenters can rarely afford to pay each individual very much, and so only trivial incentives are offered per choice. This is hardly satisfying if incentives are being used to help us emulate real-world decision-making, or to understand phenomena that are intensified when outcome magnitude is high. With regards to the latter, we might consider Harrison's (1994) study of the Allais paradox. Recall that he found the number of Allais-type choices was reduced when there was real money on the table. In Harrison's study the amounts were \$1, \$20 and \$0, whereas in the original example from Allais (1953) they were FF 100 million, FF 500 million and FF 0 – in 1950s money! Although we cannot (in the absence of real incentives) know this with certainty, it is not at all unlikely that with real millions on the table the incentives would increase the frequency of the paradoxical effect – it might be easier to risk losing \$20, but not to risk losing \$10 million. The small real amounts might be less diagnostic than the gigantic large amounts.

One common way to overcome the incentive size problem is to make the rewards probabilistic (Wilcox 1993). For instance, we give one participant a reward based on a single choice – the participant and the judgment to be chosen randomly after the study is finished. This allows the experimenter to offer large nominal amounts as incentives. It is hoped that this enables us to

obtain, at modest cost, the advantages of large hypothetical payoffs with real incentives. But probabilistic payoffs are not much of a solution. First, they do not increase task realism in any way, and can even be said to decrease it – I know of no situations outside of the laboratory and the national lottery offering such a reward schedule.

Perhaps even more importantly, however, probabilistic incentives are almost always intended to serve as stand-ins for certain ones.[3] Yet this can only be justified by assuming that participants can ignore the real task contingencies, and act 'as if' each choice was going to be rewarded with certainty. But if they can do that, then surely they can just as easily act 'as if' hypothetical incentives were real. Indeed, if we did observe a difference between conditions, it would be impossible to justify the belief that probabilistic-real incentives were better than certain-hypothetical incentives at predicting behaviour when faced with certain-real incentives.[4]

The problems just cited are specific cases of a more general problem – the frequent lack of realism in incentivised tasks. One other case of unrealism is that while real-world decisions invariably involve trade-offs between real losses as well as gains, the use of incentives in the laboratory rarely allows the imposition of real losses. Therefore, we are forced to make trade-offs between the *realism* and the *reality* of incentives. We can easily ask about millions of hypothetical dollars, but not about millions of real dollars; and we can easily ask people if they would pay hypothetical money, but (rarely) to really pay. It may be that real-world behaviour is better exemplified by realistic hypothetical choices than unrealistic real ones.

Using incentives can also commit us to complex and cumbersome research designs that get in the way of the questions we want to answer. Monetary incentives often (although not always) introduce new problems that might be more trouble than they are worth. For example, they can add to the intellectual effort required from the participant, often with a corresponding reduction in the realism of the task. One example is the use of a proper scoring rule to reward well-calibrated probability judgement.[5] The injunction to 'be accurate' is replaced with complex formula or payoff table assigning payoffs to different probability judgments given different outcomes. I suspect that participants either more-or-less ignore the rule, or else get so caught up in understanding it that this becomes the focus of their activity. Likewise, it is very difficult to explain a second-price auction (Kagel and Levin, 1993) or a Becker-DeGroot-Marschak price mechanism (Becker *et al.* 1964), and it is unlikely that participants who have been carefully taught will then give responses like they would give in their everyday life.

Incentives can also transform a task into something other than what we want to study. Consider, for example, studies in which many choices are made about different gambles. If all the gambles are to be played, respondents *should* make a 'portfolio' decisions about them, and consequently be more risk-neutral than they would otherwise be (Redelmeier and Tversky

1992). To get around this we must either ask each participant a single question, hence sharply reducing sample size, or use a probabilistic procedure which, as just discussed, has its own problems. With hypothetical choices, the portfolio problem is a much smaller threat literally because participants have no incentive to bundle the gambles together – after all, they are not going to play any of them for real.

