How Do People Choose?

An Experimental Investigation of Models of 'Sub-optimal' Decision Making.

Nuttaporn Rochanahastin

PhD

The University of York

Economics and Related Studies

July 2018

Abstract

This research experimentally tests some recent theories of 'sub-optimal' behaviour in individual decision making. The first chapter experimentally tests a theory by Manski (2017), addressed to explaining 'satisficing' behaviour. He addresses two key questions: *when* should the decision-maker (DM) satisfice?; and *how* should the DM satisfice? The theoretical results are simple and intuitive; we have tested them experimentally. Our results show that some of his propositions (those relating to the 'how') appear to be empirically valid while others (those relating to the 'when') are less so. The second chapter tests two 'limited attention' theories, namely, those of Masatlioglu *et al* (2012), and Lleras *et al* (2017). These theories are built upon axioms which are weakenings of WARP and are experimentally testable using standard choice data. We found that one weakening is a more plausible weakening of WARP than the other. The third chapter involves the concept of salience. Leland and Schneider (2016) proposes axioms of salience perception. We experimentally test these. We also test the implications of these axioms for risky choice as encapsulated in their SWUP model. The results show general support for the axioms; while those from the implication section show some support for the CARA SWUP model, in that, for the majority of the subjects, SWUP fitted better than EU.

Contents

Lists of Figures

List of Tables

Acknowledgements

First and most importantly, I would like to express my deepest appreciation to my supervisor, Professor John D Hey. His invaluable advice and guidance helped me through every single step of completing this thesis. Since my first day here at York, he always ensure that I have everything I needed for my research. I could not have asked for a better advisor and mentor for my Ph.D study. However, I find that the most merit aspects I learn from him are not applicable just only in academics but also in everyday life. I wish I could be as passionate, attentive, honest, and respectful as him.

I also would like to thank Mr. John Bone and Dr. Martin Forster for their useful comments under their role in the Thesis Advisory Panel. I am grateful to Professor Itzhak Gilboa and Dr. Dominic Spengler for serving on my examination committee. This thesis could not be completed had I only worked in front of a computer screen. I am indebted to everyone involved in my long but fruitful journey here in York. I thank you all.

Last but not the least, I would like to thank my family: my father, my brother and sister for supporting me unconditionally, throughout the good and not-so-good time spent writing this thesis and all my life. And, to my mother in heaven, this one is for you.

Declaration

I, Nuttaporn Rochanahastin, declare that this thesis titled, ’How do people choose? An experimental investigation of models of ‘sub-optimal’ decision making.’ and the work presented in it are either my own or my joint work. The first chapter is joint work with my supervisor, Prof. John Hey, and my collegue, Yudistira Permana. The version presented in this thesis is the same version which has been published in Theory and Decision (doi: 10.1007/s11238-017-9600-5). There are slight amendments of notation and numbering in order to be consistent with the rest of the thesis. The experimental interface programming for this experiment was done by Paolo Crosetto while calculations and simulations are our own work.

The second chapter is my own with extensive guiding from and discussions with Prof. John Hey. It is currently under review at Journal of Behavioral and Experimental Economics. The third chapter is developed jointly with Prof. John Hey. There is an alternative version of this paper where we combined it with a theoretical work and jointly submitted with Mark Schneider and Jonathan Leland. The combined version is under review at Management Science. All programming in the second and third chapter is my own.

I also confirm that:

- This work was done wholly while in candidature for a research degree at this University. This work has not been submitted at any other institute for any other award.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given.
- I have acknowledged all main sources of help.

- Where the thesis is based on work done by myself jointly with others, I have made reference to that.

Introduction

This thesis consists of three distinct chapters. The over-riding theme of the thesis is the experimental investigation of some recent theories of 'sub-optimal' behaviour in individual decision making. Traditional economic theory assumes that a decision maker (DM) is fully rational. By this, it means that the DM must have (near) complete relevant information as well as powerful cognitive power in order to achieve global optimisation. For example in a choice context, a DM must have a complete preference ordering then he/she must be able to deliberate all available options and choose consistently according to those ordering. However, as postulated in Simon (1955), DMs face constraints as to limits to the amount of information one can possess or limits on computational ability. Empirical evidences[1] suggest that these lead to departures from complete rationality and/or a shortfall from optimal outcomes. A real life example would be when someone performs a Google search, most likely he/she will not look beyond the first couple of pages of the results. One might argue that the costs involved in coming into decisions are already implicitly incorporated in economic models. However, unlike any other costs that can be easily incorporated as constraints, deliberation cost or the cost of thinking has a special characteristic that prevents itself from entering into conventional optimization problem. Savage (1954) was the first to mention informally that an attempt to consider these costs as an act that the people must decide will lead to an endless regression. He states:

'It might ... be stimulating, and it is certainly more realistic, to think of consideration or calculation as itself an act on which the person must decide. Though I have not ex-

[1] See Grether and Plott (1979) and Caplin *et al* (2011), for example.

plored the latter possibility carefully, I sus- ect that any attempt to do so leads to fruit-less and endless regression.' (Savage (1954), p. 30).

If there is a cost of thinking then there must also be a cost to think whether it is worth thinking and so on. This infinite regress problem prevents optimization from being a logical model from the behavioural point of view[2]. Therefore, there could be other models that human actually use. One plausibal hypothesis, which is being investigated here, is that DMs simply use heuristics or behavioural rules.

With respect to decision theory, Subjective Expected Utility Theory (Savage (1954)) is considered as the standard, and normatively appealing, theory of rational choice. The theory suggests that choices are made under the assumption of a fixed set of alternatives with probability or belief of each outcome or alternative formed subjectively. These underlying assumptions are simple and elegant. However, they might not be true empirically. Another example is the Weak Axiom of Revealed Preference (WARP) (Samuelson (1938)) which is regarded as the foundation stone of preference theory. Its inference method is sound empirically in that the preferences are inferred from observed choices. Again, it receives much empirical criticism possibly because it also operates under the full rationality assumption. Hence, the descriptive validity of these theories is not clear. A recent development of theoretical research has been aimed at describing or addressing 'sub-optimal' or 'boundedly rational' behaviour by incorporating different heuristics or biases. Most of the experimental works that test normative models do not perform particularly well[3] as they observe 'what' people choose based on theories that internally assume complete consistency. The failure of these theories according to experimental results could stem from the lack of a deliberation process in the models. This thesis attempt to shed some light on how behavioural heuristics might play a role in bridging the gap between rational theories and actual behaviour. Each chapter of the thesis investigates a 'sub-optimal' theory,

---

[2] I refer the reader to Conlisk (1966) for the detailed discussion regarding this problem.
[3] See the results of Hey and Pace (2014) and Bone et al. (2009), for example.

concentrating particularly on recently-proposed areas that include the question on 'how' do people come to make a decision.

The first chapter experimentally tests a new theory by Manski (2017), addressed to explaining 'satisficing' behaviour. Satisficing occurs when the decision-maker (DM) does not go for the 'best' option, but is satisfied with something less. Rather tautologically, this is when decision-makers are satisfied with achieving some objective, rather than in obtaining the best outcome. The term was coined by Herbert Simon in 1955, and has stimulated many discussions and theories. Prominent amongst these theories are models of incomplete preferences, models of behaviour under ambiguity, theories of rational inattention, and search theories. However, all seem to lack an answer to at least one of two key questions: *when* should the DM satisfice?; and *how* should the DM satisfice? In a sense, search models answer the latter question (in that the theory tells the DM when to stop searching), but not the former; moreover, usually the question as to whether any search at all is justified is left to a footnote. Manski addresses these questions by setting the decision problem in an ambiguous situation (so that probabilities do not exist, and many preference functionals can therefore not be applied) and by using the Minimax Regret criterion as the preference functional. The theoretical results are simple and intuitive. Deliberation costs play a central role. 'Optimising' or 'Satisficing' will be the decision if their respective associated cost is low enough. If both costs are sufficiently large then 'No Deliberation' will be preferred. The theory also suggests what is the level of the threshold level of satisficing (aspiration level) should be. We have tested these propositions experimentally. Our results show that some of them (those relating to the 'how') appear to be empirically valid while others (those relating to the 'when') are less so. Subjects do not follow the theory in term of what strategy to choose. However, when they chose to satisfice, the aspiration level is close to theory prediction that is, is half way between the relevant upper and lower bounds of the payoffs. Choosing a strategy is a particularly difficult task but choosing an aspiration level is less difficult and is more intuitive. These results may not be surprising. The more straightforward bit of the theory is more likely to be empirically accurate.

The second chapter investigates one of the plausible alternative explanations for the assumption of complete and exhaustive deliberation process that has recently received attention. The motivating example is that it is unlikely that people go through every item in the supermarket shelves or go through every result of their Google search. Instead, DMs have limited attention. We employ an experimental procedure similar to that of Manzini and Mariotti (2010) to test and compare two 'limited attention' theories, namely, those of Masatlioglu *et al* (2012), and Lleras *et al* (2017). These theories are built upon axioms which are weakenings of WARP. The validity of the predictions coming out of these theories depends upon the validity of the underlying axioms. This paper uses standard choice data to determine the (relative) violation rate of their underlying axioms. We observe the number of actual violations of the axiom underlying each theory and compare them with a benchmark so that we can penalise them for different degree of restrictions. A 'benchmark' was derived from simulations of random behaviour. We found that Lleras *et al* is the more restricted version when compared to Masatlioglu *et al*. Its ability to extract preferences is higher. Masatlioglu *et al* seems to perform the best in the consistencies analyses which is the main observation for the axiom violations and appears to be the empirically more plausible weakening of WARP using this criteria.

The third chapter involves the concept of salience. Leland and Schneider (2016a) propose three principles to characterise properties of the salience perceptual system. We experimentally test these directly using non-symbolic stimuli. We also test the implications of these axioms for risky choice as encapsulated in their SWUP model on the same experimental participants. SWUP model differs from Expected Utility (EU) theory in that outcomes and probabilities are weighted by their salience. The experiments involved pairwise choice questions which were divided into two sections. The first section tests the axioms directly. The axioms are designed to say when a pair of items (*x,y*) is 'more salient' than another pair (*x',y'*). We had to interpret what this means. Dictionaries define 'salient' as something important or noticeable, or, occasionally, as something very important or very noticeable. Nowhere is 'more salient' defined, but, in the spirit of the dictionary definitions, we take it to mean 'more noticeable'. So, in keeping with this spirit, we devised an experiment to see whether subjects could detect 'more noticeable'. To do so, in each problem subjects were presented with two boxes, each containing red and blue circles. The task implicitly, but not explicitly, was to choose whichever box was more salient according to the axioms. The second section tests the implications of these axioms for risky choice. Subjects were presented with a series of pairwise lottery choices. The results from the first section show general support for the axioms; while those from the second section show some support for the CARA SWUP model, in that, for the majority of the subjects, SWUP fitted better than EU. There is also a modest connection between the violations of the axioms and the violations of the predictions. The finding provides a link between a basic property of salience perception and risk seeking behaviour in long-shot lotteries.

What we have learned from these experiments is that it is likely people employ heuristics somewhere in their decision process. This is part of the reason why optimality or consistency is difficult to achieve. However, there are sceptics, particularly from theorists' points of view, that heuristics are inconsistent and that we could need one model for each type of deviation from optimality. We also find some support for the traditional theories such as WARP in the second chapter and EU in the third chapter. This suggests that DMs are not solely using heuristics but, possibly with learning and more careful deliberation, also conform to consistent models. Hence, the answer for better explanatory model possibly lies between the end of two spectrums in which one extreme is the complete consistency to full rationality and another is heuristics and biases. 'How' people come to make a decision must be taken into account.

Chapter 1

When and How to Satisfice: An Experimental Investigation

1.1 Introduction

This paper is about satisficing behaviour. Way back in 1955 Herbert Simon made a call for a new kind of economics stating that:

> "the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environment in which such organisms exist". (p 99)

There is a fundamental conflict here provoked by the use of the word 'rational', and economists' obsession with it. The problem is that the expression 'rational behaviour' covers virtually all forms of behaviour, as long as it is motivated by some 'rational' objective function, and the decision-maker has all relevant information available to him or to her, and the decision-maker (henceforth, DM) can perform all the necessary calculations costlessly. If calculations are costly, then we are led into the infinite regression problem, first pointed out by Conlisk in 1996, and rational behaviour, as defined by economists, cannot exist. We are therefore constrained to operate with rational models, defined as above. The way forward, within the economics paradigm, is therefore to weaken our ideas of what we mean by rational behaviour. This is the way that economics has been moving. Prominent amongst these latter weaker theories are theories of incomplete preferences (Ok *et al* (2012), Nau (2006), Mandler (2005), Dubra *et al* (2004)); theories of behaviour under ambiguity (Etner *et al* (2012), Gajdos *et al* (2008), Ghirardarto *et al* (2004), Hayashi and Wada (2010), Klibanoff *et al* (2005), Schmeidler (1989) and Siniscalchi (2009)); theories of rational inattention (Sims (2003), Manzini and Mariotti (2014), Matejka and McKay (2015), Caplin and Dean (2015)); and search theories (Masatlioglu and Nakajima (2013), McCall (1970), Morgan and Manning (1985), and Stigler (1961)). A useful survey of satisficing choice procedures can be found in Papi (2012).

Almost definitionally, *models of incomplete preferences* have to be concerned with satisficing: if the DM does not know his or her preferences, it is clearly impossible to find the best action. These models effectively impose satisficing as the only possible strategy. The problem here is that complete predictions of behaviour must also be impossible. Prediction is possible in *models of behaviour under ambiguity*. But here again satisficing behaviour 'must' occur, if only because not all the relevant information is available to the DM. Unless the DM's information is objectively correct, there is presumably always some action that is better than the one chosen by the DM. But here the DM does not choose to satisfice; nor does he or she choose how to satisfice. *Models of rational inattention* also capture the idea of 'satisficing' behaviour – in that choice is made from a subset of the set of possible actions – those which capture the attention of the DM, that is, those which are in the consideration set of the DM. However, these theories are silent on the reasons for the formation of a consideration set, and, in some of them, on how the consideration set is formed.

We examine a new theory – that of Manski (2017) – which might be classified as an extended search model. *Search models* seem to be closest to the scenario in which Manski's paper is set. Standard search models assume that the DM is searching for the highest number in some distribution, and that there is a cost of obtaining a drawing from that distribution. Because of this cost, the DM does not keep on searching until he or she finds the highest number: generally he or she should keep on searching until a 'sufficiently' high number is found. This could be termed the DM's *aspiration level*. One interpretation of Manski's paper is that he generalises the story: in addition to being able to search for numbers greater than some (or several) aspiration level(s), the DM can pay a higher search cost and be able to find the highest number, and also the DM can choose not to indulge in any search and simply receive a lower number. Manski not only considers choice between these three strategies, but also the choice of the aspiration level(s). This is the 'how' of Manski's theory: he explains how many times satisficing should be implemented, how aspiration levels should be formed and how they should be changed in the light of the information received[4].

We experimentally test this new theory. Some of the other models that we have discussed have also been tested experimentally; for incomplete preferences we refer the reader to Cettolin and Riedl (2016), Costa-Gomes *et al* (2014) and Danan and Ziegelmeyer (2006); for behaviour under ambiguity to Abdellaoui *et al* (2011), Ahn *et al* (2010), Halevy (2007), Hey and Pace (2014) and Hey *et al* (2010); for rational inattention to Chetty *et al* (2009), De Los Santos *et al* (2012); and for search theories to Caplin *et al* (2011), De Los Santos *et al* (2012), Hayashi and Wada (2010) and Reutskaja *et al* (2011). Our experimental test has some similarities in common with some of these and some differences. In some ways our test is closest to that of Hayashi and Wada (2010), though they test minimax, α-maximin and the (linear) contraction model

[4] There are echoes of this in Selten (1998), though he notes on page 201 that "In this respect, the role of aspiration levels in [Selten's] model is different from that in the satisficing processes described by Simon, where it is assumed that it can be immediately seen whether an alternative satisfies the aspiration level or not. The situation of the decision maker in [Selten's] model is different."

(Gajdos *et al* (2008)). We test Manski's model and have a different way of generating imprecise information/ambiguity.

In the next section we describe the Manski model, while in section 3 we discuss the experimental design. Our results are in section 4, and section 5 concludes.

1.2 Manski's Model of Satisficing

In the model the DM has to choose some *action*. The DM knows that there is a set of actions, each member of the set implying some payoff. The payoffs of these actions are bounded between a lower bound, *L,* and an upper bound, *U,* which are known to the DM. Hence, without costly deliberation, the DM faces a problem under ambiguity as he or she does not have sufficient knowledge to determine the optimal decision – that of choosing the action which yields the highest payoff. However, the DM can learn more about the payoff values subject to different costs, which in turn, yield different benefits. There are three available deliberation strategies: 'No Deliberation', 'Satisficing', and 'Optimising'. 'No Deliberation' incurs no cost and yields only the value of the payoff of an arbitrarily chosen action. 'Optimising' has a positive cost (*K*) and reveals the maximum payoff value. 'Satisficing' has a positive cost (*k*) and provides information whether there are actions that are at least as large as some specified aspiration level.

Crucial to the model is that the assumed objective of the DM is the minimisation of maximum regret (MMR). One reason for this is that there is no known probability distribution of the payoffs, so, for example Expected Utility theory and its various generalisations cannot be applied[5]. Additionally, and crucially for our experiment, the solution is an *ex ante* solution, saying what the DM *should* plan to do as viewed from

[5] Manski notes that "The maximin criterion gives the uninteresting result that the person should always choose the null option when deliberation is costly."

the beginning of the problem. As Manski writes “I study *ex ante* minimax-regret (MMR) decision making with commitment”. So the DM is perceived of as choosing a strategy at the beginning of the problem, and then implementing it. This implies a *resolute* decision-maker. If the DM is not resolute the solution may not be applicable.

The paper applies the *ex ante* minimax-regret rule to this environment and derives a set of simple, yet intuitive, decision criteria for both the static and the dynamic choice situation. Simon (1955) also suggested that there can be a *sequence* of deliberations/satisficing where the DM adjusts his or her aspiration level in the light of information discovered. Hence, the dynamic choice situation is of particular interest. Manski’s theory (in his Proposition 2) is that:

(1) The optimal (maximum) number of rounds of deliberation (M^*) if the DM uses a satisficing strategy is given by:

$$M^* = \operatorname{int}[\frac{\log(\frac{U-L)}{k}}{\log(2)}]$$

(2) If the DM uses a satisficing strategy, the DM sets the aspiration level t_m in the m’th round of satisficing as follows:

$$t_m = \frac{L_m + U_m}{2}$$

Here t_m denotes the aspiration level in round m and L_m and U_m are the lower and upper bounds on the payoffs given what the DM has observed up to round m.

(3)

a. Optimisation is an MMR decision if:

$$K \leq U - L \text{ and } K \leq kM^* + \frac{U-L}{2^{M^*}}$$

b. Satisficing with M^* and t_m ($m=1,\ldots,M^*$) is an MMR decision if:

$$k \leq \frac{U-L}{2} \text{ and } K \geq kM^* + \frac{U-L}{2^{M^*}}$$

c. No Deliberation is an MMR decision if:

$$k \geq \frac{U-L}{2} \text{ and } K \geq U-L$$

The intuition of the theory is simple. Deliberation costs play a central role. 'Optimising' or 'Satisficing' will be the decision if their respective associated cost (*K,k*) is low enough. If both costs are sufficiently large then 'No Deliberation' will be preferred. If 'Satisficing' is chosen, the aspiration level is midway between the relevant lower bound and the relevant upper bound, while the number of deliberation rounds is decreasing in its associated cost. This theory is different from the existing search literature in that it provides the concept of satisficing search that follows more closely Simon's perception of adaptive aspiration levels than standard search models. It clearly states *when* the DM should satisfice. It also provides a solution to the choice of aspiration levels.

Before we move on to the experiment, let us briefly translate the above theory into a description of behaviour. The DM starts with knowing that there is a set of payoffs (the number of them unknown) lying between some lower bound *L* and some upper bound *U*. The DM is told the values of *k* and *K*. The first thing that the DM needs to do is to design a *strategy.* This depends on the values of *k* and *K*. If these are sufficiently large (see 3c above), the DM decides not to incur these costs and chooses 'No Deliberation'. The DM is then told and given the payoff of the first action in the choice set, and that is the end of the story.

If *K* is sufficiently small (see 3a above) the DM decides to incur this cost and 'Optimise' and hence learn the highest payoff. He or she gets paid the highest payoff minus *K*, and that is the end of the story.

The interesting case is 3b, where *k* is sufficiently small and *K* sufficiently large. The DM then decides to satisfice with (a maximum[6] of) *M** rounds (as given by 1 above)[7] of satisficing. In each of these *M** rounds, the DM sets an aspiration level, pays *k*, and is told at the end of the round whether or not there are payoffs greater than or equal to

[6] Depending on what the DM learns he or she may not implement all *M** rounds.

[7] After these *M** rounds, the DM should choose 'No Deliberation'. Subjects were informed about that.

the stated aspiration level. More precisely, the DM is told whether there are 0, 1 or more than 1 payoffs greater than or equal to the stated aspiration level .The DM then updates his or her views about the lower and upper bounds on the payoffs in the light of the information received. This updating procedure is simple:

- If there are **no** payoffs greater than aspiration level t_m then $L_{m+1} = L_m$ and $U_{m+1} = t_m$
- If there **are** payoffs greater than aspiration level t_m then $L_{m+1} = t_m$ and $U_{m+1} = U_m$

where L_m and U_m are the lower and upper bounds after m rounds of satisficing.

When at most M^* rounds have been completed the DM gets paid the payoff of the first action in the range between his or her current lower bound and the current upper bound minus kM (the costs of deliberation), where M is the actual number of rounds of satisficing implemented ($M \leq M^*$).

This paper reports on an experiment to test the theory. We test whether subjects choose between 'No Deliberation', 'Satisficing' and 'Optimising' correctly (as in (3) above). We also test, when subjects choose to satisfice, whether they choose the correct number of rounds of satisficing (as in (1) above), and whether aspiration levels are chosen correctly (as in (2) above).

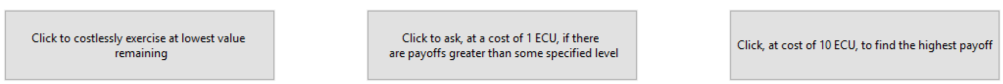
1.3 Experimental design

The actual experimental design differs in certain respects from the design of the theory. First, we told subjects that if they implemented 'No Deliberation' they would be paid the *lowest* payoff in the choice set, rather than the payoff of the first-ordered element of the choice set. Second, we only told subjects, when they chose to satisfice with an aspiration level t, whether there were or were not payoffs greater than or equal to t, and not whether there were 0, 1 or more than 1. Moreover, if after satisficing for m

rounds, and discovering that there were payoffs in a set [L_m,U_m], if they chose 'No (further) Deliberation' at that point they would get a payoff equal to the *lowest* payoff in the set [L_m,U_m] minus *mk*. These differences do not change the predictions of the theory in that an MMR decision-maker will always assume that the first element is the lowest element. Additionally, the *ex ante* choice of *M** remains the same.

Let us give an example (which was included in the Instructions to the subjects). To make this example clear, we need to introduce some notation: the variable *lvgeal* is defined as the *lowest payoff greater than or equal to the highest aspiration level for which there are payoffs greater than or equal to the aspiration level.*

On the screen (see the screenshot below) there were three buttons



The one on the left corresponds to 'No Deliberation', the one in the middle to 'Satisfice' and the one on the right to 'Optimise'. In this example k=1 and K=10.

Suppose – **though the DM does not know this and our subjects were not told this** – that the payoffs are

55 18 75 19 9

If the DM clicks on the left-hand button straight away the income would be 9 (the lowest payoff).

If the DM clicks on the right-hand button straight away the income would be 65 (the highest payoff, 75, minus K).

If the DM clicked on the middle button and specified an aspiration level of 40, he or she would be told that there *are* payoffs greater than this, but would not be told how many nor what they are. The software would, however, note that the lowest payoff greater than or equal to 40 is 55. This would be the *lvgeal* defined above. If the DM clicked on the left-hand button at this stage his or her income would be 54 (*lvgeal* minus *k*). After this first round of satisficing the DM's L_1 and U_1 are 40 and 100 respectively.

If the DM now clicks on the middle button again and now specifies an aspiration level of 70, he or she would be told that there *are* payoffs greater than this, but would not be told how many nor what they are. The software would, however, note that the lowest payoff greater than or equal to 70 is 75. This would become the *lvgeal*. If the DM clicks on the left-hand button at this stage the income for this problem would be 73 (*lvgeal* minus *2k*). After this second round of satisficing the DM's L_2 and U_2 are 70 and 100 respectively.

If the DM now clicks on the middle button a third time, and now specifies an aspiration level of 80, he or she would be told that there are *no* payoffs greater than this. The software would, however, keep the *lvgeal*, 75, in memory. If the DM clicks on the left-hand button at this stage the income for this problem would be 72 (*lvgeal* minus *3k*). After this third round of satisficing the DM's L_3 and U_3 are 70 and 80 respectively.

Subjects could keep on clicking on the middle button as often as they wanted, but they were told that the cost would be deducted from the payoff each time.

Note that in this particular case, it is better to click on the middle button twice (with aspiration levels of 40 and 70) and then on the left-hand button, rather than to click on either the left-hand button or the right-hand button straight away, and better than to click on the middle button one or three times (with aspiration levels of 40, 70 and 80) and then on the left-hand button. **But this is not always the case**.

In the experiment, 48 subjects were sequentially presented with 100 problems on the computer screen, all of the same type. They were given written Instructions and then shown a PowerPoint presentation of the instruction before going on to the main experiment. Subjects were informed of the lower *(L)* and upper *(U)* bounds on the payoffs in each problem; these were fixed at 1 and 100 respectively. They were also told the two types of cost; the cost of finding out whether there are any payoffs greater or equal to some specified aspiration level *(k)* and the cost of finding the highest payoff *(K)*. The number of payoffs *(N)* was fixed at 5, though subjects were not given this information[8]. We used the procedure in Stecher *et al* (2011) to generate the ambiguous distributed payoffs. This procedure creates complete ambiguity for subjects as they have no way to put any probabilities on the payoffs. To make this clear to the subjects we inserted Figures which can be found in the 1A and 1B in the Instructions. Each of them contains 49 distributions, each of 10,000 replications. In the Figure in appendix 1B the drawings were from a uniform distribution over the entire range; in the Figure in appendix 1C from an ambiguous distribution as derived using the Stecher *et al* (2011) method. It will be seen that all the distributions in Figure in the appendix 1B are approximately uniform, while those in Figure in the appendix 1C are all completely different. We told the subjects that "this means that one cannot attach probabilities to each of the numbers coming up. Probabilities are undefined."

We ran two different treatments, Treatment 1 and Treatment 2. In each of these subjects were presented with 100 problems. In Treatment 1, we had four different values for k and K (with N, L and U fixed across the 100 problems); and we gave the subjects these 4 problems in 4 blocks of 25, with the order of the blocks randomised across subjects. In Treatment 2, we had 100 different values for k and K in each of the 100 problems, and presented the problems in a randomised order (again with N, L and U fixed across the 100 problems). Figures 1.1 illustrates. Figure 1.2 shows the predictions of the theory.

[8] This is not relevant to the theory.

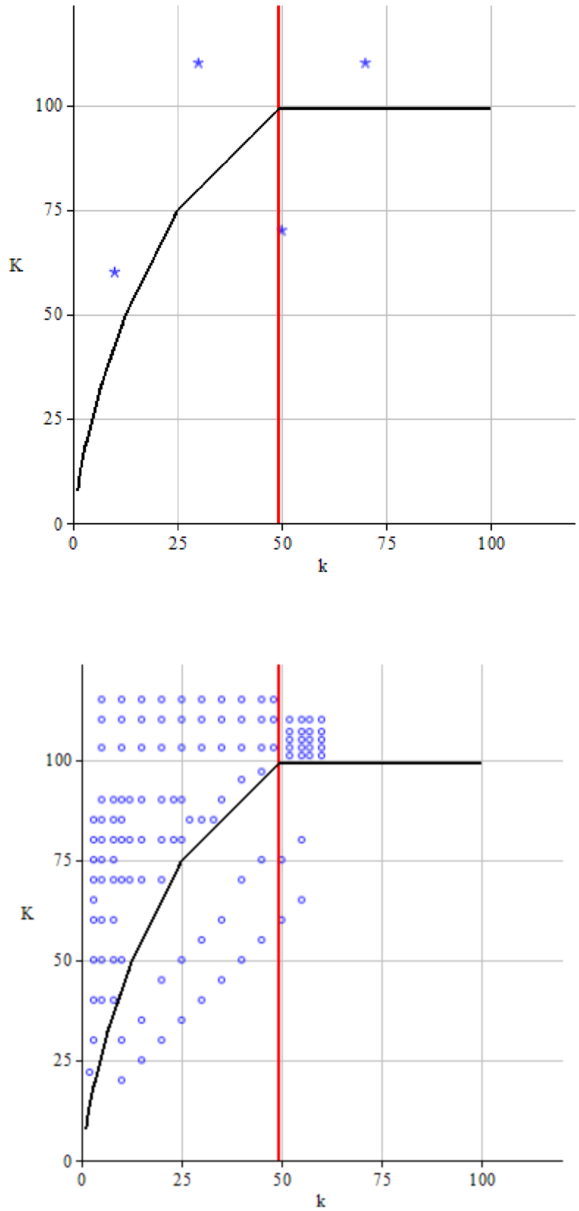


Figure 1.1: Sets of *k* and *K* for Treatment 1 (top) and Treatment 2 (bottom) plotted in the parameter space.

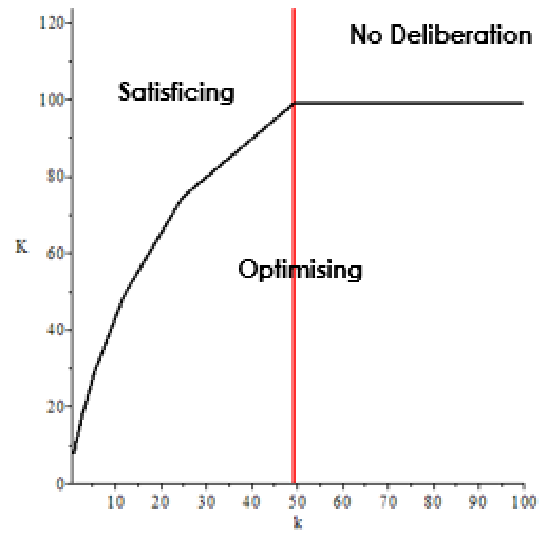


Figure 1.2: Partition of the parameter space into areas corresponding to the theoretical predictions

All 48 subjects completed the experiment which was conducted in the EXEC Lab at the University of York. Subjects' ages ranged from 18 to 44 years. Educational backgrounds were: high school graduate or equivalent (9 subjects); college credit (8); bachelor degree (19); master degree (11); and professional degree (1). 46 subjects reported themselves as a student (8 subjects in a bachelor degree, 9 subjects in a master degree and 11 subjects in doctoral degree); one subject was a member of staff at the University of York; one subject did not report his/her current degree/position. Subjects' ethnicities were mainly White (26 subjects) while 18 were Asian/Pacific Islander, 3 were Black or African American and 1 other. There were only 5 subjects who had any work experience related to finance or economics, but most of them (34 subjects) had previously participated in an economics experiment.

To be a fair test of the theory, we need to give incentives to the subjects to act in accordance with it. We should repeat the fact that the theory is an *ex ante* theory: it tells DMs what to do as viewed from the beginning of a problem; it assumes commitment. Clearly, given the nature of the experiment, we cannot observe what the subjects plan *ex ante*, nor can we check whether they implement their plan. All we can observe is what they do, so we are testing the theory in its entirety – meaning the validity of *all* its assumptions[9]. *Ex ante* the objective of the theory is to minimise the maximum regret. *Ex ante* Regret is the difference between the maximum possible income and their actual income. The maximum possible value of the former is exogenous – it depends upon the problem which in our case is always 100 *ex ante*. So minimising the *ex ante* maximum regret is achieved by maximising their income. So we paid them their (average[10]) income.

A subject's payment from the experiment was their average income from all 100 problems plus the show-up fee of £2.50. Average income was expressed in Experimental Currency Units (ECU). Each ECU was worth 33⅓p; that is 3 ECU was equivalent to £1. They filled in a brief questionnaire after completing all problems on the computer screen, were paid, signed a receipt and were free to go. The average payment was £13.05. This experiment was run using purpose-written software written (mainly by Paolo Crosetto) in Python 2.7.

1.4 Results and analyses

The purpose of the experiment was to test Proposition 2 of Manski (2017) as stated in section 1.2. First, we compare the actual and theoretical decisions for all subjects and in each treatment. Second, we compare the actual and theoretical predictions for income and regret. Third, we analyse the number of rounds of satisficing by comparing the theoretical and actual number for all subjects and both treatments. Finally, we

[9] An alternative design would be to ask subjects to state a plan and then *we* implement it. But 'stating a plan' is not straightforward – not only would subjects have to state whether they want to have 'No Deliberation', 'Optimise' or 'Satisfice', they would also have to specify their rules for choosing their aspiration levels. Asking subjects to do this would be immeasurably more difficult than asking them to play out the problems. We expand on this in our conclusions.

[10] If subjects are maximising their income on each problem they are maximising their average income, and *vice versa*, as problems are independent.

analyse the subjects' actual aspiration levels and compare them with those of the theory.

1.4.1 When to Satisfice

Our experiment gives us 4,800 decisions (between 'No Deliberation', 'Satisficing' and 'Optimising') across 48 subjects and 100 problems. Table 1.1 gives a comparison of the actual and the theoretical decisions; here the main diagonal indicates where subjects followed the theoretical prediction. From this table it can be seen that 2,693 out of the 4,800 decisions (56.10%) are in agreement with the theoretical. The number of theoretical predictions for each strategy can be found at the end of each row while the total number of subjects' decisions can be found at the bottom of each column. Subjects appear to choose 'No Deliberation' significantly more than the theoretical prediction (49.88% compared with 17.50%). Comparing Treatment 1 with Treatment 2 shows that Treatment 2 is closer to the Manski optimal than Treatment 1: 1,476 out of 2,400 actual decisions (61.50%) match with the theoretical in Treatment 2 compared to 1,217 out of 2,400 actual decisions (50.71%) in Treatment 1.[11]

		Subjects' choices			
		No Deliberation	Satisfice	Optimise	**Totals**
Manski's theory	No Deliberation	717 (85.36%)	98 (11.67%)	25 (2.98%)	**840** (17.5%)
	Satisfice	1,079 (34.58%)	1,895 (60.74%)	146 (4.68%)	**3,120** (65%)
	Optimise	598 (71.19%)	161 (19.17%)	81 (9.64%)	**840** (17.5%)
	Totals	**2,394** (49.88%)	**2,154** (44.88%)	**252** (5.25%)	**4,800**

*The number in parentheses indicates the percentage by row and column

Table 1.1: Actual *vs* Theoretical Decisions for All the Subjects

[11] Tables reporting results for treatment 1 and treatment 2 can be found in the Appendix 1A.

In Table 1.2 we compare the actual and theoretical average income and average regret. Obviously, it must be the case that actual regret is higher than the theoretical regret (as subjects were not always following the theory). Subjects also have a higher average income. This suggests that subjects may have been working with a different objective function[12], or making some assumption about the distribution of the payoffs that was not true[13]. Comparing the two treatments, we see that subjects in Treatment 2 have relatively better results in terms of the average income (33.40 ECU to 30.10 ECU) and regret (95.20 ECU to 121.10 ECU) than in Treatment 1. This is interesting, as the idea of Treatment 1 (where each problem was repeated 25 times) was to give subjects a chance to learn; we had expected performance to be better there. Perhaps they learnt about the 'distribution' of payoffs and therefore departed from the theory?

Average Income and Regret			
		Theoretical	**Actual**
All Subjects	**Income**	24.30	31.80
	Regret	65.70	108.20
Treatment 1	**Income**	21.60	30.10
	Regret	72.70	121.10
Treatment 2	**Income**	270	33.40
	Regret	58.70	95.20

Table 1.2: Actual Average *vs* Theoretical Average for Income and Regret

1.4.2 How to Satisfice

[12] For example, maximising Expected Utility.
[13] For example, assuming that the distribution was uniform.

Table 1.3 compares the theoretical (maximum[14]) and the actual number of rounds of satisficing (obviously restricted to the cases where they actually satisficed). There are 452 problems out of 3,120 problems (14.49%), where the subjects should satisfice, and where they choose the same number of rounds of deliberation as the theoretical prediction. The difference between treatments is small: 16.67% and 11.89% matches of theoretical and actual number of rounds of deliberation, for treatments 1 and 2 respectively. Generally they choose fewer rounds of satisficing than the theory predicts[15].

	Actual number of rounds of satisficing												
	M	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**11**	**Totals**
Manski's theory	**0**	1,448	200	19	8	1	2	1	0	0	0	1	**1,680**
	1	852	312	46	8	4	0	0	1	0	1	0	**1,224**
	2	132	163	34	7	0	0	0	0	0	0	0	**336**
	3	190	532	248	69	13	4	0	0	0	0	0	**1,056**
	4	19	89	85	38	27	4	1	0	1	0	0	**264**
	5	18	71	67	44	26	10	1	2	0	1	0	**240**
	Tots	**2,659**	**1,367**	**499**	**174**	**71**	**20**	**3**	**3**	**1**	**2**	**1**	**4,800**

Table 1.3: Actual *vs* Theoretical Number of Rounds of Satisficing

Figure 1.3 shows a plot of actual *vs* theoretical aspiration levels for all subjects (and separately for those in Treatments 1 and 2) where the subjects chose to satisfice[16]. We calculate the theoretical aspiration level based on the relevant lower and upper bounds at the time of choosing satisficing. The forty-five degree line shows what subjects should do if they select their aspiration level according to the theory. The figure shows that subjects' aspiration levels increase with the theoretical levels, although the mean

[14] Note that if subjects were following the theory with our design, the actual number of rounds would be equal to the *M**, while in the theory the actual number could be less than *M** (because they would stop satisficing if they discovered the highest payoff).
[15] This is not a consequence of our experimental design which encourages subjects to choose the maximum number of rounds. Indeed with the theory we might observe numbers below the theoretical maximum.
[16] We exclude the few outliers when the subjects put their aspiration level above 100. There were 39 (1.2%) out or out of 3347 cases where this happened.

equality test shows a rejection of equal means between the actual and theoretical aspiration level when subjects do satisficing (t-test = 15.19, p = 0.000) for all the subjects. Doing this analysis for each treatment separately shows the same result.

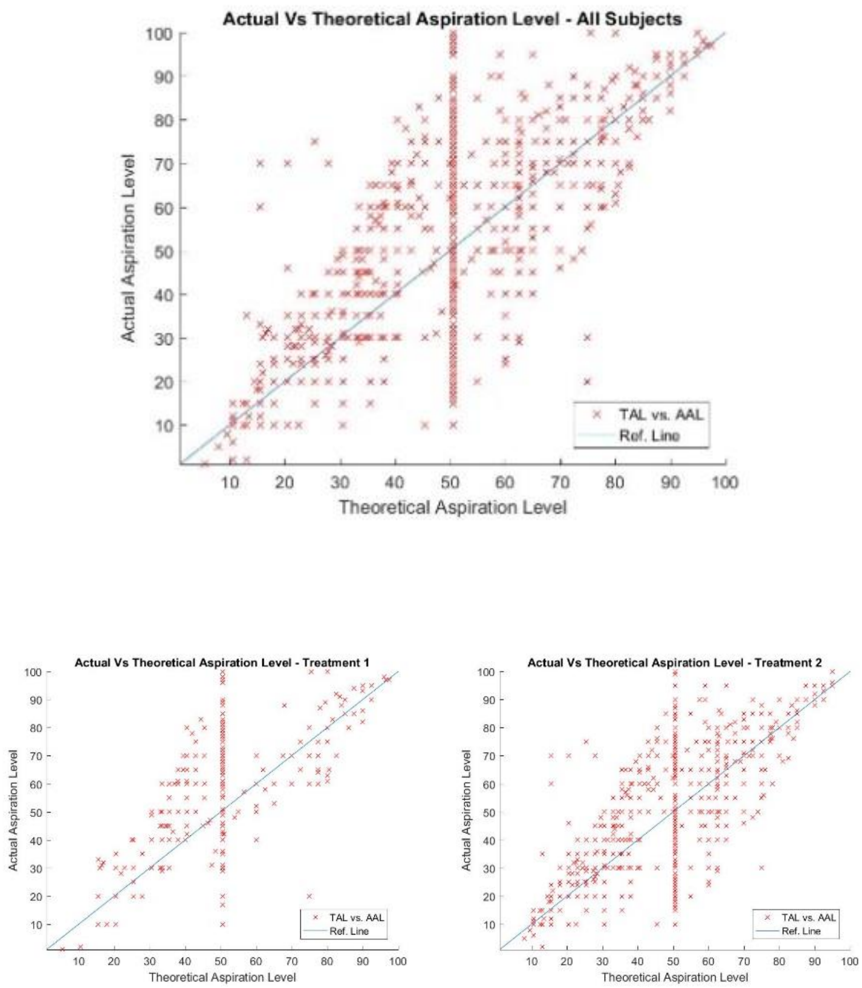


Figure 1.3: Actual *vs* Theoretical Aspiration Level

We now investigate more closely whether subjects set their aspiration level as the theory predicts: equal to the mid-point between the relevant upper and lower bounds. We report below regressions of the actual aspiration level against the optimal level. If the theory holds, the intercept should be zero and the slope should be equal to 1. We omit observations where the aspiration level was above the upper bound (see footnote 9), and accordingly, carry out truncated regressions. Before we proceed to the regressions, we note that the correlations between the actual and theoretical aspiration level 0.544 over all subjects, 0.513 for Treatment 1 and 0.569 for Treatment 2.

	Model 1	Model 2
Theoretical aspiration level	0.994 (0.0208)	1.144 (0.0071)
Constant	7.662* (1.035)	
Observations	3,308	3,308
Wald chi^2	2,273.52	25,592.94

Note: *indicates significance at 1% against the null that the true is 1.0 or 0.0 as appropriate.