A final problem with using incentives is that it involves a high cost in research efficiency. A simple questionnaire can be administered to a hundred or so people in an afternoon, a carefully conducted experiment with monetary incentives might take a week or longer.[6] If the results from the two studies are the same, as they are likely to be, the use of real incentives has produced nothing of scientific value (except, perhaps, reassurance). Even if the results are different, and if the incentivised study gives 'cleaner' or 'more rational' results, it is still an open question whether it was worth offering monetary incentives. In Ortmann and Hertwig's calculus, 'the benefits of being able to run many studies do not outweigh the costs of generating results of questionable reliability'. But if incentivised studies merely contain less error (and the reviews of Hogarth and Camerer, and Ortmann and Hertwig both suggest this) then it is not an easy matter to compute at what point the benefit from doing more studies outweighs that from doing marginally better ones.

4 CONCLUSION

The primary message of this paper is that the use of monetary incentives is one of the tools researchers have at their disposal, and the decision to employ them, just like any other research decision, should be made based on what the incentives will achieve. The best research strategy is to first ask 'what are the ideal studies for investigating this problem?' If these studies can easily be conducted using real monetary incentives, and if the money is available, then there is every reason to do so: there are some advantages to the use of incentives, and so the incentivised design dominates the non-incentivised one. But usually the ideal studies cannot be done, and so the experimenter will have to compromise and choose the best achievable study. This compromise will usually involve a trade-off between realism and reality. Because monetary incentives impose a constraint on what is achievable, it is likely that in many cases the best possible study will be a hypothetical one that relies on participants being able to report what they would do if placed in a situation that cannot be precisely reproduced in the laboratory.

Daniel Read
University of Durham
Daniel.read2@durham.ac.uk

NOTES

1 My own experience is in line with this. I have taught the Allais paradox in many classes, and I have never met a group that did not violently resist the idea that it was in any way a paradox. Even when the logic is clear, the preference remains.
2 *Greed*; *Nostromo*; *Shallow Grave*; and *A Simple Plan* all provide some variations of this theme of unexpected wealth leading people to violate well-planned principles.
3 When future outcomes are uncertain many classic anomalies in intertemporal choice are reduced or even disappear (Keren and Roelofsma 1996).
4 In fact, what studies have been done suggest that real-probabilistic, real-certain and hypothetical-certain payments do yield approximately the same results (Starmer and Sugden 1991; Beattie and Loomes 1997; Cubitt *et al.* 1998).
5 The quadratic scoring rule is the most widely used proper scoring rule. The payoff $Q_j(p)$ is given by:

$$Q_j(p) = a + 2bp_j - b\sum_{i=1}^{n} p_i^2$$

Where p_j is the probability judgement, p_i is the probability the event occurs, and a and b are payoff parameters chosen by the experimenter. Using this rule, a respondent who truthfully reveals their beliefs will maximise the expected value of their payoffs.
6 It is not merely the cost of incentives, but the logistics involved in making and recording payoffs can be burdensome.

REFERENCES

Allais, M. (1953) 'Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine', *Econometrica* 21: 503–46.

Beattie, J. and Loomes, G. (1997) 'The impact of incentives upon risky choice experiments', *Journal of Risk and Uncertainty* 14: 155–68.

Becker, G.M., DeGroot, M.H. and Marschak, J. (1964) 'Measuring utility by a single-response sequential method', *Behahvioral Science* 9: 226–32.

Camerer, C.F. and Hogarth, R.M. (1999) 'The effects of financial incentives in experiments: a review and capital-labor-production framework', *Journal of Risk and Uncertainty* 19: 7–42.

Castellan, N.J. (1969) 'Effect of change of payoff in probability learning', *Journal of Experimental Psychology* 79: 178–82.

Cubitt, R., Starmer, C. and Sugden, R. (1998) 'On the validity of the random lottery incentive system', *Experimental Economics* 1: 115–38.

Cummings, R.G., Harrison, G.W. and Rutstrom, E.E. (1995) 'Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive-compatible?', *American Economic Review* 85: 260–66.