Table 1.4: Regressions of the Actual Aspiration Level on the Theoretical Aspiration Level for All Subjects

Table 1.4 shows that the coefficient on the theoretical aspiration level is not significantly different from 1 in Model 1. However in Model 1 we have included a constant term which should not be there; unfortunately it is significantly different from 0, which it should not be. If we remove the constant term to get Model 2, we find that the slope coefficient is almost significantly different from 1. So this table tells us that subjects are almost but not quite following the Manski's rule.

We broke down the analysis of Table 1.4 by treatments. The results are similar for Model 1 in both treatments. In Model 2, we find that the slope coefficient is significantly different from 1 in both treatments.

We now delve deeper and try to understand how the actual aspiration levels are determined, and in particular, how they are related to the upper and lower bounds. We present below regressions of the subjects' aspiration level as a function of these bounds. If following the theory the relationship should be $AL_{im} = 0.5L_{im} + 0.5U_{im}$ (where AL_{im} is subject *i*'s aspiration level in round *m* of satisficing and L_{im} and U_{im} are the relevant lower and upper bounds). As before, we have excluded outliers (aspiration levels greater than the upper bound) from the regression and performed truncated regressions.

	Model 1	Model 2
Lower bound	0.439*	0.441*
	(0.0156)	(0.0158)
Upper bound	0.546*	0.583*
	(0.0153)	(0.00421)
Constant	3.489*	
	(1.315)	
Observations	3,308	3,308
Wald chi^2	2,457.69	32,335.35
Likelihood ratio	710.82	2,113.65

Note: *indicates significance at 1% against the null that the true is 0.5 or 0.0 as appropriate.

Table 1.5: Regressions of the Actual Aspiration Level on the Lower and Upper Bounds for All Subjects

Table 1.5, over all the subjects, shows that the estimated parameters on the bounds are significantly different from the theoretical value of 0.5, and that the subjects put more weight on the upper bound and less on the lower bound when they select their aspiration levels.

If we break down the analysis of Table 1.5 by treatments, we see some differences between them. In Treatment 1 the estimated parameters are significantly different from the theoretical 0.5 (with more weight put on the upper bound than the lower), while in Treatment 2 they are much closer (and indeed only significantly different from 0.5 for one estimated parameter). So in Treatment 2 the subjects are closer to the theory in this respect than in Treatment 1. This confirms an earlier result. Possibly it was a consequence of the fact that in Treatment 2 each problem was an entirely new one, while in Treatment 1 (where 4 problems were given in blocks of 25) subjects were 'learning' about the distribution of payoffs[17] and thus departing from the theory: as the subjects were working through the 25 problems they felt that they were getting some information about the 'distribution'.

1.5 Conclusions

The overall conclusion must be that subjects were not following the part of the theory regarding the 'when' question: the choice between 'No Deliberation', 'Satisficing' and 'Optimising', possibly as a consequence of our experimental design[18]. However, the choice of the number of rounds of satisficing is closer to the theory. The first of these is a particularly difficult task and the second slightly less difficult, and therefore these results may not be surprising. In addition, subjects may have experienced difficulties in understanding what was meant by an ambiguous distribution. However, when it comes to the choice of the aspiration levels, subjects are generally close to (though sometimes statistically significant from) the optimal choice of *(L+U)/2*. This latter task is easier and more intuitive. So it seems that the 'when' part of the theory is not empirically validated, while part of the 'how' part receives more empirical support.

One serious problem with our experimental test (which we have already mentioned) is that the theory is an *ex ante* theory, and one with commitment (so the DM is resolute), while our experimental test involves observing what subjects actually do. A full *ex ante* test is difficult as we would have to ask subjects to specify, not only their choice of deliberation strategy, but also their choice of conditional aspiration levels. Perhaps we could go part-way there by getting the computer to implement some stated aspiration levels, telling subjects the computer algorithm, and asking subjects simply to choose between 'No Deliberation', 'Satisficing' and 'Optimisation'. This would be a partial test

[17] This raises an interesting theoretical point: if we observe 25 repetitions of an ambiguous process, can we learn about it?
[18] Though we should re-iterate that, even though our design differs from that of the theory, the theoretical predictions should be the same.

– one answering only the 'when' of the title. Other variations are possible, but all appear to be difficult.

Let us restate that the theory is an *ex ante* theory and one with commitment: the DM is committed to his or her *ex ante* plan and implements it resolutely. The theoretical predictions may be different if the DM is not resolute. Let us illustrate this with the choice of M^*. At the beginning of the problem the DM calculates M^* – which depends on L and U at the beginning. After m rounds of satisficing the DM will have updated lower and upper bounds. Suppose he or she re-calculates the relevant M^* – call this M_m^*. Will it be true that M_m^* is equal to M^*-m? We see no reason why that should be so – it depends upon the information that the DM has acquired. So it seems perfectly reasonable that a DM should revise his or her plan as he or she works through a problem. But then this is not the optimal way to solve the problem even if the DM is a MMR agent – backward induction should be employed. Perhaps this is what our subjects were doing?

In conclusion we should note that there are three crucial elements to the theory: the use of the MMR preference functional, commitment and the perception of the payoffs as having an ambiguous 'distribution'. The violation of any of these would lead to a breakdown of the theory. We tried to ensure that subjects perceived the 'distribution' as being ambiguous in our experiment. We tried to incentivise the use of the MMR preference functional by our payment rule, but the subjects could well have had a different objective function[19]. Unfortunately it seems difficult to force commitment on the subjects, and they may well have been revising their strategy as they were working through a problem. Nevertheless subjects seem to have been following the theory in at least one key respect – the choice of their aspiration levels.

[19] For example they could have been Expected Utility maximisers operating under the (wrong) assumption that the distributions were uniform.

Appendix 1

1A Actual vs Theoretical Decisions

		Subjects' choices			
		No Deliberation	Satisfice	Optimise	**Totals**
Manski's theory	No Deliberation	524 (87.33%)	64 (10.67%)	12 (2.00%)	**600**
	Satisfice	522 (43.50%)	645 (53.75%)	33 (2.75%)	**1,200**
	Optimise	452 (75.33%)	100 (16.67%)	48 (8.00%)	**600**
	Totals	**1,498** (62.42%)	**809** (33.71%)	**93** (3.88%)	**2,400**

Note: the number in parentheses indicates the percentage by row.

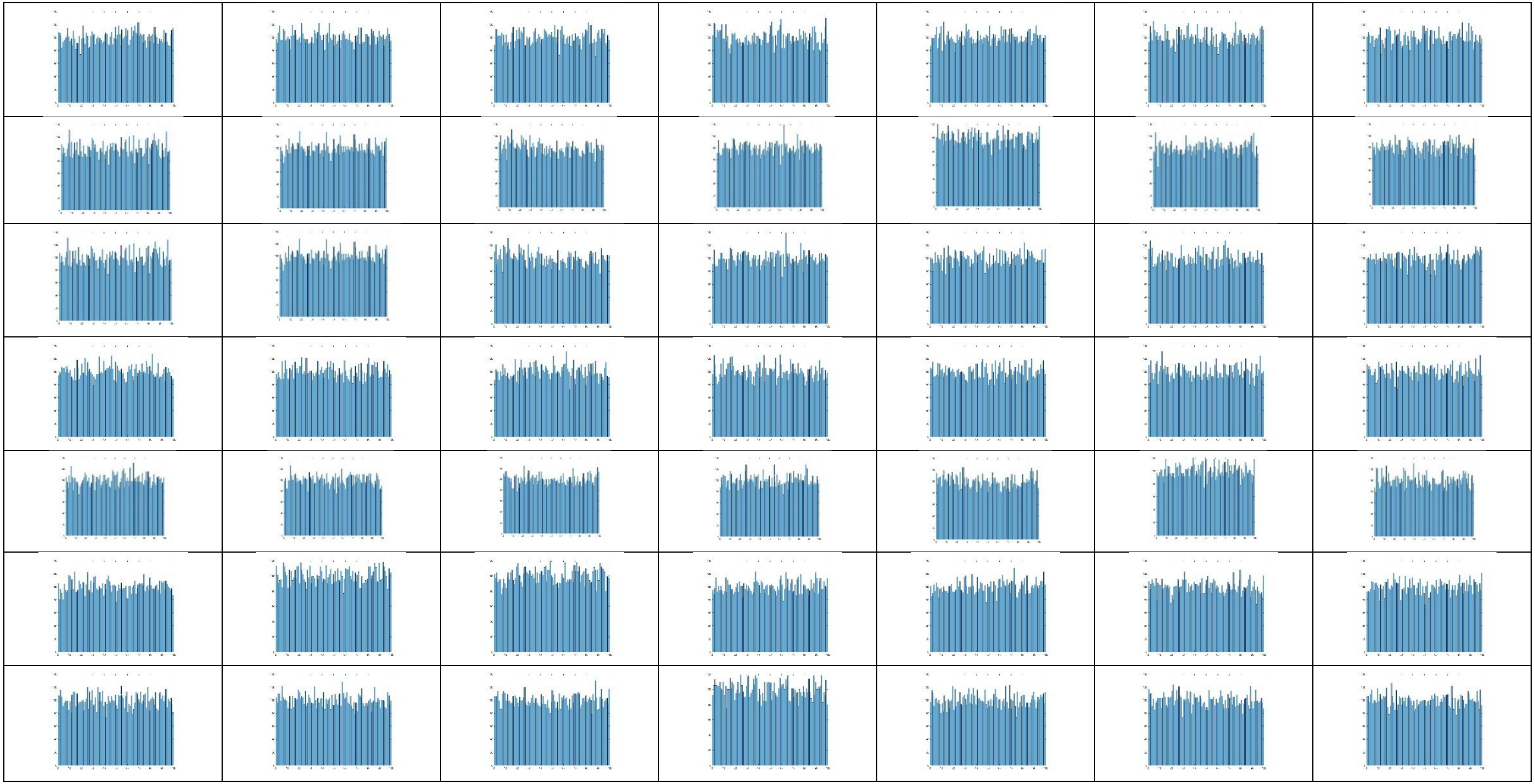
Table 1A.1: Actual *vs* Theoretical Decisions in Treatment 1

		Subjects' choices			
		No Deliberation	Satisfice	Optimise	**Totals**
Manski's theory	No Deliberation	193 (80.42%)	34 (14.17%)	13 (5.42%)	**240**
	Satisfice	557 (29.01%)	1,250 (65.10%)	113 (5.89%)	**1,920**
	Optimise	146 (60.83%)	61 (25.42%)	33 (13.75%)	**240**
	Totals	**896** (37.33%)	**1,345** (56.04%)	**159** (6.63%)	**2,400**

Note: the number in parentheses indicates the percentage by row.

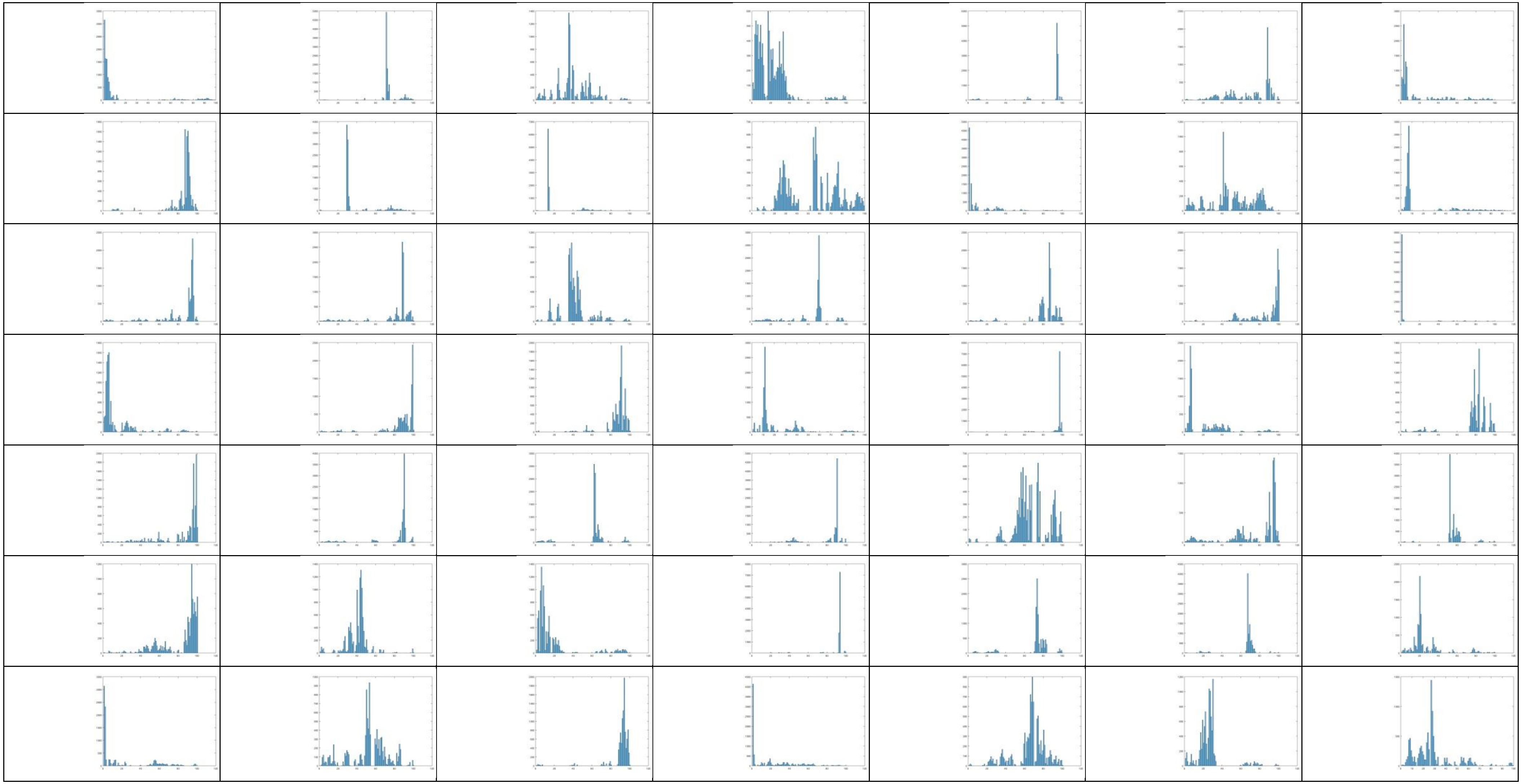
Table 1A.2: Actual *vs* Theoretical Decisions in Treatment 2

1B Uniform Risky distributions[20]



[20] This is figure 1 in the instructions shown to subjects.

1C Ambiguous Distributions[21]



[21] This is figure 2 in the instructions shown to subjects.

1D Instructions

Instructions

Preamble

Welcome to this experiment. Thank you for coming. Please read carefully these instructions. They are to help you to understand what you will be asked to do. You are going to earn money for your participation in the experiment and you will be paid immediately after its completion.

The Experiment

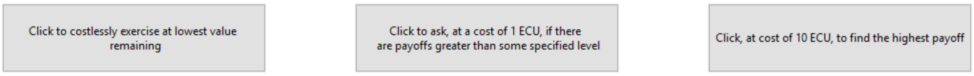
You will be presented with a series of 100 problems, all of the same type. In each problem, there are a set of integer *payoffs*, about which you initially know nothing. During any problem, you might choose to incur some costs to get information about the payoffs. At the end of any problem you will get a particular one of these payoffs. We call your *income* for any problem this payoff *minus* any costs of information that you expended in that problem. Your payment for participating in this experiment will be determined by the average income from these problems, plus a £2.50 show-up fee.

At the beginning of each problem you will not be told anything about these payoffs other than they are between 1 and 100; the payoffs can be anywhere between and including 1 and 100. In fact, they will be randomly distributed between these bounds with what is known as an *ambiguous* distribution. As such a distribution is important to the experiment; we should describe it in more detail.

Ambiguous and uniform risky distributions

Examine Figures 1 and 2 at the end of these instructions. To produce each of these figures we replicated 49 times the drawing of 10,000 random numbers. For Figure 1 we generated them as *uniformly* distributed random numbers. You will see that the number of times that each number between 1 and 100 came up was roughly the same (around 100) on each replication; so one can conclude that the probability of any number coming up in the experiment is 1 in 100. For Figure 2, we generated them as *ambiguously* distributed random numbers. You will notice that, whereas in Figure 1, each of the 49 replications the distributions are approximately the same, in Figure 2, this is emphatically not the case: the distributions vary enormously across the replications. This means that one cannot attach probabilities to each of the numbers coming up. Probabilities are undefined.

<u>Part of the screen</u>



On the screen you will see some information about the payoffs and you will also see three buttons – an example is above. These relate to information that you can buy if you wish.

<u>Information</u>

You can choose, if you want, to buy information about the payoffs, but you do not need to.

If you do *not* want to buy information, then you should click on the left-hand button shown above, and then your income for that problem will simply be the *lowest* payoff in the set of payoffs.

If you *do* decide to buy information, there are two types you can buy – with high (denoted by *K*) and low (denoted by *k*) costs.

If you spend the high cost, *K*, by clicking on the right-hand button above, then the software will tell you the highest payoff in the set of payoffs, so that your income for that problem would be the highest payoff minus the high cost. In the example screen shot above, the high cost is 10 ECU.

If you want to spend the low cost, *k*, then you should click on the middle button above (in the screen shot above this low cost is 1 ECU), and then you will be asked to specify an *aspiration level*. The software will tell you whether there are any payoffs greater than or equal to this value. You will be told *either* that "there *are* payoffs greater than or equal to your aspiration level" *or* that "there are *no* payoffs greater than or equal to your aspiration level". If there are payoffs greater than or equal to the aspiration level, then the software will keep a record of these payoffs, and, in particular, will keep a record of the lowest one of these payoffs (greater than or equal to the aspiration level). We call this payoff the *lowest payoff greater than or equal to the highest aspiration level for which there are payoffs greater than or equal to the aspiration level.* For succinctness in what follows, we denote this by *lvgeal*. We note that the software automatically updates *lvgeal* in the sense that if you try a higher aspiration level and there are payoffs greater than or equal to this aspiration level, then *lvgeal* will become the lowest payoff greater than or equal to this new aspiration level.

You can pay this low cost as many times as you wish (though the costs will be deducted from your final payoff to determine your income for this problem) and you can change your aspiration level.

When you have decided that you have obtained enough information, simply click on the left-hand button, and your income for that problem will be *lvgeal* minus the costs you incurred in finding it. You could, of course, click on the right-hand button and your income for that problem will be the highest payoff minus all the costs you incurred up to that point, including the *K*.

Payment

Your payment from the experiment will be the average income from these problems plus the show-up fee of £2.50. When you have finished all 100 problems, the software will calculate your average income across all 100 problems. In the experiment all amounts are denominated in ECU (Experimental Currency Units). Each ECU is worth 33⅓p; that is 3 ECU is equivalent to £1. The show up fee is £2.50 and this will be added to your payment from the experiment, as described above.

Example (**Note crucially – you will NOT be told the values of the payoffs**. This example is simply to demonstrate how the software works.)

Suppose that k=1 and K=10. Suppose – **though you will not be told this** – that the payoffs are

55 18 75 19 9

If you clicked on the left-hand button straight away your income for this problem would be 9 (the lowest payoff).

If you clicked on the right-hand button straight away your income for this problem would be 65 (the highest payoff, 75, minus the high cost).

If you clicked on the middle button and specified an aspiration level of 40, you would be told that there *are* payoffs greater than this, but you would not be told how many nor what they are. The software would, however, note that the lowest payoff greater than or equal to 40 is 55. This would be the *lvgeal* referred to earlier. If you clicked on the left-hand button at this stage your income for this problem would be 54 (*lvgeal* minus the low cost).

If you now clicked on the middle button again and now specified an aspiration level of 70, you would be told that there *are* payoffs greater than this, but you would not be told how many nor what they are. The software would, however, note that the lowest payoff greater than or equal to 70 is 75. This would become the *lvgeal*. If you clicked on the left-hand button at this stage your income for this problem would be 73 (*lvgeal* minus the low cost twice).

If you now clicked on the middle button again and now specified an aspiration level of 80, you would be told that there are *no* payoffs greater than this. The software would, however, keep the *lvgeal*, 75, in memory. If you clicked on the left-hand button at this stage your income for this problem would be 72 (*lvgeal* minus the low cost three times).

You can keep on clicking on the middle button as often as you want, but you should note that the costs will be deducted from the payoff each time. You should also note that your income from a problem can be negative.

Note that in this particular case, it is better to click on the middle button twice (with aspiration levels of 40 and 70) and then on the left-hand button, than to click on either the left-hand button or the right-hand button straight away, and better than to click on the middle button three times (with aspiration levels of 40, 70 and 80) and then on the left-hand button. **But this is not always the case**.

What to next

Your screen is off when you enter the lab. After every subject has read and understood these Instructions (and had any doubts clarified by asking an experimenter), we will tell you to switch the screen on (by pressing the bottom right button). You will see a PowerPoint presentation of these Instructions. To run this, click on the 'From Beginning' button which is located on the top left of your screen. The presentation goes at a predetermined speed and lasts about 5 minutes. When it gets to the end – to a screen saying 'THANK YOU' – please call over an experimenter, and, if necessary, clarify any doubts with him or her. You will then be told how to start the experiment proper.

If you have any questions, please raise your hand and an experimenter will come to you.

John Hey Yudistira Permana Nuttaporn Rochanahastin

May 2016

Chapter 2

Assessing Axioms of Theories of Limited Attention

2.1 Introduction

Most economic theories are built upon axioms. This is particularly true for decision theories and social choice theories. The validity of the predictions coming out of these theories depends upon the validity of the underlying axioms. In a strict sense an axiom can only be right or wrong: one observation violating an axiom can be considered proof that it is wrong. A cynic would argue that all axioms are wrong, and, while that is almost certainly true of axioms in economics, it is not particularly helpful. One way of rationalising violations is to posit that decision-makers make 'occasional' mistakes – that is, there is some noise in their behaviour. We then need to find a way of measuring the amount of noise in behaviour (relative to the theory being tested). We need to measure 'how right' an axiom is.

Research on a new batch of theories addressed to *satisficing*, or sub-optimal, behaviour has stimulated the direct test of axioms as most of them are axiomatically based, but the testing methodology can also be applied in other contexts. The word *satisficing* was coined by Simon (1955) as describing behaviour which is not optimising – behaviour in which the decision-maker aims for a satisfactory outcome rather than an optimal one. There are many theories which try and describe satisficing behaviour, including models of incomplete preferences, models of behaviour under ambiguity, theories of rational inattention, and search theories.

An area which is particularly active is that of theories of rational inattention. This current paper focuses on a particular sub-branch of work in this area, namely Limited Attention. This has recently caught the attention of researchers following the pioneering works of Sims (1998, 2003, 2010). The applications of this body of research extend to wide areas such as Macroeconomics, Games, and so on.[22] Here we focus on the works that use standard choice data and Revealed Preference in two-stage shortlisted procedures.

Manzini and Mariotti (2007), Masatlioglu *et al* (2012), and Lleras *et al* (2017) all have the same structure: a decision-maker (DM) is being asked to choose one element out of some large choice set, but the set is so large that the DM, in order to simplify a complex problem, pays attention to, and hence chooses from, a subset of this set – a subset called the *Consideration Set*. Axioms characterise how the DM does this. The similarity of these three papers is that all (have to) weaken a standard axiom of decision theory, namely the Weak Axiom of Revealed Preference (WARP) (Samuelson (1938)). They do this in different ways with different weakenings. This sub-branch of limited attention theories can be tested with standard choice data and hence can be investigated with the experimental design used in this research.

Some of the example of theories include Manzini and Mariotti (2007) which suggests a 'shortlisting method'. The method is a two-stage procedure, in which the DM in the first stage weeds out unacceptable choices using one criterion and then proceeds in the second stage to a choice using another criterion. Such a procedure is called a 'Rational Shortlist Method' (RSM). They later, in 2012, suggested a relaxation to RSM called 'Categorise Then Choose' (CTC) which allows the DM to

[22] Caplin (2016) provides a useful and comprehensive review.

compare sets of alternatives in the first stage. They suggest that the theories are testable by testing axioms such as WARP, weakening of WARP, and the Expansion axiom (an alternative chosen from each of two sets is also chosen from their union).

Cherepanov *et al* (2013) suggest a similar procedure to CTC in that the DM compares several rationales or 'motivations' in the first stage and then maximises preference among shortlisted alternatives in the second stage. Manzini and Mariotti (2010) experimentally tested some of these theories; their results are reported in the next section.

2.2 Theories being tested and relevant literatures

Masatlioglu *et al* (2012) suggest the use of an 'attention filter'[23] to be the main property of, and which determines, whether some element is a member of a consideration set. In essence, their attention filter requires that a consideration set is unaffected when an alternative that a DM does not pay attention to becomes unavailable. As a result of this property, the DM is revealed to pay attention to some alternatives. A direct contradiction or violation of WARP is needed in order to elicit the DM's preference. An example of the violation is when a DM chooses a_1 from the choice set $\{a_1, a_2, a_3\}$ and chooses a_3 from the choice set $\{a_1, a_3\}$. We will refer to this situation as a choice inconsistency. The model is empirically testable by testing the axiom of WARP with Limited Attention (WARP(LA)) which is:

For any nonempty S, there exists $x^* \in S$ such that, for any T including x^*

if $c(T) \in S$ and $c(T) \neq c(T \setminus x^*)$, then $c(T) = x^*$.

According to this model, *x* is revealed preferred to *y* if and only if when *y* is taken out of the choice set, *x* is no longer chosen. *y* is revealed to attract attention from the DM while *x* is chosen in the original set. This axiom provides an interesting and crucial implication which is the acyclicity property. Preferences implied using the

[23] The property of the attention filters is that the consideration sets are unaffected when an alternative the DM does not pay attention to becomes unavailable.

theory are acyclic. An example of a cycle preference relation would be: $a_1 \succ a_2 \succ \cdots \succ a_k \succ a_1$.

Lleras *et al* (2017) also point out that a consideration set and its primitives might not be directly observable. However, the acyclicity property of the axiom is empirically testable. Their paper bases its theory on the assumption of *contraction consistency*[24] and it coined the term 'competition filter' to be the main property of a consideration set, and subsequently, the revealed preference. Their paper's main axiom is Limited Consideration WARP (LC-WARP):

For any nonempty S, there exists b*∈ S such that for any T including b*,

if (i) c(T) ∈ S, and

(ii) b* = c(T') for some T' ⊃ T

then c(T) = b*

Again, this axiom's main implication is that it does not allow any cycle in the implied preference relation. Masatlioglu *et al* (2012) and this paper are the main focus of this research. We should note that WARP assumes full attention, unlike Lleras *et al* and Masatlioglu *et al* assume limited attention. These two papers are, to the author's knowledge, the first papers in the rational inattention field that provide a method to reveal attention to a particular item and do not impose unobservable criteria to try to identify an item in a consideration set[25].They have not been experimentally tested. Their characterisations are also based on the revealed preference method which is empirically testable from directly observed choices.

There are also other related models that involve two-stage shortlisted procedures. However, they are not investigated experimentally in this paper because the models impose some requirements or assumptions on consideration set formation or the shortlisting procedures which make them incompatible with the

[24] If an alternative was considered in a set then it must be considered in its subset.

[25] Consideration set formation is crucial in rational inattention but it is very difficult to observe or pin down. These two papers provide methods (through characterisations) on inferring the existence of a particular element in a consideration set. Papers assuming a stochastic consideration set for example Manzini and Mariotti (2014) raise questions as to whether they are testable.

experimental design in this paper[26]. Manzini *et al* (2013) provides characterisation for Two-Stage Threshold representation (TST). Alternatives survive the first stage screening if a threshold value is reached. In a series of papers, Tyson (2008, 2013, 2015) developed extensively two-stage incomplete preference models which provide some connections between satisficing to attention. A DM maximises a binary relation from imperfectly perceived preferences in the first stage and maximises a binary relation over alternatives that survived in the second stage. Finally, search and costly information acquisition was developed in Caplin *et al* (2011), Caplin and Dean (2011, 2015), and Matějka and McKay (2015). Manzini and Mariotti (2014) provides a probabilistic version of a consideration set. A random consideration set is a randomly drawn subset of the choice set. The actual choice is the most preferred item in the consideration set.

In terms of empirical literatures, Manzini and Mariotti (2010) is the closest in spirit to this research. They report on a choice experiment using remuneration instalments as alternatives. There are 4 instalment plans and subjects were presented with all combinations of them. The paper investigates axioms from standard decision theory (WARP) as well as other three theories, namely RSM, CTC and Cherepanov *et al* (2013)'s version of Rationalisation. They find that one aspect of WARP (Condorcet) is violated substantially more than the other (pairwise choice). Therefore, models that are more compatible with the Condorcet property, for example, CTC, are more likely to be successful in the experiment. As expected, WARP is violated the most. On the other hand, CTC and Order Rationalisation perform well. They also use Selten's Measure of Predictive Success, which introduces a parsimony factor, to take into consideration some of the nested-ness of these models.

Chetty *et al* (2009) observed inattention in the case of taxation. They conducted a field experiment observing the difference between tax-inclusive and tax-exclusive price tags in a grocery store and find that changes in tax policy affect demands more in tax-inclusive price tags. De los Santos *et al* (2012) uses data on web browsing and

[26] The design in this chapter focuses only on choice data.

online book purchasing to test search models. The paper rejects a sequential search model in which a consumer always buys from the last store she visited, when he/she crosses the reservation benchmark and favours the fixed sample size search strategy.

Caplin *et al* (2011) report on a search experiment within which there were four 'Experiments'. In Experiment 1 search was over a set of payoffs expressed as simple sums ("two plus eight minus six") differing in their number and complexity, with no time constraint; the sums were generated from an exponential distribution (shown to the subjects). In Experiment 2, subjects were told that their payment will be at a random time in a decision problem. This is to incentivise subjects to always choose the best alternative at that moment in time. Experiment 3 was designed to explore how screen position and object complexity impacts search order. Experiment 4 was the same as Experiment 1 with a two-minute time constraint. The novelty of this paper is that it recorded provisional choice data and contemplation times. They find evidence supporting the sequential search and satisficing model.

2.3 Experimental Design

The purpose of this research is to experimentally determine which of Masatlioglu *et al* (2012) or Lleras *et al* (2017) appears to be the empirically more plausible WARP weakenings, if at all empirically better than WARP itself. Therefore, the experimental procedure is relatively close to Manzini and Mariotti (2010). It is a choice-function-eliciting experiment where the alternatives are risky lotteries. Lotteries are used because the experimenter can provide a real monetary incentive through them without the objects having any objective value. Neither theory imposes a minimum number of alternatives facing the DM, but the more alternatives presented to the subject the better, as this is an attention-related experiment. This gives a higher chance of preventing subjects from recognizing the pattern of the alternatives or carefully deliberating through each of the problems. The drawback is that there is a limitation on the number of problems that the experimenter can present to the subjects. Therefore, we presented only subsets of problems.

This experiment had 10 baseline lotteries[27] which imply a total of 1,023 possible subsets. The subjects were presented with 118 of them[28]. The analyses and comparison between theories will be done using these problems[29]. The lotteries are designed to be similar but contain some differences[30]. The expected value of the lotteries varied from a minimum of £8.00 to a maximum of £9.80. The randomisation process for selecting the subsets started from randomly selecting a subset of two alternatives and based on that randomly selecting a higher number of alternatives' subsets. We, first, randomly selected 5 2-alternative subsets and based on that, we randomly selected 3-alternative subsets that are supersets of one of those 5 2-alternatives subsets. After that, we randomly selected 4-alternative subsets that are supersets of one of those 3-alternatives subsets, and so on.

The lottery visualisation is in a two-dimensional figure where the x-axis represents probabilities and the y-axis represents money outcomes; we used this because there are two important attributes that comprise a lottery: the money outcome and the probability. We feel that a two-dimensional figure best captures this concept as well as giving the subject some idea of the expected value of a lottery in the form of the shaded area. Figure 2.1 shows an example of a lottery and how it

[27] Lotteries details can be found in Appendix 2A.

[28] Randomised lotteries in each problem can be found in Appendix 2B.

[29] de Clippel and Rozen (2014) notes that there could be a problem in a limited dataset. One would need a choice function which is defined for all choice problems in order to conclude that observed choices are consistent with a theory. However, we intend to measure 'how right' the axiom is by identifying *how many* violations a subject makes, given the information available from these problems. Unlike Manzini and Mariotti (2010), we do not assume that a subject is either consistent or inconsistent which require complete knowledge over dataset. We will further address to this point in the results and analysis section.

[30] There are stochastically dominated lotteries. If the DM paid attention and weeded out these lottery, WARP, which assumes full attention, will capture this. Therefore, the DM's behaviour should satisfy WARP. This does not affect the analysis of the two limited attention axioms either because if the DM paid attention to these lotteries, the DM's preference inference method remain unchanged.

was presented to the subjects. This example lottery has 8 in 10 chance of gaining £7 and 2 in 10 chance of gaining £11.

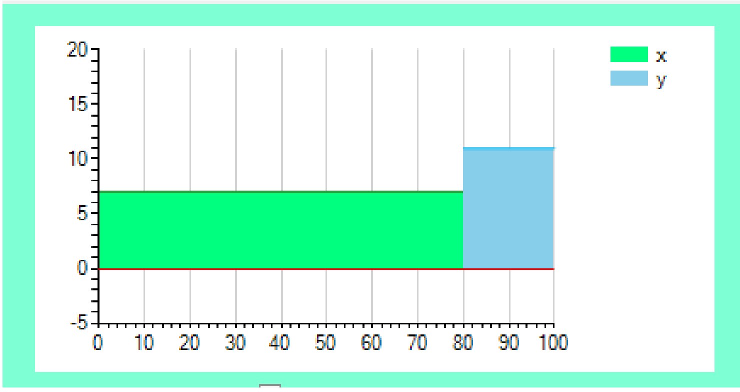


Figure 2.1: A visualisation of a lottery.

Figure 2.2 shows an example screenshot of a problem faced by the subjects. The subject’s task is relatively straightforward: to choose the most preferred lottery in each problem. Taking into account that this is an attention experiment, an upper bound of 45 seconds per problem was imposed. Subjects had to wait a minimum of 10 seconds before confirming their choice, to minimise them clicking at random.

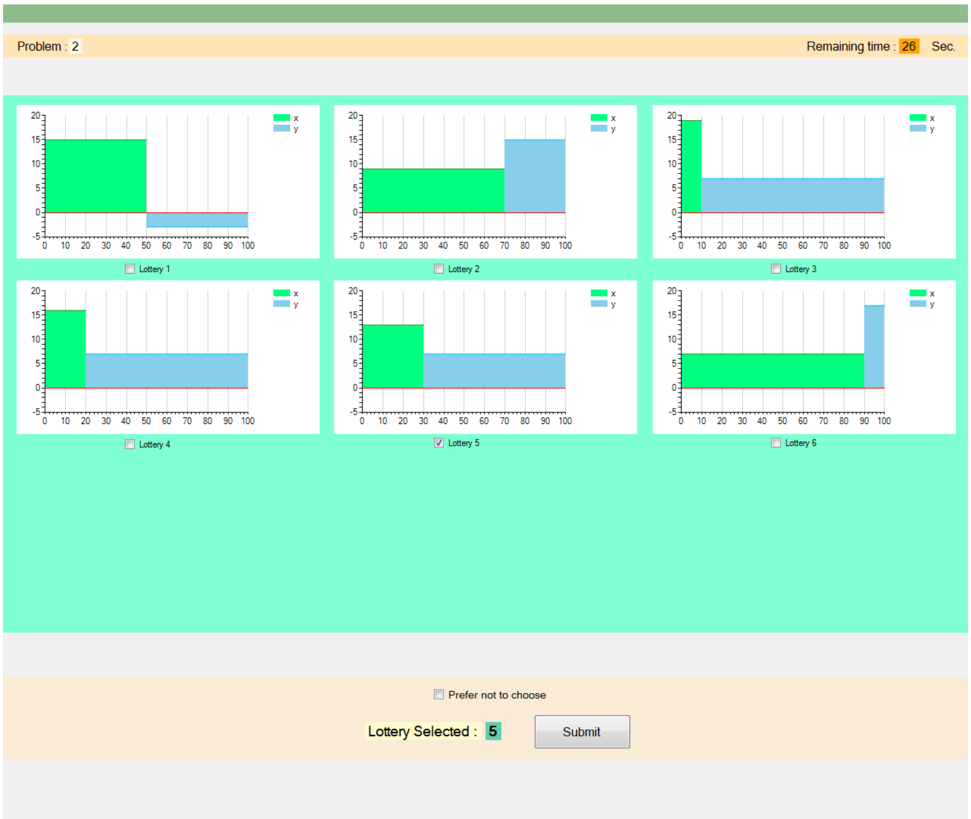


Figure 2.2: An experimental screenshot.

At the end of the experiment, for each subject the chosen lottery in a randomly selected problem was played out for real. Each subject randomly selected a problem for their payment in private by drawing a disk from a bag containing numbered disks from 1 to 118. Their lottery choice in that problem was then played out for real by drawing from another bag containing 10 disks, a multiple of 10 from 10 to 100. The total payment for the experiment was the lottery payoff plus a £3 show-up fee. Subjects were informed that some lotteries involve losses. The maximum loss outcome of any lottery is £3. If subject's lottery payoff is a loss then this is taken out from the show-up fee. After the payment, subjects were free to go.

We recruited a total of 65 subjects for the experiment which was conducted in the EXEC Lab at the University of York. Subject's ages ranged from 18 to 44 years. 64 of whom were students and one was a member of staff at the University of York. There were 34 females (52.31%) and 31 males (47.69%). The average total payment per subject was £11.25. Subjects spent an average of less than one hour in the laboratory. This experiment was run using purpose-written software written in Visual Studio.

2.4 Results and Analyses

The Weak Axiom of Revealed Preference (WARP) is the textbook normative axiom and a baseline description of utility maximisation behaviour. Choice inconsistencies violate the axiom which is essential to the utility maximisation model but they have long been confirmed by a large volume of empirical literature[31]. Much of this literature uses observed individual choice because it is the most obvious and appealing as a measure of preference. As an outside observer, we can infer that the chosen item is revealed to be at least as preferred as items that were not chosen (assuming WARP) . Therefore, preference from observed choices plays an important role in this analysis. Choice inconsistencies and cycles of revealed preference serve as a key measurement of axiom violations in the various models.

Suppose a DM faces a choice set which consists of a complete subset of *n* alternatives. First, let us consider the preference inference. If a DM fully conforms to and behaves according to WARP, we will be able to uncover a complete and transitive preference ordering: there will not be any choice inconsistencies. Each problem will provide us with information on the preference relation. In addition, if we assume that a DM obeys WARP, we will not be able to infer anything (any preference) using the two limited attention theories. However, if we drop WARP and want to use one of the other two limited attention theories, we need at least one violation of WARP or choice inconsistency from two choice problems in order to be able to infer something from these theories. How much we can infer depends upon the extent of choice inconsistencies (as defined earlier) we observe from the data.

[31] for example, Grether (1978), Grether and Plott (1979).

Now let us take a look at the violations. Choice inconsistency is a violation of WARP. We need at least two choice problems to extract this information. This is not necessarily true with Masatlioglu *et al* (2012) or Lleras *et al* (2017). They allow for a choice inconsistency in two problems. In order for a cycle to happen to the inferred preferences, using any of the two limited attention method, we need at least four choice problems: two problems to infer a binary relation and other two problems to infer a contradiction. This is when a violation of the two limited attention axioms happen. We can be certain that this is a true violation rate of the axiom for this particular DM as a result of a complete subset. Also, this conjecture is under a crucial assumption that a DM has a complete and transitive underlying preference. The cycles represent the violation of the axiom and not that the subject does not have complete underlying preference. Otherwise, a DM is behaving irrationally and we cannot infer anything from the data.

Our experiment design presented subjects with a subset of a complete set of 10 alternatives. In term of preference inference, if a subject fully conformed to WARP, we might not be able to uncover a complete preference under WARP as we did not present subjects with every possible pairwise problems. If a subject behaves according to WARP but we assume Masatlioglu *et al* (2012) or Lleras *et al* (2017), we still will not be able to infer anything from the two limited attention theories. However, if there is a choice inconsistency and WARP is violated, we will be able to infer some preferences according to the two limited attention axioms. Again, how much we can infer depends on the choice inconsistency displays in the data. If there are cycles in those preferences inferred, they are the violations. But there could be more violations to the axiom from the unobserved choices. In the actual analysis, we will compare the relative violations based on the problems given and with the simulations of random behaviour using the same problems that the subject faced.

The preferences will be extracted from choice(s) given each model's requirement. First, the preference inference for WARP is direct and straight forward. The axiom states that, in every choice set, there is the best alternative that must be chosen. It means that the chosen alternative from a choice set is revealed preferred to the other alternatives in the set. Therefore, for every problem, pairwise preference(s) can be inferred. Clearly, the axiom has an implicit assumption of full attention, namely that a DM considers every alternative in the choice set.

For Masatlioglu *et al* (2012), limited attention consideration is taken into account for the inferred preferences. We need to make sure that the alternative attracts attention in order to be able to extract a preference. An alternative x is revealed preferred to y if and only if when y is taken out of the choice set, x is no longer chosen. For example, if a DM choses a_1 from a choice set $\{a_1, a_2, a_3\}$ and a_3 from a choice set $\{a_1, a_3\}$, we can conclude that a_1 is revealed preferred to a_2 because dropping a_2 changes the choice which means that DM must have paid attention to a_2 when he/she chose from $\{a_1, a_2, a_3\}$. Notice that this is a direct contradiction to WARP. WARP needs to be violated in order for Masatlioglu *et al* model to infer anything. Also, we need at least two problems and a choice inconsistency for it to be possible to infer any preference. Note that a 'choice inconsistency' here is not caused by inconsistencies in the underlying preferences but by limited attention as assumed in the analysis.