Forsythe, R., Horowitz, J.L., Savin, N.E. and Sefton, M. (1994) 'Fairness in simple bargaining experiments', *Games and Economic Behavior* 6: 347–69.

Grether, D.M. and Plott, C.R. (1979) 'Economic theory of choice and the preference reversal phenomenon', *American Economic Review* 69: 623–38.

Harrison, G.W. (1994) 'Expected utility theory and the experimentalists', *Empirical Economics* 19: 223–53.

Hertwig, R. and Ortmann, A. (2001) 'Experimental practices in economics: a methodological challenge for psychologists?', *Behavioral and Brain Sciences* 24: 433–51.

Hoffman, E., McCabe, K. and Smith, V. (1996) 'On expectations and monetary stakes in ultimatum games', *International Journal of Game Theory* 25: 289–301.

Holt, C.A. and Laury, S.K. (2002) 'Risk aversion and incentive effects', *American Economic Review* 92: 1644–55.

Kachelmeier, S.J. and Shehata, M. (1992) 'Examining risk preferences under high monetary incentives: experimental evidence from the People's Republic of China,', *American Economic Review* 82: 894–920.

Kahneman, D. and Tversky, A. (1979) 'Prospect theory: an analysis of decision under risk,', *Econometrica* 47: 263–91.

Kagel, J. and Levin, D. (1993) 'Independent private value auctions: bidder behavior in first-, second- and third price auctions with varying numbers of bidders', *Economic Journal* 103: 868–79.

Keren, G. and Roelofsma, P.H.M.P. (1995) 'Immediacy and certainty in intertemporal choice', *Organizational Behavior and Human Decision Processes* 63(3): 287–97.

Kuhberger, A., Schulte-Mecklenbeck, M. and Perner, J. (2002) 'Framing decisions: hypothetical and real', *Organizational Behavior and Human Decision Processes* 89: 1162–75.

Lepper, M.R., Greene, D. and Nisbett, R.E. (1973) 'Undermining children's intrinsic interest in with extrinsic reward: a test of the ''overjustification'' hypothesis', *Journal of Personality and Social Psychology* 28: 129–37.

List, J.A. and Gallet, C.A. (2001) 'What experimental protocol influence disparities between actual and hypothetical stated values? Evidence from a meta-analysis', *Experimental and Resource Economics* 20: 241–54.

Loewenstein, G., O'Donoghue, T. and Rabin, M. (2003) 'Projection bias in predicting future utility', *The Quarterly Journal of Economics* 118: 1209–48.

Rabin, M. (1998) 'Psychology and economics', *Journal of Economic Literature* 36: 11–46.

Read, D. and Loewenstein, G. (1999) 'Enduring pain for money: decisions based on the perception and memory of pain', *Journal of Behavioral Decision Making* 12: 1–17.

Redelmeier, D. and Tversky, A. (1992) 'On the framing of multiple prospects', *Psychological Science* 3: 191–93.

Savage, L.J. (1954) *The Foundations of Statistics*, New York: Wiley.

Shafir, E. and Tversky, A. (1992) 'Thinking through uncertainty: nonconsequential reasoning and choice', *Cognitive Psychology* 24: 449–74.

Slovic, P. and Tversky, A. (1974) 'Who accepts Savage's axioms?', *Behavioral Science* 19: 368–73.

Smith, V.L. and Walker, J.M. (1993) 'Monetary rewards and decision costs in experimental economics', *Economic Inquiry* 31: 245–61.

Starmer, C. and Sugden, R. (1991) 'Does the random-lottery incentive system elicit true preferences? An experimental investigation', *American Economic Review* 81: 971–78.

Tetlock, P.E. and Lerner, J. (1999) 'The social contingency model: identifying empirical and normative boundary conditions on the error-and-bias portrait of human nature', in S. Chaiken and Y. Trope (eds) *Dual Process Models in Social Psychology*, New York: Guilford Press.

Wilcox, N. (1993) 'Lottery choice, incentives, complexity and decision time', *Economic Journal* 103: 1397–417.