Lleras *et al* (2017) based their consideration set formation under the assumption that if an alternative attracts the DM in the menu with more alternatives, it will also attract his/her attention in subsets of the menu. A choice change in a smaller menu suggests that the choice is preferred to that from the bigger menu that is its superset. For example, if a DM chose a_1 from a choice set $\{a_1, a_2, a_3\}$ and a_3 from a choice set $\{a_1, a_3\}$, we can conclude that a_3 is revealed preferred to a_1 because the DM must have seen a_1 from the choice set $\{a_1, a_3\}$. This behaviour is also a direct contradiction to WARP and there is a possibility that preferences inferred are incomplete. These three models uncover preferences from observed choices under different (and contradictory) assumptions. The analysis tries to identify the relative strength in term of explanatory power of each axiom. It begins by examining how complete in terms of preference inference. Then, the violations in various aspects are analysed. The preference inference and the violations are two separate issues. Even though there are no violations of the axioms, the inferred preference relation may be only partial. Let us provide an example where this happens using Masatlioglu *et al* model. For simplicity and without the loss of generality, suppose there are four alternatives, a1, a_2, a_3, and a_4 and a DM complete preference ordering is $a_1 \succ a_2 \succ a_3 \succ a_4$. Suppose a DM's choices are as follow:

Problem No.	Choice set	Decision
1	$\{a_1, a_2, a_3, a_4\}$	a_1
2	$\{a_1, a_2, a_3\}$	a_1
3	$\{a_1, a_2, a_4\}$	a_1
4	$\{a_1, a_3, a_4\}$	a_1
5	$\{a_2, a_3, a_4\}$	a_3
6	$\{a_1, a_2\}$	a_1
7	$\{a_1, a_3\}$	a_3
8	$\{a_1, a_4\}$	a_4
9	$\{a_2, a_3\}$	a_2
10	$\{a_2, a_4\}$	a_2
11	$\{a_3, a_4\}$	a_3

Table 2.1: Example of problems and decision (1).

Because preferences inferred from all three theories are in term of binary of pairwise comparisons, the completeness or consistency analysis will be in term of the pairwise (2 alternatives) permutations or combinations of a set of the total alternatives which can be represented by the matrices in the example below. In this example, there are 4 alternatives, hence, the analysis matrix is 4*4 in dimension. In the actual experiments, there are 10 alternatives, hence, 10*10 matrices. The following table below show the matrix of all inferred preference from the above example using Masatlioglu *et al*. The X mark in each cell represents the preference of the *row* alternative over the *column* alternative.

Masatlioglu *et al*	a_1	a_2	a_3	a_4
a_1		X	X	X
a_2				
a_3				X
a_4				

Table 2.2: All inferred pairwise preference according to Masatlioglu *et al*.

Using the model, we can infer that $a_1 \succ a_2$, $a_1 \succ a_3$, $a_1 \succ a_4$, and $a_3 \succ a_4$ This is only partial preference as we cannot elicit the relationship between $a_2 \succ a_3$. The preference inferences in this case are without any inconsistencies to Masatlioglu *et al'*s axiom.

To make it clearer, let us provide another example to show how (partial) preferences can be inferred and that there are violations from each theory given the experimental design. Suppose that the DM is now choosing at random and his/her arbitrary decisions are those in column 3 of table 2.3.

Problem No.	Choice set	Decision
1	$\{a_1, a_2, a_3, a_4\}$	a_3
2	$\{a_1, a_2, a_3\}$	a_1
3	$\{a_1, a_2, a_4\}$	a_2
4	$\{a_1, a_3, a_4\}$	a_1
5	$\{a_2, a_3, a_4\}$	a_4
6	$\{a_1, a_2\}$	a_1
7	$\{a_1, a_3\}$	a_3
8	$\{a_1, a_4\}$	a_4
9	$\{a_2, a_3\}$	a_2
10	$\{a_2, a_4\}$	a_2
11	$\{a_3, a_4\}$	a_4

Table 2.3: Example of problems and decision (2).

Figure 2.3 reports all the *pairwise* preferences that can be inferred from table 2.3[32] for this particular DM. For an example of how to infer preferences according to WARP, let us take a look at problem number 1. a_3 is chosen, therefore, we can infer $a_3 \succ a_1$, $a_3 \succ a_2$, and $a_3 \succ a_4$. Applying the same process to other problems give us the top table of Figure 2.3. Masatlioglu *et al* requires at least two problems to infer any preference. Let us take a look at the first and the second problems: we note that dropping a_4 changes the choice; therefore, we can conclude that a_4 must attract attention of the DM in problem 1 but a_3 is chosen. Hence, a_3 is revealed preferred to a_2. Applying the same process to every pair of problems give us the middle table of Figure 2.3. The inference for Lleras *et al* also requires at least two

[32] Note that these are all *direct* inferences. We do not assume transitivity at this point.

problems. Again, let us take a look at the first two problems: a_3 is chosen in the first problem suggesting that a_3 must attract DM attention in every smaller subset that contains a_3. Hence, we can infer $a_1 \succ a_3$ because a_1 is chosen in the second problem. All preferences inferred by Lleras *et al* for this example are shown in the bottom table of figure 2.3.

WARP	a_1	a_2	a_3	a_4
a_1		X	X	X
a_2			X	X
a_3	X	X		X
a_4	X	X		

All inferred pairwise preference according to WARP.

Masatlioglu *et al*	a_1	a_2	a_3	a_4
a_1		X		X
a_2				X
a_3	X	X		X
a_4			X	

All inferred pairwise preference according to Masatlioglu *et al*.

Lleras *et al*	a_1	a_2	a_3	a_4
a_1		X	X	
a_2			X	X
a_3	X			
a_4	X		X	

All inferred pairwise preference according to Lleras *et al*.

Figure 2.3: All inferred pairwise preference.

This analysis of all inferred pairwise preference aims to measure the 'completeness' of the revealed preference from each model. As discussed in Masatlioglu *et al* (2012), their approach depends upon choice inconsistency and the preference inferred can be very incomplete. We calculate the percentage of the revealed preference that can be extracted from the data, by dividing the number of the inferred preferences over the number of possible permutation, to show and compare the 'completeness' of these different models[33]. For this example, WARP can infer 83.33% (10/12) of the binary preference while Masatlioglu *et al* and Lleras *et al* can infer 58.33% (7/12). The hypothesis is that the higher the inferred preference percentage, the more information we can extract from the data.

It is noticeable that there are conflicting pairwise preference relationships in tables in Figure 2.3 For example, both $a_1 \succ a_3$ and $a_3 \succ a_1$ are observed for WARP. These represent cycles which are violations to the axioms. In this next analysis, we extract these conflicting or cycles of pairwise preferences for this DM. We use the term 'depth of the cycles' to represent this analysis.

[33] Because each model has different degree of restrictions, we compare these numbers to the simulations for the actual analysis. The simulation procedure will be explained in the next section.

WARP	a_1	a_2	a_3	a_4
a_1			X	X
a_2			X	X
a_3	X	X		
a_4	X	X		

Depth of the cycles from preferences inferred according to WARP.

Masatlioglu *et al*	a_1	a_2	a_3	a_4
a_1				
a_2				
a_3				X
a_4			X	

Depth of the cycles from preferences inferred according to Masatlioglu *et al*.

Lleras *et al*	a_1	a_2	a_3	a_4
a_1			X	
a_2				
a_3	X			
a_4				

Depth of the cycles from preferences inferred according to Lleras *et al*.

Figure 2.4: Depth of the cycles.

The percentages of the violations can be calculated by dividing these number of conflicting preferences over the number of the all inferred preference as observed in Figure 2.3 For this example, WARP cycle preferences account for 80.00% (8/10) of the all inferred preference while Masatlioglu *et al* and Lleras *et al* cycles account for 28.57% (2/7). This analysis assesses violations of the axioms. The hypothesis is that the lower the violation rate, the better the axiom is, descriptively.

The cycles raise the question about the underlying preferences. We cannot infer actual preferences from these violations. If we extract only non-conflicting or acyclic pairwise preference from Figure 2.3 for this DM, we can fill another table which reports only the pairwise combinations of the alternatives. The marks in the following table show that we can infer the valid cycle-free[34] relationships within the pair. For example, a mark in the cell (a_1, a_2) can be either $a_1 \succ a_2$ or $a_1 \prec a_2$. This is the analysis to show both the validity of the axioms and the completeness of the inferred preference in term of only the consistent preferences.

[34] This refers to the cycle of length two as transitivity is still not assumed at this point.

WARP	a_1	a_2	a_3	a_4
a_1		X		
a_2				X
a_3				
a_4				

Valid inferred pairwise preference according to WARP.

Masatlioglu *et al*	a_1	a_2	a_3	a_4
a_1		X	X	X
a_2			X	X
a_3				
a_4				

Valid inferred pairwise preference according to Masatlioglu *et al.*

Lleras *et al*	a_1	a_2	a_3	a_4
a_1		X		X
a_2			X	X
a_3				X
a_4				

Valid inferred pairwise preference according to Lleras *et al.*

Figure 2.5: Valid inferred pairwise preference.

The validity percentage is calculated by dividing the number of these consistent preferences over the number of the possible combinations. For this example, WARP cycle-free preferences account for 33.33% (2/6) of the all inferred preference while Masatlioglu *et al*'s and Lleras *et al*'s valid part of the preferences account for 83.33% (5/6). The hypothesis for this analysis is that the higher the percentage, the more valid and complete the axiom is. Next section we present sequentially these analyses from the actual data.

2.4.1 Inferred preferences

First, we take a look at how much of (or how complete are) the preferences each theory can infer, given the same set of problems (choice sets) and decisions. All these theories can infer direct pairwise preferences according to the methods mentioned above. In term of permutations, there are 90 possible pairwise preference relationships which can in principle be inferred from the experimental data[35]. The inference percentage per subject is calculated out of these 90 relationships. Table below reports the average percentages over all subjects. The simulation's method and procedure will be explained after the table.

All inferred pairwise preferences	Actual	Simulation	Absolute Difference	Relative Difference
WARP	69.62%	96.32%	-26.70 p.p.[36]	-38.35%
Masatlioglu *et al*	28.10%	45.89%	-17.79 p.p.	-63.61%
Lleras *et al*	47.18%	89.24%	-42.06 p.p.	-89.15%

Table 2.4: All inferred pairwise preferences.

Since crudely comparing these numbers to determine the relative validity of axioms will not work, we have developed a method of providing a 'benchmark'. This is derived from simulations of random behaviour and counting the number of inferred preferences. At the end of the day we can compare the observed number of violations of each axiom with the benchmark figures and hence provide a relative measurement of 'how good' is each axiom. Different subsets of problems will also give different inference. Therefore, simulations using the same set of problems penalise for different degrees of restriction of each theory to give a fairer competing ground in the comparisons. In this case, the simulation is done by creating 100,000 repetition of random decisions using the same 118 problems that subjects faced. These decisions are used to extract preferences in the same manner

[35] The extraction procedures are similar to example provided in Figure 2.3. The difference is that in the actual experiment, the grand set of 10 alternatives were presented to subjects and table 2.4 reports the average over all subjects or repetitions.
[36] Percentage points.

as the actual data (the same procedures explained in the example given in table 2.2 and 2.3 above). This method can serve as one of the suggestions for a direct and non-parametric test of these theories.

The actual inference percentages show that we can partially infer the preferences from the choice data. The *t*-test for difference between two population means is employed to verify that the average from the actual experimental data is significantly different from the average from the simulations for each theory. We found that the *p*-values for WARP, Masatlioglu *et al*, and Lleras *et al* are$7.71 * 10^{-23}, 1.60 * 10^{-15}\ and 5.66 * 10^{-23}$respectively, suggesting that the two means are significantly different. The hypothesis here is that the higher the inference percentage, the more information we can extract from the data. Also, since every theory's percentage is a decrease relative to the simulation, a lower decline rate is preferred. WARP is more restricted than the other two models so the random behaviour provide the highest inferred preference percentage at 96.32%. This is followed by Lleras *et al* at 89.24% and then by Masatlioglu *et al* at 45.89%. The actual data shows that Lleras *et al* declines the most in term of absolute percentage points, from 89.24% to 47.18%. The ranking in term of the ability to extract preferences remains the same which suggests that WARP is the most restricted model followed by Lleras *et al* and the then by Masatlioglu *et al* In terms of relative differences, both weakening theories have significantly falls, suggesting that they are less restricted than anticipated when compared to WARP.

2.4.2 The inconsistencies

Next, the cyclicity of the inferred preferences are analysed in different ways. Choice inconsistencies or revealed preference cycles are the main criteria that can be used to measure the relative degree of validity of the three axioms, since all three characterisations involve a common acyclicity property. We are going to look at the breadth, depth, and length of the cycles based on the categorisation by Bouacida and Martin (2017). The breadth of the cycles are how spread cycles are observed among experimental subjects. The depth and length of the cycles delve deeper into individual behaviour. The depth investigates direct inferred pairwise preferences while the length applies the transitivity assumption to the direct inferred preferences.

2.4.2.1 The breadth of the cycles

First, we want to look at how widespread the inconsistencies are shown among experimental subjects. The data revealed that the inconsistencies are much more extensive in this experiment (10 alternatives) as compared to Manzini and Mariotti (2010) (4 alternatives). We found that 100% of the subjects showed some degree of choice inconsistencies according to WARP and Lleras *et al* while 73.84% (48 out of 65 subjects) display inconsistent preferences according to Masatlioglu *et al*[37]. This finding is as expected and consistent with empirical literatures[38] observing pervasive preference cycles in choice behaviour.

2.4.2.2 The depth of the cycles

Next, we take a look at how much cycles invaded into the inferred preference. We begin by analysing the all inferred pairwise preference (as presented in table 2.5). The depth of the cycles are represented by the proportion of those inferred preference that exhibit inconsistencies. This can be calculated by dividing the number of inferred pairs that exhibit cycles by the total number of pairs inferred. Table 2.5 reports these percentages[39]. The hypothesis is that the higher the percentage of the cycles, the more violation of the axioms are shown in the data. In term of comparison with the simulations, the greatest decline represent the best relative performance.

---

[37] Note that this is the lower bound of the subject that display inconsistency at least one inconsistency. There might be more given the complete dataset. However, the focus of this analysis is to demonstrate that majority of subjects are inconsistent and cycle behaviour is extensive.

[38] Again, see Grether (1978), Grether and Plott (1979), for example.

[39] The calculation procedure is the same as in Figure 2.4. However, table 2.5 reports the average over all subjects or repetitions.

Depth of the cycles	Actual	Simulation	Absolute Difference	Relative Difference
WARP	58.73%	96.70%	-37.97 p.p.	-64.65%
Masatlioglu *et al*	20.30%	39.98%	-19.68 p.p.	-96.95%
Lleras *et al*	70.17%	91.38%	-21.21 p.p.	-30.23%

Table 2.5: Depth of the cycles.

Because the denominator in the calculations is the inferred preferences, these percentages already take into consideration the degree of restriction. Therefore, the comparison of the actual percentages is also applicable. Again, we use the *t*-test for the difference between two population means and find that the *p*-values for WARP, Masatlioglu *et al*, and Lleras *et al* are $1.18 * 10^{-19}$, $1.39 * 10^{-13}$, $8.88 * 10^{-11}$respectively, rejecting that the null hypotheses of equal means. Masatlioglu *et al* shows the greatest improvement, in term of the relative difference, compared to the simulations while the Lleras *et al* violation percentage shows the least improvement. Masatlioglu *et al* also has the lowest actual violation percentage while Lleras *et al* has the highest. WARP shows a significant improvement from the simulation that displays almost 100% violation rate and the actual violation percentage is still relatively higher than Masatlioglu *et al*

Next, we delve deeper into the validity of each axiom by focusing on the valid (consistent) inferred preference. This can be done by observing the inferred preferences calculated in section 4.1 and extracting only those pairwise preference combinations that do not exhibit any inconsistency[40]. The proportion of these valid relations over the total number of pairwise choice combinations (45 pairs) are calculated, and reported in the second column of table 2.6 – 'Valid inferred pairwise preference'. This also shows how complete are the inferred preferences, taken into the account only consistent preferences. The hypothesis here is the higher the valid inferred preference, the more consistent and complete is the axiom, given the problems used in the experiment. Also, an improvement over simulations is preferred.

[40] The procedures are similar to those examples in Figure 2.5.

Valid inferred preference	Actual	Simulation	Absolute Difference	Relative Difference
WARP	50.97%	0.63%	44.71 p.p.	87.72%
Masatlioglu *et al*	40.75%	54.57%	-13.82 p.p.	-33.91%
Lleras *et al*	22.50%	15.16%	7.34 p.p.	32.62%

Table 2.6: Valid inferred pairwise preferences.

Intuitively, the valid inferred pairwise preferences are inversely related with the all inferred preferences. More restricted models result in higher preference inference, which in turn, translates into higher chance of cycles and less valid inferred preferences. The *t*-test of difference in means rejects the null hypothesis of equality in means between actual data and simulation for all theories. The results show that Masatlioglu *et al* has the higher percentage of valid inferred preferences when compared to Lleras *et al*; this is as expected because it is the less restrictive model. However, it declines by 13.82% relative to the simulation. WARP has the highest increment on the relative difference. It is an improvement relative to only 0.63% of the inferred preference in the simulation because it is the most restrictive model. This shows modest support to both Masatlioglu *et al* and WARP. Lleras *et al* perform relatively better than Masatlioglu *et al* in this category. The valid inferred preference is 22.50%. Its validity and completeness improved 32.62% over the simulation of random behaviour. Lleras *et al* is more restrictive compared to Masatlioglu *et al* as observed from higher percentage of all inferred pairwise preference (table 2.4). It also shows greater improvement over the simulations.

2.4.2.3 The length of the cycles

The violations in each cycle length can be obtained by observing the violations assuming transitive preference. The shortest possible length here is a cycle of length 2. This is a direct inconsistency or a reversal in preference inference. This type of cycle has already been analysed in the previous section. Longer lengths are obtained from applying the transitivity assumption and the longest length is 10. We calculate the violation percentage[41] at each length. It is unclear whether the longer or the shorter the length is more problematic for the underlying complete preference ordering: one can argue that a cycle of length 2 is a direct contradiction but also on the other hand, it is not sensible for a longer transitive preference to have a contradiction as the preference ranking should be clearer.

[41] The procedures are similar to those in table 2.5.

Cycle Length	WARP	Masatlioglu *et al*	Lleras *et al*
2	58.73%	20.30%	70.17%
3	75.48%	35.34%	85.45%
4	79.81%	54.83%	90.42%
5	80.18%	63.18%	90.90%
6	80.18%	63.18%	90.90%
7	80.18%	63.18%	90.90%
8	80.18%	63.18%	90.90%
9	80.18%	63.18%	90.90%
10	80.18%	63.18%	90.90%

Table 2.7: Violations in different cycle length from actual data.

Cycle Length	WARP	Masatlioglu *et al*	Lleras *et al*
2	96.70%	39.98%	91.38%
3	99.88%	78.37%	99.62%
4	99.88%	95.61%	99.68%
5	99.88%	96.03%	99.68%
6	99.88%	96.03%	99.68%
7	99.88%	96.03%	99.68%
8	99.88%	96.03%	99.68%
9	99.88%	96.03%	99.68%
10	99.88%	96.03%	99.68%

Table 2.8: Violations in different cycle length from simulations.

Cycle Length	WARP	Masatlioglu *et al*	Lleras *et al*
2	-37.97 p.p.	-19.68 p.p.	-21.21 p.p.
3	-24.40 p.p.	-43.04 p.p.	-14.17 p.p.
4	-20.07 p.p.	-40.77 p.p.	-9.26 p.p.
5	-19.70 p.p.	-32.85 p.p.	-8.78 p.p.
6	-19.70 p.p.	-32.71 p.p.	-8.78 p.p.
7	-19.70 p.p.	-32.71 p.p.	-8.78 p.p.
8	-19.70 p.p.	-32.71 p.p.	-8.78 p.p.
9	-19.70 p.p.	-32.71 p.p.	-8.78 p.p.
10	-19.70 p.p.	-32.71 p.p.	-8.78 p.p.

Table 2.9: Absolute difference of violations in cycle lengths.

Cycle Length	WARP	Masatlioglu *et al*	Lleras *et al*
2	-64.65%	-96.95%	-30.23%
3	-32.32%	-121.78%	-16.59%
4	-25.15%	-74.35%	-10.24%
5	-24.57%	-52.00%	-9.66%
6	-24.57%	-52.00%	-9.66%
7	-24.57%	-52.00%	-9.66%
8	-24.57%	-52.00%	-9.66%
9	-24.57%	-52.00%	-9.66%
10	-24.57%	-52.00%	-9.66%

Table 2.10: Relative difference of violations in cycle lengths.

The percentage increases with the length because of the transitivity assumption. Table 2.7 reports the results of cycle lengths from the experimental data while table 2.8 reports results from the simulations. Tables 2.9 and 2.10 report their differences. The differences in means are significant at every length for every theory. We can see that the pattern remains through every length in term of actual violations. Masatlioglu *et al* shows the lowest violation rates at every length while Lleras *et al* shows the highest. The maximum violations for Masatlioglu *et al* is 63.18% compare to WARP at 80.18% and Lleras *et al* at 90.90%. All the theories improve at every cycle length when compared to the simulations. In term of relative difference, Masatlioglu *et al* still shows greatest improvement over simulations when transitivity is fully explored, follows by WARP and Lleras *et al.*

2.4.3 Results summary and comment

We provide a summary of the direct comparisons between the two weakenings of WARP. Table 2.11 shows which theory performs better in term of actual percentage in accordance with the hypothesis in each theory. Masatlioglu *et al* has fewer axiom violations (the depth of the cycles)[42] and has higher consistent inferred preference rates. Lleras *et al* is shown to be more restricted and can infer more preference in general.

	All inferred preference	Depth of the cycles	Valid inferred preference
Masatlioglu *et al*		X	X
Lleras *et al*	X		

Table 2.11: Theories comparison in term of actual percentage.

Table 2.12 reports the results in term of relative difference when compared to simulations. Masatlioglu *et al* shows a smaller decline in the inference percentage. It also shows the most improvement over the simulations in term of violation

[42] The pattern remains for every cycle length.

percentages[43]. Lleras *et al* improves more in the valid inferred preference category. This results from it being the more restricted theory which causes relatively less valid inferred preference in the simulations.

	All inferred preference	Depth of the cycles	Valid inferred preference
Masatlioglu *et al*	X	X	
Lleras *et al*			X

Table 2.12: Theories comparison in term of relative difference when compared to simulations.

Since these three theories provide different predictions and contain overlapping areas, one might argue that there is a need to penalise in order to compare their explanatory power. We have tried to address this issue by using the 'benchmark' procedure. One possible alternative method is using Selten's measure of predictive success (Selten (1991)). The measure is given by:

$$m = r - a$$

where r is the relative frequency of correct predictions (the number of observed outcomes divided by the number of possible outcomes); and a is the penalisation parameter – which is given by the size of the predicted subset compared to the set of all possible outcomes.

There is a practical difficulty of this measure namely, the number of outcomes increases drastically with the number of the number of problems or alternatives. For this study, the predictive parsimony variable (a) in the measure for WARP, given there are 10 alternatives and 118 problems is (3628800/3*10^74) which is approximately zero. The a variable is also the same (zero) for Masatlioglu *et al* and Lleras *et al* since the denominator is also very large. Therefore, the measure is left with just the variable r in our case and it is the violation percentage itself.

[43] The pattern also remains for every cycle length.

2.5 Conclusion

One of the set of theories that attempts to address the sub-optimality of decision making behaviour which has recently emerged and has received much recognition is the set of theories of limited attention or rational inattention. Most of these are founded upon axioms which make the validity of the predictions coming out of these theories dependent upon the validity of the underlying axioms. We experimentally test the axioms underlying two of these new theories directly, namely those of Masatlioglu *et al* (2012), and Lleras *et al* (2017), which are based on the revealed preference framework. The experimental procedure elicits standard choice data. We observe the number of actual violations and then compare these with a 'benchmark' which was derived from simulations of random behaviour.

Out of the two weakenings of WARP, Lleras *et al* is the more restricted version when compared to Masatlioglu *et al.* Therefore, its ability to extract preferences is higher. Masatlioglu *et al* seems to perform the best in the inconsistencies analyses which is the main observation for the axiom violations. Their key axiom is the only axiom for which some subjects do not violate it at all. The axiom displays the least percentage in term of the depth of the cycles and also shows the greatest improvement over simulations. Lleras *et al* does not perform so well in term of the depth of the cycles. However, its validity of the inferred preference shows the greatest improvement over simulations. These patterns remain when full transitivity is assumed. WARP, which is the standard and normative way of describing choice behaviour, received some modest support from the data, in that, it is the most informative model. The crude percentage of the valid inferred preference according to WARP is the highest.

Appendix 2

2A Lottery details

No.	*Px (x100)*	*X*	*Py (x100)*	*Y*	*E*
1	0.5	19	0.5	-3	8
2	0.8	6	0.1	17	8.2
3	0.6	18	0.4	-3	9.6
4	0.7	8	0.3	13	9.5
5	0.1	19	0.9	7	8.2
6	0.2	16	0.8	7	8.8
7	0.3	14	0.7	7	9.1
8	0.4	11	0.6	8	9.2
9	0.9	7	0.1	17	8
10	0.8	13	0.2	-3	9.8

2B List of alternative(s) in each problem

Prob. No.	Alternatives								
1	1								
2	2								
3	3								
4	4								
5	5								
6	6								
7	7								
8	8								
9	9								
10	10								
11	1	2							
12	3	9							
13	4	9							
14	5	9							
15	5	10							
16	1	2	3						
17	1	2	6						
18	1	2	9						
19	1	2	10						
20	1	3	9						
21	1	5	9						
22	3	5	10						
23	3	9	10						
24	4	5	10						
25	4	6	9						
26	4	8	9						
27	5	6	9						
28	5	8	9						
29	5	8	10						
30	1	2	3	10					
31	1	2	4	9					
32	1	2	5	6					
33	1	3	5	9					
34	1	3	5	10					
35	1	3	9	10					
36	1	5	6	9					
37	1	5	7	9					
38	2	3	9	10					

39	2	4	5	10					
40	2	4	6	9					
41	3	4	5	10					
42	3	4	6	9					
43	3	5	6	9					
44	3	5	6	10					
45	3	5	9	10					
46	4	5	8	9					
47	4	5	8	10					
48	4	8	9	10					
49	1	2	3	5	10				
50	1	2	3	9	10				
51	1	2	4	7	9				
52	1	2	4	8	9				
53	1	2	5	6	8				
54	1	3	4	5	10				
55	1	3	5	6	10				
56	1	3	5	7	10				
57	1	3	5	8	10				
58	1	3	7	9	10				
59	1	3	8	9	10				
60	1	5	6	7	9				
61	2	3	4	9	10				
62	2	3	5	6	9				
63	2	3	6	9	10				
64	2	3	7	9	10				
65	2	4	5	6	10				
66	3	4	5	7	10				
67	3	4	6	9	10				
68	3	5	6	7	10				
69	3	5	6	8	10				
70	3	5	7	9	10				
71	4	5	7	8	9				
72	4	6	8	9	10				
73	1	2	3	4	8	9			
74	1	2	3	4	9	10			
75	1	2	3	5	6	10			
76	1	2	4	7	9	10			
77	1	2	5	6	7	8			
78	1	2	5	6	8	10			
79	1	3	4	5	6	10			
80	1	3	4	5	8	10			
81	1	3	4	5	9	10			

82	1	3	4	6	9	10			
83	1	3	5	6	8	10			
84	1	3	5	6	9	10			
85	1	4	5	6	7	9			
86	1	5	6	7	9	10			
87	2	3	4	5	6	9			
88	2	3	4	5	7	10			
89	2	3	4	5	9	10			
90	2	3	4	8	9	10			
91	2	4	5	6	7	10			
92	2	4	5	6	8	10			
93	3	4	5	7	9	10			
94	1	2	3	4	5	7	10		
95	1	2	3	4	5	8	9		
96	1	2	4	5	6	7	9		
97	1	2	4	5	6	7	10		
98	1	2	4	5	7	9	10		
99	1	2	5	6	7	9	10		
100	1	3	4	5	6	7	9		
101	1	3	4	5	6	7	10		
102	1	3	4	5	6	8	10		
103	1	3	4	5	6	9	10		
104	1	3	4	5	7	9	10		
105	1	3	4	5	8	9	10		
106	1	5	6	7	8	9	10		
107	2	3	4	5	6	8	9		
108	2	3	4	5	7	9	10		
109	3	4	5	6	7	9	10		
110	1	2	3	4	5	6	7	10	
111	1	2	3	4	5	6	8	9	
112	1	2	3	4	5	6	9	10	
113	1	2	3	5	6	7	9	10	
114	1	3	4	5	6	8	9	10	
115	2	3	4	5	6	7	9	10	
116	1	2	3	4	5	6	7	8	9
117	1	2	3	4	5	6	7	8	10
118	1	3	4	5	6	7	8	9	10

2C Instructions

Instructions

Preamble

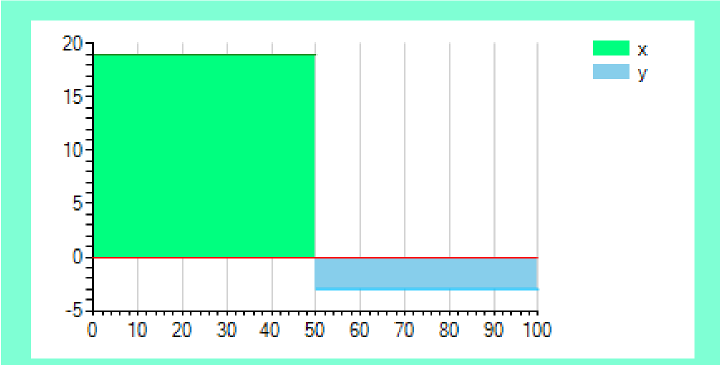
Welcome to this experiment. Thank you for coming. Please read carefully these instructions. They are to help you to understand what you will be asked to do and how will you get paid. The experiment is simple and gives you the chance to earn a considerable amount of money. You will be paid in cash immediately after the experiment is completed.

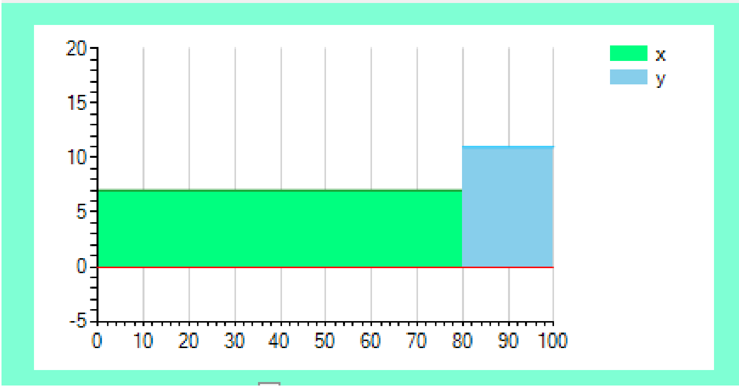
The Experiment

The experiment is interested in how you take decisions. There are no right or wrong answers. You will be presented with a series of 118 problems, all of the same type. In each problem, there is a set of *lotteries*. We will describe in detail what we mean about a lottery in the next section. Your task is to choose one of these lotteries or not to choose any lottery at all in a problem. The outcome of playing out this lottery will lead to a *payoff* to you. Your payment for participating in this experiment will be the payoff from a randomly chosen one of these problems, (playing out the lottery of your choice), plus a £3 show-up fee. If it occurs that you did not choose any lottery in the randomly selected problem, your *payoff* will be your show-up fee. Details of all the payment procedures will be explained in the *payment* section.

A Lottery

We describe now what we mean by a 'lottery'. Here we represent each lottery visually. The visual representation will be like the two examples below,





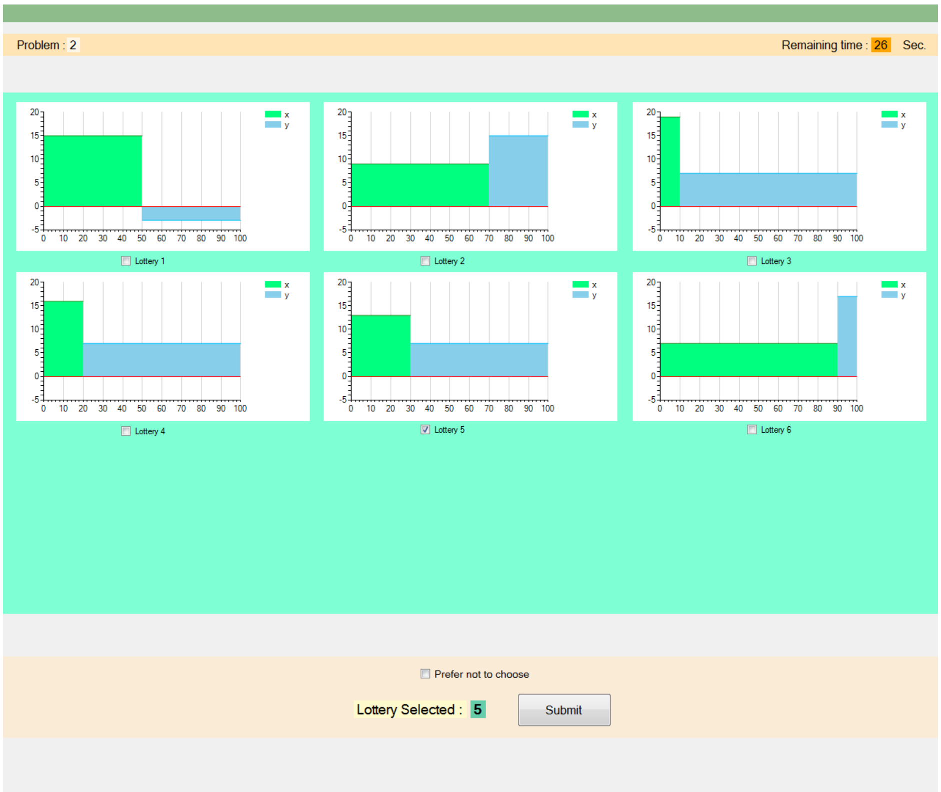
It is simplest to explain these in terms of the implications for your payment if one of these is randomly selected to be played out at the end of the experiment. What we will do in all cases is to ask you to draw – without looking – a disk out of a bag containing 10 disks numbered from 10, and an increments of 10, to 100. (You will be able to check that the bag contains all these disks before you do the drawing.) The number on the disk that you draw will determine a point on the horizontal axis; your payment would be the amount on the vertical axis implied by that point through the figure. At the point on the horizontal axis where the vertical axis changes it value, the payment would equal to the value of the vertical axis to the *left* of that point. In each lottery, there are two possible outcomes or *payoffs*.

So, for example, in the *top* lottery, if the number on the disk that you draw is between 10 and 50 *inclusive* you would get £19, notice that if the number on the disk is 50 you would get £19; if it is between 60 and 100 *inclusive* you would make a loss of £3. This loss will be deducted from your show-up fee. This implies that the chance of you getting paid £19 is 50 percent and the chance of you making a loss of £3 is also 50 percent. This will also be written in words. The caption will appear when you move the mouse cursor over the shaded areas. If the *bottom* lottery is to be played out, if the number on the disk that you draw is between 10 and 80 *inclusive* you would get £7, notice that if the number on the disk is 80 you would get £7; if it is between 90 and 100 *inclusive* you would get £11.

Let us give specific examples. In the *top* lottery, suppose the number on the disk that you draw is 70, then you would make a loss of £3 out of your show-up fee. In the *bottom* lottery, suppose the number on the disk that you draw is 30, you would receive £7.

<u>Choices</u>

In each problem, there is a set of *lotteries*. The number of lotteries varies from problem to problem. Your task is to choose one of these lotteries, or not to choose any lottery. You can choose a lottery by clicking at the box below the lottery of your choice. If you do not want to choose any lottery, you can do that by clicking the '*Prefer not to choose*' button at the bottom part of the screen. Below is an example of a problem screen.



Payment

When you complete all 118 problems, please raise your hand and the experimenter will come to you. You will be lead to a separate room where the payment will take place. You will randomly choose one of the problems to play out for real. This is done by you drawing a disk from a bag containing 118 disks, each labelled number 1 to 118. The number on the disk that you draw is the problem that will be played out for real.

If you chose one of the lotteries in that problem

Your payment from the experiment will be from playing out a lottery of your choice from the randomly-chosen problem of the experiment plus the show-up fee of £3. You will randomly choose one numbered disk from another bag containing 10 disks numbered from 10, 20, 30, …, 100, and the number on the disk chosen will determine your payoff according to the procedure describe in the *lottery* section. If the *payoff* in the randomly chosen problem is zero you will receive only a show-up fee. If the *payoff* in the randomly chosen problem is negative, this will be deducted from your show-up fee. The maximum loss from a problem is -£3, therefore, at worst; you will be receiving £0 from this experiment.

If you did not choose any lottery in that problem

Your payment from the experiment will be only the show-up fee of £3.

What to do next (About the Experimental Software)

When you finish reading these Instructions, you should click on the '*start*' button at the bottom of the screen (you will not be able to click this button until at least 5 minutes have passed). This will lead you to the actual experimental problems, and you will then be starting the experiment proper. Each problem screen has a countdown timer at the top right corner of the screen. You cannot click any button until 10 seconds have passed from when you started on that problem. There is a time limit of 45 seconds to make a decision on any problem. You can change your decision as many times as you want during this time period. You can click '*Submit*' button before the time limit is reached. If you chose a choice and the time limit is over, that choice will automatically be your choice. If you do not choose any choices and the time limit is over, the default option, which is '*Prefer not to choose*', will be taken as your choice on that particular problem.

If you have any questions at any stage of the experiment, please raise your hand and an experimenter will come to you.

Thank you for your participation.

Nuttaporn Rochanahastin
October 2017

Chapter 3

Axioms of Salience Perception and Choice under Risk: An Experimental Investigation

3.1 Introduction

Formal models of human decision making have been proposed for over half a century to explain empirical violations of rational choice theory (Dhami (2016), Wakker (2010)). A more unified perspective is now emerging, based on the idea that fundamental properties of the perceptual system lead people to focus on larger differences in payoffs. These distortions in perceived salience produce deviations from models of rational behaviour. This 'salience-based' account of decision making has been applied to explain decisions under risk (Bordalo *et al* (2012)), decisions over time (Kőszegi and Szeidl (2013)), consumer choice (Bordalo *et al* (2013b)), asset prices in financial markets (Bordalo *et al* (2013a)), judicial decisions (Bordalo *et al* (2015)), competition between firms (Bordalo *et al* (2016)), and strategy selection in games (Leland and Schneider (2015, 2018)). All of these models rely on essentially the same basic properties of the perceptual system. However, the link between these properties of the perceptual system and economic decision making has not been directly tested.

In this paper, we directly test axioms that characterise salience perception and their linkage with economic choices under risk. Our experiment also tests for the presence of a broader link between the representation of non-symbolic stimuli and symbolic stimuli in human decision making. A variety of evidence summarised in Dehaene (2011) suggests that both animals and human infants form approximate non-symbolic number representations, and that such representations are also used by human adults, even when confronted with information that has been presented symbolically. A study on number perception, Moyer and Landauer (1967), concluded that the mind converts symbolic numerals to analogue magnitudes and that "a comparison is then made between those magnitudes in much the same way that comparisons are made between physical stimuli." More recently, Halberda *et al* (2008) found that individual differences in the perception of non-symbolic stimuli (perceptions of arrays of coloured dots) correlates with math achievement on symbolic tasks. Authors in Schley and Peters (2014) found that diminishing sensitivity in the perception of symbolic stimuli correlates with diminishing sensitivity to numerical magnitudes presented symbolically in choices under risk. Yet, it is unknown whether there is a link between the perception of symbolic and non-symbolic stimuli in human decision making.

We first introduce three axioms to formalise properties of the perceptual system that are implied by salience models of decision making. We experimentally test the validity of these axioms in a perceptual task involving boxes containing red and blue dots. We then have participants make choices between lotteries with different possible payoffs and probabilities of winning money. We observe a relationship between individual differences in salience perception and participants' preferences for skewness in choices under risk, as would be expected if salience perception influences behaviour. The preference for positively skewed lotteries in situations involving risk is well known and it generates the purchase of lottery tickets (Friedman and Savage (1948), Kahneman and Tversky (1979), Tversky and Kahneman, (1992)), the 'longshot' bias in betting markets (Weitzman (1965)), and the over-valuation of positively skewed financial assets (Barberis and Huang (2008)). However, the determinants of skewness preference are not well understood. Our findings provide a step in addressing this gap by establishing a direct link between salience perception and the economic preference for skewness.

The properties of salience perception have typically been justified on intuitive and empirical grounds. Now consider how these properties may also be derived from first principles that one might postulate to characterise the perceptual system. Consider salience perceptions between pairs of quantities (x, y). These quantities may be monetary payoffs, for example.

3.2 Axioms of Salience Perception

Leland and Schneider (2016a) considers salience perceptions between pairs of quantities(x, y). These quantities may be pairs of payoffs or probabilities or time delays, for example. Denote the set of quantities being compared by a closed and convex set $\Omega \subset \mathbb{R}_+^2$. Let $\trianglerighteq_s$ be a binary relation called a salience relation carrying the interpretation "at least as salient as" over pairs in Ω, with strictly greater salience and equivalence denoted by $\vartriangleright_s$ and $\sim_s$. The following axioms are imposed on $\trianglerighteq_s$:

Axiom 1 (ORDERING AND CONTINUITY). $\trianglerighteq_s$ is a continuous[44] weak order on Ω. For $x, y, x', y' \geq 0$, let $\Delta(x, y) := x - y \geq 0$, and let $r(x, y) := x/y \geq 1$. Note that our definitions of differences, $\Delta(x, y)$, and ratios, $r(x, y)$ have, without loss of generality, set $x \geq y$.

A salience relation that ranks the salience of pairs of quantities (*x, y*) is complete, transitive, and continuous.

Axiom 2 (SYMMETRY). For any $(x, y), (y, x) \in X$, $(x, y) \sim_s (y, x)$.

[44] The notion of continuity invoked here is that used in consumer preference theory in economics: The relation $\trianglerighteq_s$ is continuous if it is preserved under limits. That is, for any sequence of pairs $\{(x^n, y^n)\}_{n=1}^{\infty}$ with $x^n \trianglerighteq_s y^n$ for all n, $x = \lim_{n\to\infty} x^n$, and $y = \lim_{n\to\infty} y^n$, we have $x \trianglerighteq_s y$ (Mas-Colell, Whinston, and Green (1995), Definition 3.C.1).

For any quantities (*x, y*), the salience of (*x, y*) is equivalent to the salience of (*y, x*).

Axiom 3 (MONOTONICITY IN INTERVALS). For any $(x, y), (x', y') \in X$, if $[y', x']$ is a strict subset of $[y, x]$ then $(x, y) \rhd_s (x', y')$.

For any pairs of quantities (*x, y*), (*x′, y′*), with $x \leq y$ and $x' \leq y'$, if interval [*x′, y′*] is a strict subset of interval [*x, y*] then (*x, y*) is more salient than (*x′, y′*).

Axiom 4 (MONOTONICITY IN RATIOS). For any $(x, y), (x', y') \in X$, if $\Delta(x, y) = \Delta(x', y')$ with $x' > x,\ y' > y$, then $r(x, y) > r(x', y')$ implies $(x, y) \rhd_s (x', y')$.

For any pairs of quantities with the same absolute difference, the pair with the larger ratio is more salient.

Axiom 5 (MONOTONICITY IN DIFFERENCES). For any $(x, y), (x', y') \in X$, if $r(x', y') = r(x, y)$, where $x' > x$ and $y' > y$, then $\Delta(x', y') > \Delta(x, y)$ implies $(x', y') \rhd_s (x, y)$.

For any pairs of quantities with the same ratio, the pair with the larger absolute difference is more salient.

These properties characterise[45] a general class of salience functions including those used in salience models of decision making. Axiom 4 is equivalent to the property of diminishing absolute sensitivity (DAS) which is a form of 'Weber's law,' and Axiom 5 is equivalent to increasing proportional sensitivity (IPS). The general definition of a salience function follows:

Definition 1: (Salience Function): For any pair of quantities (*x, y*), a *salience function* σ(*x, y*) is any symmetric and continuous function that satisfies the following three properties:

[45] Appendix 3A provides a proof which is extracted from Leland and Schneider (2016a).

1. Ordering: If [*x′, y′*] is a subset of [*x, y*], then $\sigma(x', y') < \sigma(x, y)$

2. Diminishing Absolute Sensitivity (DAS): For any $x, y, \varepsilon > 0$, $\sigma(x + \varepsilon, y + \varepsilon) < \sigma(x, y)$.

One other natural property for a salience function is the following:

3. Increasing Proportional Sensitivity (IPS): For any $x, y > 0$, *and* $k > 1$, $\sigma(kx, ky) > \sigma(x, y)$.

Definition 2: A function σ *represents* a salience relation $\trianglerighteq_s$ if for all $(x, y), (x', y') \in X$, we have $(x, y) \trianglerighteq_s (x', y')$ if and only if $\sigma(x, y) \geq \sigma(x', y')$.

To test whether there is a link between salience perception and economic choices, we conduct an experimental test of Axioms 3, 4, and 5 in a perceptual task involving arrays of red and blue dots (Part I of the experiment) and then test whether salience perception predicts risky choice (Part II). Sample tasks from the experiment are shown in Figure 3.1

The approach that we use to test the validity of each axiom is to compare the actual behaviour in discriminating between the differences or ratios of red and blue dots (such as those displayed in Figure 3.1) with random choice[46]. A distribution indistinguishable from random choice would falsify the axioms.

[46]A related benchmark to evaluate DAS (Axiom 4) and IPS (Axiom 5) is 'constant sensitivity'. For a salience function s that satisfies Constant Absolute Sensitivity (CAS) and any $\varepsilon > 0$, $s(x + \varepsilon, y + \varepsilon) = s(x, y)$. For a salience function s that satisfies Constant Proportional Sensitivity (CPS) and any $k > 1$, $s(kx, ky) = s(x, y)$. In our experiment, CAS and CPS make the same predictions as random choice.

3.3 An experimental investigation

We carried out an experiment in two sections. The first section was to test Axioms 3, 4 and 5; the second section was to test the implications of the axioms for risky choice.

Let us start with the first section – a test of the axioms. The axioms are designed to say when a pair of items (x,y) is 'more salient' than another pair (x',y'). We had to interpret what this means. Dictionaries define 'salient' as something important or noticeable, or, occasionally, as something very important or very noticeable. Nowhere is 'more salient' defined, but, in the spirit of the dictionary definitions, we take it to mean 'more noticeable'. So, in keeping with this spirit, we devised an experiment to see whether subjects could detect 'more noticeable'. To do this, we needed a task where subjects were asked to choose between two objects. We could not ask them to choose the object that was more salient, as we would have had to say what that meant. We decided to reward them if the thing that they chose had more of something – that is, was more noticeable. So we would be observing what they found as more noticeable – and therefore more salient. Rather tautologically, if they noticed whatever it was that we were asking them for, they must have found it noticeable.

The way that we implemented this was to give them a series of problems, in each of which they had to choose one out of two boxes. Their payment depended upon whether the box they chose satisfied a question posed to them. If it did their payment was a positive sum of money; if it did not, their payment was zero. This method enables us to test these axioms as properties of perception, independent of any classical domain of choice behaviour. We give details below.

3.3.1 The experimental design in the first section

This section was designed to test the axioms. Subject were presented with 120 pairwise choice questions; each problem testing one of the three axioms. In each problem, subjects were presented with two boxes (*Left* and *Right*). Each box contained *red* and *blue* circles. The numbers of *red* and *blue* circles represent a pair of alternatives according to the axioms. An example of a problem's screenshot for this section is given in Figure 3.1. The subject's task implicitly, but not explicitly, was to choose whichever box was more salient according to the axioms. We incentivised them by paying them £10 if they chose correctly according to a question posed to them, and paying them nothing otherwise. Their total payment for part I was the average payoff from all 120 problems.

For testing Axiom 3 (Monotonicity in Intervals) the question was '*Which box has the greatest difference between the number of blue balls and the number of red balls?*'.

For testing Axiom 4 (Monotonicity in Ratios) the question was '*Which box has the greatest ratio of blue balls to red balls?*'. To make the question a proper test of the axiom, both boxes had the same absolute difference.

For testing Axiom 5 (Monotonicity in Differences) the question was *'Which box has the greatest difference between the number of blue balls and the number of red balls?*'. To make the question a proper test of the axiom, both boxes had the same ratio.

Although subjects were facing similar questions for Axiom 3 and Axiom 5, the underlying parameters (the number of blue balls and red balls) were designed differently. For Axiom 3, in one box – the more salient box – the number of blue balls and red balls is a super set of the other box - the less salient box. This suggests that, in the more salient box, there are more blue balls as well as less red balls compared to the less salient box. Hence, the subject task is to identify which box has more blue balls ***and*** fewer red balls.

For Axiom 5, we kept the ratio of the blue balls and red balls constant and varied the difference. Therefore, the subject's task was to choose the box with the greatest difference, ***given*** a fixed ratio. The questions that asked '*Which box has the greatest difference between blue balls and red balls?*' is sufficient for the condition for both Axiom 3 and Axiom 5.

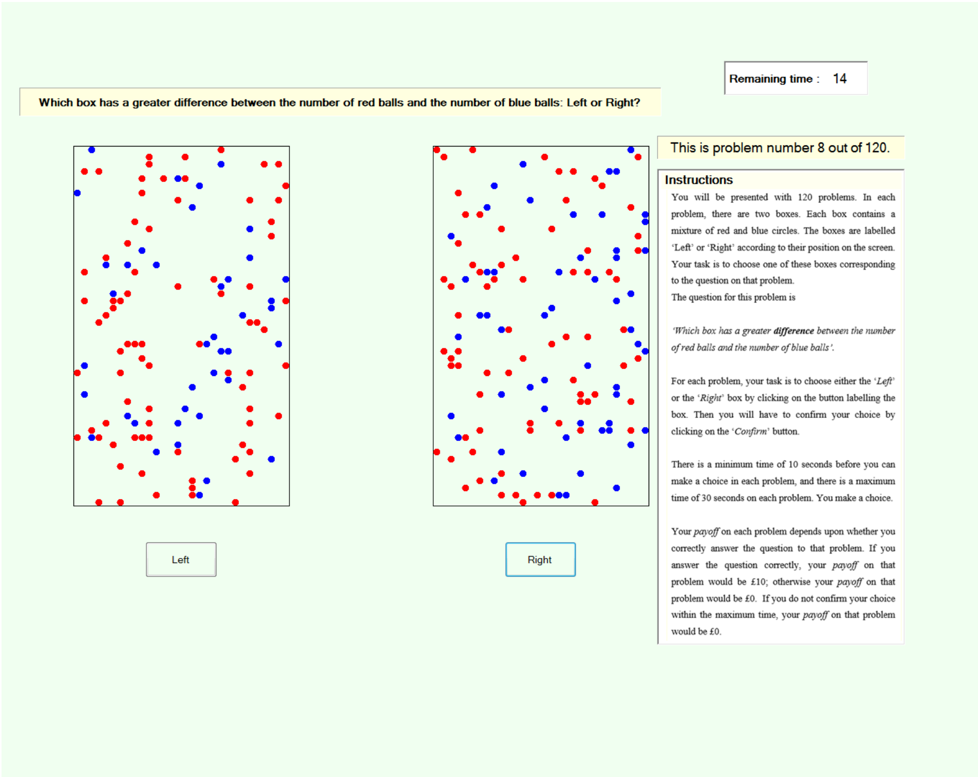


Figure 3.1: A screenshot of a problem from the first section of the experiment.

There were 16, 22 and 22 problems corresponding to Axioms 3, 4, and 5 respectively[47]. These make up 60 baseline problems; each of these was repeated twice, giving a total of 120 problems. Subjects were given written instructions which were read to them before starting the experiment. After the end of this first section of the experiment, subjects were given written instructions for the second section; these were also shown on their screens; they had to wait for at least five minutes before they could start the second section. This was a way of forcing them to read carefully the Instructions.

[47] Parameters for each problem can be found in Appendix 3B.

3.3.2 The experimental design in the second section

The second section of the experiment involved a set of risky choice problems. These were pairwise lottery choice problems aiming at investigating the relation between the salience axioms and violations of rational choice theory under risk. The lotteries in the problems were designed so that we could test the implications of the axioms. One of the applications of the salience function (implied by the salience axioms) is Salience Weighted Utility over Presentations model (SWUP) (Leland & Schneider, 2016b). The model is derived based on Expected Utility (EU) model with weights place on both probabilities and payoffs. Suppose there are two lotteries, A and B. There are finite set of outcomes denoted A_i and B_i, $i = 1,2,\dots,n$. Each A_i occurs with probability p_i and each B_i with probability q_i. A decision maker (DM) is strictly preferred A to B if and only if the following holds:

$$\Sigma_{i=1}^{n}\left[\phi(p_i, q_i)(p_i - q_i)\big(U(A_i) + U(B_i)\big)/2 + \mu(A_i, B_i)\big(U(A_i) - U(B_i)\big)(p_i + q_i)/2\right] > 0$$

This model assumes that a DM evaluate different attributes across lottery. Difference in attribute values are perceived according to salience perception and attract disproportionate attention. Therefore, they are weighted in the expected utility evaluation process. The weight $\phi(p_i, q_i)$ is placed on probability differences and $\mu(A_i, B_i)$ on payoff differences.

The implications of SWUP include the Fourfold Pattern of Risk Attitudes. Axioms 3 and 4 – operating through payoff salience functions that exhibit Diminishing Absolute Sensitivity (DAS) and Ordering - have implications that push DMs toward apparent risk aversion for high-probability gains. In contrast, Axiom 5 – which operates through payoff salience functions that exhibit Increasing Proportional Sensitivity (IPS) – has implications that push DMs toward apparent risk seeking for low-probability gains. In addition, Axiom 5 has also an implication of predicting the common ratio effect, operating through probability-weight salience function.

There were 36 problems in this section[48]. The first 26 of these were designed to detect the "Fourfold Pattern of Risk attitudes" which is the implication of salience perception on *lottery payoffs,* and the last 10 to detect the effect of IPS on choice.

The characteristics of IPS (and Ordering) determined the design of problems 1 to 12. SWUP predicts that, *assuming risk neutrality*, DMs would prefer the risky lottery over the certainty in these problems. For example, problem 1 is a choice between:

	A_1, B_1	P_{A1}, P_{B1}	A_2, B_2	P_{A2}, P_{B2}
A	100	0.01	0.01	0.99
B	1	0.01	1	0.99

SWUP implies that A is preferred to B if and only if $\mu(100,1)[u(100)-u(1)](0.01)+\mu(0.01,1)[u(0.01)-u(1)](0.99)$ is positive. Assuming risk-neutrality (as in Proposition 7 of Leland and Schneider (2016a)), the preference depends on whether $\mu(100,1)(100-1)(0.01) + \mu(0.01,1)(0.01-1)(0.99)$ is positive, or on whether $0.99\mu(100,1) - (0.9801)\mu(0.01,1)$ is positive. According to IPS, $\mu(100,1) > \mu(0.01,1)$. This makes the left hand side of the equation positive, and hence A is preferred to B.

The characteristics of DAS (and Orderings) determined the design of problems 13 to 26. SWUP predicts, *assuming risk-neutrality*, that DMs would appear to be risk averse and would prefer the certainty in these problems. For example, problem 26 is a choice between:

	A_1, B_1	P_{A1}, P_{B1}	A_2, B_2	P_{A2}, P_{B2}
A	100	0.99	0.01	0.01
B	99	0.99	99	0.01

SWUP implies that B is preferred to A if and only if $\mu(100,99)[u(100)-u(99)](0.99)+\mu(0.01,99)[u(0.01)-u(99)](0.01)$ is negative. Assuming risk-neutrality (as in Proposition 7 of Leland and Schneider (2016a)), preference depends on whether $\mu(100,99)(100-99)(0.99)+ \mu(0.01,99)(0.01-99)(0.01)$ is negative; or, on whether $0.99\mu(100,99) - (0.9899)\mu(0.01,0.99)$ is negative. According to DAS and Ordering,

[48] Parameters for this section can be found in Appendix 3B and 3C.

μ(0.01,99) > μ(100,99) which makes the equation negative and thus, B is preferred to A.

The characteristics of IPS determined the design of Problems 27-36[49]. Each lottery has one non-zero payoff in which we scaled the amount of the payoffs according to the common ratio of their relative probabilities. For example, problem 27 was:

	A_1, B_1	P_{A1}, P_{B1}	A_2, B_2	P_{A2}, P_{B2}
A	9	0.90	0	0.10
B	18	0.45	0	0.55

There are two sets of common ratio problems. Problems 27-32 have a common ratio of 2 while 33-36 have a common ratio of 3. Subjects who conform to Expected Utility Theory should not change their choice within a set. However, SWUP predicts a pattern of change in choice when the probabilities have been scaled down.

We used the random lottery incentive mechanism for this section. At the end of the experiment, each subject randomly selected a problem for their payment by drawing a disk from a bag containing 36 disks, numbered from 1 to 36. Their lottery choice in that problem was then played out for real by drawing from another bag containing 100 disks, numbered from 1 to 100. The total payment for the experiment is the sum of the payments from two sections plus £2.50 show-up fee. An example of a screenshot for the second section is shown in Figure 3.2

We recruited a total of 80 subjects for the experiment which was conducted in the EXEC Lab at the University of York. Subject's ages ranged from 18 to 44 years. 78 of whom were students and 2 reported themselves as a member of staff at the University of York. There were 54 females (67.50%) and 26 males (32.50%). The average total payment was £12.85. This experiment was run using purpose-written software written in Visual Studio.

[49] Parameters for each problem can be found in Appendix 3D.

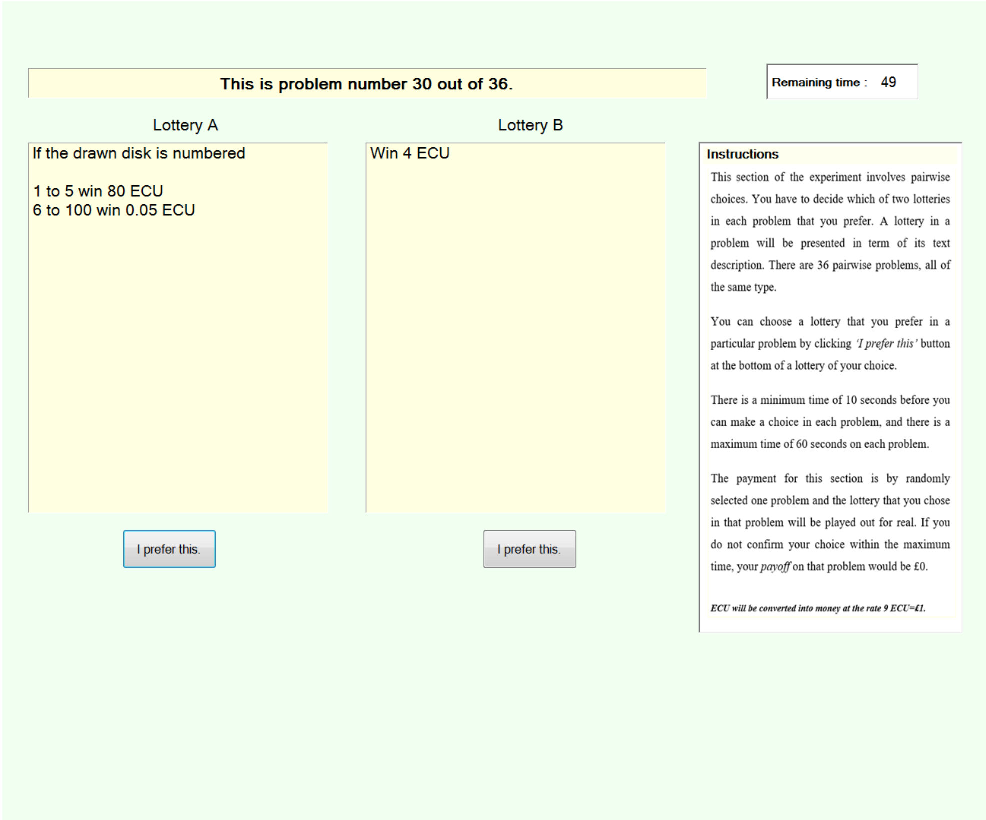


Figure 3.2: A screenshot of a problem from the second section of the experiment.

3.4 Results and Analyses

We begin by reporting the results of the first section – the direct tests of the axioms. Then we report the results from the second section of the experiment – investigating the axioms' implications for risky choice. Finally we report on the relationship between salience perception (section one) and behavioural biases in risky choice (section two). This analysis can be further broken down into (1) the relationship between axioms 4 and 5 and their implications on the effects of changes in the lottery payoffs, and (2) the relationship between axiom 5 and its implications on the effects of changes in the lottery probabilities. So we should be able to see which axioms are driving which behavioural biases.

3.4.1 Tests of Axioms for Salience Perception

First, we want to test and compare the strength of each axiom to determine their validity. In principle a single violation of an axiom is sufficient to discredit it, but that seems rather harsh. The 'benchmark' that we will use in measuring 'how good' is each axiom is by comparing the actual behaviour with random choice. A related benchmark that could be considered, particularly for DAS (Axiom 4) and IPS (Axiom 5) is to compare the violations with the 'constant sensitivity' cases. DAS predicts that, for a salience function σ and any $\varepsilon > 0$, $\sigma(x+\varepsilon, y+\varepsilon) < \sigma(x,y)$. On the contrary, for a salience function σ that satisfies *Constant Absolute Sensitivity* (CAS) and any $\varepsilon > 0$, $\sigma(x+\varepsilon, y+\varepsilon) = \sigma(x,y)$. Also, IPS predicts that, for a salience function σ and any $\alpha > 1$, $\sigma(\alpha x, \alpha y) > \sigma(x,y)$. On the contrary, for a salience function σ that satisfies *Constant Proportional Sensitivity* (CPS) and any $\alpha > 1$, $\sigma(\alpha x, \alpha y) = \sigma(x,y)$. In our experiment design, CAS and CPS would make the same predictions as random choice. Therefore, a distribution indistinguishable from random choice would falsify the axioms. Random choice suggests that the violation is equal to 50 percent for each axiom. Table 3.1 shows the average violation percentage for each axiom.

	Average	St. dev.	*p*-value
Axiom 3	0.1508	0.0975	0.000
Axiom 4	0.1210	0.1641	0.011
Axiom 5	0.2011	0.1493	0.023

Table 3.1: The average violation percentage for each axiom

We employ a simple t-test to compare the sample average with the hypothesised average of 50 percent under the benchmark of random choice. The alternative hypothesis is that $\mu < 0.5$. The *p*-value of the test (the sample size is 80) is reported in Table 3.1 The null hypothesis is rejected at the 5% significance level for every axiom. Thus, we can conclude that subjects were not behaving randomly. On this criterion all axioms are valid and not falsified.

Next, we calculate the *relative* violation rates of the three axioms. The two sample *t*-test for difference in means is employed. The results are presented in Table 3.2 As the table above shows, the average violation of Axiom 4 is the least and it is lower than Axiom 3 (at a 10% significance level) and Axiom 5 (at 1%). Axiom 5 is violated the most compared to the other two axioms (both at 1%). Interestingly, Axiom 4 which involves comparing the ratios is violated less than Axioms 3 and 5 which involve comparing differences.

	Axiom 3	Axiom 4	Axiom 5
Axiom 3		0.0825	0.0064
Axiom 4	0.0825		0.0000
Axiom 5	0.0064	0.0000	

Table 3.2: *p*-values for two sample *t*-test for difference in means

To understand why subjects may be making mistakes, we use the difference between the differences, or the ratios of the two boxes, in each problem as a measure of the degree of difficulty. We find that the correlations between the number of violations and the degree of difficulty are negative for all the axioms – which is as expected. Axiom 4 is the most negatively correlated. The table below shows the correlations.

	Correlation
Axiom 3	-0.263
Axiom 4	-0.699
Axiom 5	-0.362

Table 3.3: Correlation between violations and the degree of difficulty

3.4.2 Testing the implications of the axioms for risky choice

We now investigate the behavioural biases found in the results from the second section of the experiment. The first 26 problems were designed to test the implications of the Axioms on the Fourfold Pattern of Risk Preferences (Tversky and Kahneman 1992). In these problems, DAS (and Ordering) have implications for risky choices at moderate and high probabilities payoffs. This induces risk averse behaviour for moderate and high-probability gains. On the other hand, IPS (and Ordering) have implications for risky choices involving low probability payoffs, and induces risk seeking behaviour for long-shot lotteries. There were 5 problems that test Axioms 3 and 5, and there were 7 problems that test Axiom 5 by itself. In these problems (problems 1-12 in Appendix 3C), *assuming risk neutrality*, SWUP predicts that subjects should choose lottery A. There were 11 problems that involve testing Axioms 3 and 4, and 3 problems that test Axiom 4 independently. *Again assuming risk neutrality*, in these problems (problems 13-26 in Appendix 3C), SWUP predicts that subjects should choose lottery B. The percentages of violations of these predictions are reported in Table 3.4. The p-value in the final column tests the violation rate against the benchmark of random choice.

	Mean	St. dev.	*p*-value
Axioms 3 and 4 (problems 16-26)	0.1875	0.2032	0.063
Axiom 4 (problems 13-15)	0.3542	0.2722	0.291
Axioms 3 and 5 (problems 2,4,7,9,11)	0.3675	0.3093	0.334
Axiom 5 (problems 1,3,5,6,8,10,12)	0.5268	0.3299	0.468

Table 3.4: The average violation percentage according to SWUP predictions

Axiom 5 seems to be the weakest, and Axioms 3 and 4 are stronger. Although Axiom 5 has the weakest support, nearly 50% of the choices designed to test it do result in the choice of the risky lottery as predicted by Axiom 5. In contrast, even slightly risk-averse decision makers under the standard economic model of rational choice would select the safer lottery, suggesting that under the standard economic model with risk-averse agents, the violation percentage for Axiom 5 should be close to 100%. Under the salience model, choices under risk are determined by properties of perception (represented by salience functions) and the agent's risk preferences (represented by a utility function). Since concavity of the utility function (risk aversion) operates against Axiom 5, the violation rates of the salience model for subjects satisfying Axiom 5 should vary between those who are more sensitive to large differences (such that lottery choices are driven primarily by the salience function), and subjects who are very risk-averse (such that lottery choices are driven primarily by the utility function).

An obvious alternative story is that in fact the subjects are not SWUP agents who are risk-neutral, but are in fact EU subjects who are not risk-neutral. If they were the latter, they would not switch from A to B, but would either choose A throughout (if risk-averse) or B throughout (if risk-loving). The table below shows the violations of the SWUP and EU predictions.

Predictions	Violations
EU - Risk Averse	36.92%
EU - Risk Lover	63.08%
SWUP	33.27%

Table 3.5: Percentage of violations of three hypotheses

This table suggests that SWUP performs better than EU in explaining behaviour.

Risk-neutral SWUP predicts that subjects will choose A in problems 1 to 12 and B in problems 13 to 26. We test this against the alternative hypothesis that subjects are not risk-neutral SWUP agents but EU agents who are not risk-neutral. We have already tested this in our table above, but here we add a regression of the proportion choosing A (*P*)[50] against the problem number (*n*) with an interactive dummy, *d*, that takes the value 0 for problems 1 to 12, and the value 1 for problems 13 to 26:

$$P = \alpha + \beta d + \gamma n + \delta dn$$

If SWUP holds, $\alpha = 1$, $\beta = -1$, $\gamma = \delta = 0$; under EU α takes some value depending upon the risk-attitude of the subjects, $\beta = \gamma = \delta = 0$. The estimated relationship is:

[50] The proportion for each problem can be found in Appendix 3E.

$$P = 0.628 - 0.113d - 0.014n - 0.001dn$$
$$(0.075)\ (0.177)\quad (0.010)\quad (0.013) \qquad\qquad R^2 = 0.69$$

Standard errors are in parentheses. β, γ and δ are not significantly different from 0 while α is significant at 1% level and its value is 0.628. The regression suggests that SWUP does not do as well as the EU explanation.

Another alternative explanation is that subjects are SWUP agents who are *not* risk-neutral. In this case, SWUP predictions depend both on salience perception and the degree of risk aversion. The most risk-averse subject would choose B throughout, but agents who have a more moderate degree of risk aversion would choose A in the first few problems and switch to B at some point. We run a regression of the proportion choosing A (*P*) against the problem number (*n*) without a dummy to test this.

$$P = \alpha + \beta n$$

The hypotheses in this case are, $\alpha > 0$ and$\beta < 0$. The estimated relationship is:

$$P = 0.664 - 0.022n$$
$$(0.049)\quad (0.003) \qquad\qquad R^2 = 0.664$$

The estimated coefficient for the intercept is positive and significant. The slope coefficient is negative and significant indicating that subjects do choose A more often in the first few problems. SWUP with risk attitudes is a valid explanation of the data.

Next, we look at the common ratio type of problems which involve the implications of IPS related to probabilities. For this type, SWUP predicts that people who exhibit IPS are also more likely to exhibit the general common ratio effect. The problems are designed to detect a change in the pattern of choice from the first few problems to the last few. There are two sets of common ratios. The first set[51] are designed so that there is a common ratio in probabilities equal to 2 while the second set[52]

[51] Problems 27-32 in Appendix 3D.
[52] Problem 33-36 in Appendix 3D.

has a common ratio in probabilities equal to 3. The percentage of subjects who choose Lottery A and Lottery B is calculated and the results are shown in Table 3.6.

	A	B
27	0.725	0.275
28	0.688	0.313
29	0.738	0.263
30	0.725	0.275
31	0.638	0.363
32	0.500	0.500
33	0.700	0.300
34	0.663	0.338
35	0.650	0.350
36	0.688	0.313

Table 3.6: The percentage of subjects who choose respective lotteries in the common ratio type of problems

We do not observe the common ratio effect in aggregate. Subjects are consistent as the majority chose Lottery A over Lottery B for every problem except problem 32. Here the alternative hypothesis that subjects are EU agents who are not risk-neutral is supported by the data.

We now compare the goodness-of-fit of non-risk-neutral SWUP with non-risk-neutral EU subject by subject. We use Maximum Likelihood Estimation and assume the Constant Absolute Risk Aversion (CARA) functional form for EU. For non-risk-neutral SWUP, to estimate the SWUP equation, we assume a CARA utility function and the following salience functional forms:

$$\mu(x,y) = \frac{|x-y|}{(x+y+1)} \; and \; \varphi(p,q) = \frac{|p-q|}{(p+q)}$$

These specifications generate the Fourfold pattern of Risk Attitudes and the Allais common ratio effect[53] which are investigated in this paper. They are based on the parameter-free functional form introduced by Bordalo *et al* (2013b). The payoff salience functional is adopted so that it satisfies the IPS property. The IPS property for probability is needed for the generalisation of the common ratio effect to the case where all outcomes are risky. This effect is not a prevalent pattern in our data, thus, we maintain the same functional form used in Bordalo *et al* (2013b). Another advantage in using these form is that it has only one parameter (curvature of the utility function) which is comparable to EU.

We find that CARA SWUP has a higher log-likelihood than CARA EU for 65% of the subjects, and it has the same number of parameters. We perform a robustness check on these parameter-free salience functional forms by estimating SWUP using the form of the salience function on probabilities[54] suggested in equation (5) of Bordalo *et al* (2012).

$$\varphi(p,q) = \frac{|p-q|}{p+q+\varphi}$$

With an additional parameter on this functional form, we employ a likelihood ratio test to compare the two models. We found that the parameter-free CARA SWUP is better fit in 87.5% of the subjects. This confirms that the IPS parameter for probabilities is not significantly different from zero.

3.4.3 Relationships between salience perception and behavioural biases in risky choice

We now investigate whether there is a connection between the axioms of salience perception and the common behavioural biases or violations of rational choice

[53] As demonstrated in Leland and Schneider (2016b).
[54] The primary reason that our specification does not include IPS for probabilities is that it is not needed for any robust behavioural predictions (except for the general common ratio effect where all outcomes are risky). For this reason and for simplicity we do not use the parametric form for $\varphi(p,q)$ with the IPS property. Moreover, the certainty effect version of the common ratio effect is predicted even without IPS for the probability salience function.

theory under risk by finding the relationship between the first section and the second section of the experiment – starting from the first 26 problems that were designed to detect implications on payoffs, followed by problems 27 to 36 that were designed to detect common ratio effects. The data that we take from the two sections are the violations of the axioms (section 1) and from the predictions of SWUP for risk-neutral agents (section 2)[55]. We hypothesise that the more a subject violates the axioms' predictions in the first section, the more the subject should depart from the theoretical predictions in the second section as well. This is a strong prediction. The first section of the experiment did not involve risk perception but rather the perception of relationships between boxes of red and blue dots - a task far removed not only from choices under risk but from any standard decision environment studied in economics. It would be surprising if there is a systematic relationship between visual perception of red and blue dots and monetary decisions under risk.

Of the three relationships that we test (DAS for payoffs and aversion to negatively skewed lotteries, IPS for payoffs and preference for positively skewed lotteries, and IPS for probabilities)[56], there are a priori reasons why we would not expect to observe a relationship for DAS for payoffs or for IPS for probabilities. Under the salience model, risk aversion is determined by both DAS and curvature of the utility function. As a consequence, our design cannot separate these potential sources of risk aversion. In addition, in the salience model of Bordalo *et al* (2012), only the salience of possible outcomes matters (and not the salience of probabilities). It is demonstrated in Bordalo *et al* (2012) that sensitivity to the salience of payoffs is sufficient to explain the major anomalies for choices under risk. In contrast, earlier work Prelec and Loewenstein (1991) has proposed, but not tested, the hypothesis that IPS also applies to probabilities. Since IPS for probabilities is not necessary to generate observed behaviour, it might not significantly influence choices under risk.

[55] So that we have a clear measurement for violations in the second section.

[56] We need Ellsberg's paradox type of problem to test for DAS implication on probability. We do not want to involve ambiguity in our experiment. Schneider *et al* (2016) tested this.

Our design provides a more natural test of IPS for payoffs. For a linear or concave (risk-averse) utility function, the salience model predicts that any skewness preference observed must be due to IPS since concavity of the value function operates against IPS, and other properties of salience perception such as DAS also operate against IPS.

We first regress, subject by subject, the violation percentage of problems involving each Axiom from the first section against the deviations from theoretical predictions in corresponding problems from the second section for each subject. We regress V_2 against $V_1$where V_1 is violations from section 1 and V_2 is violations from SWUP predictions in section 2. The hypothesis is that β is positive and significant.

Axiom 4 (implications of DAS):

For this axiom, we regress the violation percentage from baseline problem 39-60 from Appendix 3B against violation in SWUP prediction in problem 13-26 from the second section (the implications of DAS for payoffs). The result is:

$$\begin{array}{ll} V_2 = \underset{(0.025)}{0.233} - \underset{(0.121)}{0.079}V_1 & \qquad R^2 = 0.0055 \end{array}$$

The slope coefficient is negative; however it is not significantly different from zero. We found no relationship between the two sections for this axiom. These findings are consistent with salience models in which risk preferences and DAS both contribute to risk aversion, but are inconsistent with salience models in which DAS is the only source of risk aversion.

Axiom 5 (implication of IPS):

Next, we take a look at the relationship between the violations of axiom 5 from the first section with problems involving common ratio problems in the second section (the implications of IPS for probabilities). Expected Utility Theory predicts a consistent choice pattern. However, a subject conforming to IPS would switch his/her choice. The inconsistencies display by a subject can be considered as a violation of EU. Therefore, our measure is to look at the percentage subject’s choice of B over A. We found that there are 26 out of 80 subjects (32.5 percent) and 37 out of 80 subjects (46.25 percent) who are consistent throughout the problems that have a common ratio equal to 2 and 3 respectively. This result is consistent with Table 3.6 where we do not observe a strong common ratio effect in this experiment at the average level. We can find the correlation between the inconsistencies or the common ratio effect in each subject and the violation percentage from the direct test of IPS in the first section by regressing V_2 against $V_1$where V_1 is the violations from section 1 and V_2 is the inconsistencies in section 2. If there is a relationship, we hypothesise that the more subject conform to IPS (less violations in the first section), the more inconsistencies he/she will make in the choice behaviour section (EU is violated more). Therefore, β is expected to be negative and significant. The result for those problems whose common ratio is equal to 3 is:

$$\begin{aligned} V_2 = \; & 0.262 - 0.138V_1 \\ & (0.051) \;\; (0.205) \qquad\qquad R^2 = 0.0058 \end{aligned}$$

And the result for those problems whose common ratio is equal to 2 is:

$$\begin{aligned} V_2 = \; & 0.246 + 0.166V_1 \\ & (0.050) \;\; (0.201) \qquad\qquad R^2 = 0.0087 \end{aligned}$$

On both type of problems, we fail to reject the null hypothesis that the slope coefficient is equal to zero. Therefore, we also do not find a significant relationship between the axiom violations from the direct test and the choice inconsistencies displayed by subjects. The finding for probabilities is consistent with salience models that assume only the salience of rewards (and not the salience of probabilities) affects economic decisions under risk.

For this axiom, we also regress, again subject by subject, the violation percentage from baseline problem 17-38 from Appendix 3B against violation in SWUP prediction in problem 1-12 from the second section (the implications of IPS for payoffs). The result is:

$$\begin{aligned} V_2 = \; & 0.362 + 0.491V_1 \\ & (0.053) \;\; (0.212) \qquad\qquad R^2 = 0.0643 \end{aligned}$$

The relationship here is positive and significant suggesting that a subject who violates more in the first section, also violates more from SWUP prediction with problems concerning axiom 5, Increasing Proportional Sensitivity. This indicates that subjects who are less likely to exhibit IPS in the round dot perceptual task are less likely to exhibit a preference for positively skewed lotteries in the decision task. This finding provides a link between a basic property of salience perception (IPS), and economic choices under risk (preference for skewness). This finding also indicates a link between the representation of non-symbolic and symbolic stimuli in decision making as salience perception of non-symbolic stimuli (larger differences between red and blue dots) is predictive of the salience perception of symbolic stimuli (larger differences between monetary rewards).

3.5 Conclusions

What have we learnt from the experiment? First, directly testing the axioms shows general support for them: we can conclude that the axioms in general are valid. The actual individual perceptions between pairs of quantities seem to be accurate with the characterisations. Second, we test the behavioural implications of the axioms (embedded in SWUP) in risky choice context and compare with the alternative of non-risk-neutral EU agents. We found that the CARA SWUP model is a plausible explanation and is a better fit than the CARA EU model. Applying these axioms into an application in a risky choice context in terms of Salience Weighted Utility over Presentations (SWUP) model, allowing risk aversion in the utility function, receives some statistical support from the data. Lastly, there is a modest connection between the violations of the axioms and the violations of the predictions. The finding provides a link between a basic property of salience perception (IPS) and risk seeking behaviour in long-shot lotteries.

Appendix 3

3A Proposition 1 for Axioms of Salience Perception:

Under Axioms 1-4, there exists a salience function σ that represents $\trianglerighteq_S$.

Proof: In this proposition, we establish that Axioms 1 through 5 are sufficient for the representation. Axiom 1, well known in consumer theory, guarantees the existence (see, for instance Mas-Colell, Whinston, and Green (1995), Ch. 3, Proposition 3.C.1) of a continuous function $\sigma: X \to \mathbb{R}$ such that $(x, y) \trianglerighteq_S (x', y') \Leftrightarrow \sigma(x, y) \geq \sigma(x', y')$. Given Axiom 1, it is clear that Axiom 2 implies that σ is symmetric, and that Axiom 3 implies that σ satisfies ordering. To show that in the presence of Axiom 1, Axiom 4 implies diminishing absolute sensitivity, we can write $x' = x + \epsilon$ and $y' = y + \epsilon$ for $\epsilon > 0$. Then we have the following lemma:

Lemma 1: $r(x, y) > r(x + \epsilon, y + \epsilon)$ for all $x > y > 0$, and any $\epsilon > 0$.

Proof: Inequality $r(x, y) > r(x + \epsilon\,, y + \epsilon)$ holds for all $x > y > 0$ and any $\epsilon > 0$ if

$$\frac{x(y+\epsilon)}{y(y+\epsilon)} > \frac{y(x+\epsilon)}{y(y+\epsilon)}.$$

which requires $xy^3 + 2xy^2\epsilon + xy\epsilon^2 > xy^3 + xy^2\epsilon + y^3\epsilon + y^2\epsilon^2$. Since $x > y$, we have $2xy^2\epsilon > xy^2\epsilon + y^3\epsilon$ and $xy\epsilon^2 > y^2\epsilon^2$. Thus, $r(x, y) > r(x + \epsilon, y + \epsilon) = r(x', y')$.■

By Axiom 4, the inequality $r(x, y) > r(x', y')$ implies $(x, y) \vartriangleright_S (x', y')$ which, in the presence of Axiom 1, implies $\sigma(x, y) > \sigma(x + \epsilon, y + \epsilon)$.

To show that in the presence of Axiom 1, Axiom 5 implies increasing proportional sensitivity, we can write $x' = \alpha x$ and $y' = \alpha y$ for $\alpha > 1$. Note that for any $\alpha > 1$, we have

$\Delta(x', y') = \alpha \cdot \Delta(x, y) > \Delta(x, y)$. By Axiom 5, $\Delta(x', y') > \Delta(x, y)$ implies$(x', y') \vartriangleright_S (x, y)$ which, by Axiom 1, implies $\sigma(\alpha x, \alpha y) > \sigma(x, y)$. ■

Proposition 2: *For any function σ that represents $\trianglerighteq_S$:*

(i) *σ satisfies ordering if and only if $\trianglerighteq_S$ satisfies Axiom 3.*

(ii) *σ satisfies DAS if and only if $\trianglerighteq_S$ satisfies Axiom 4.*

(iii) *σ satisfies IPS if and only if $\trianglerighteq_S$ satisfies Axiom 5.*

Proof: That Axioms 3, 4 and 5 are sufficient for ordering, DAS, and IPS, respectively was confirmed in Proposition 1. It remains for us to show that Axioms 3, 4 and 5 necessarily follow from the properties of a salience function. It is clear that Axiom 3 is implied by ordering. To see that DAS implies Axiom 4, recall that by Lemma 1, $r(x, y) > r(x + \epsilon\ , y + \epsilon)$ for any $x, y > 0$ and any $\epsilon > 0$. Also, note that $\Delta(x, y) = \Delta(x + \epsilon\ , y + \epsilon)$. By DAS, $\sigma(x, y) > \sigma(x + \epsilon, y + \epsilon)$ which implies $(x, y) \triangleright_S (x + \epsilon, y + \epsilon)$ for any σ that represents $\trianglerighteq_S$. Thus, we have $\Delta(x, y) = \Delta(x + \epsilon\ , y + \epsilon)$, and $r(x, y) > r(x + \epsilon\ , y + \epsilon)$ which, by DAS, imply $(x, y) \triangleright_S (x + \epsilon, y + \epsilon)$ and Axiom 4 follows.

To see that IPS implies Axiom 5, recall that $\Delta(x', y') = \alpha \cdot \Delta(x, y) > \Delta(x, y)$ for any $x, y > 0$ and any $\alpha > 1$. Also note that $r(x, y) = r(\alpha x, \alpha y)$. By IPS, $\sigma(\alpha x, \alpha y) > \sigma(x, y)$ which implies $(\alpha x, \alpha y) \triangleright_S (x, y)$ for any σ that represents $\trianglerighteq_S$. Thus, we have $r(x, y) = r(\alpha x, \alpha y)$, and $\Delta(\alpha x, \alpha y) > \Delta(x, y)$, which, by IPS, imply $(\alpha x, \alpha y) \triangleright_S (x, y)$ and Axiom 5 follows.

3B Parameters for the first section of the experiment.

		More Salient Box		Less Salient Box	
Problem No.	Axiom	Red	Blue	Red	Blue
1	3	100	20	70	50
2	3	100	20	95	25
3	3	80	20	50	40
4	3	80	20	70	25
5	3	80	40	61	41
6	3	80	40	51	41
7	3	80	40	75	65
8	3	80	40	75	55
9	3	80	40	65	55
10	3	40	20	30	25
11	3	40	10	35	15
12	3	40	10	38	13
13	3	20	10	16	14
14	3	10	5	8	6
15	3	10	5	9	6
16	3	10	2	5	4
17	5	80	40	60	30
18	5	80	40	20	10
19	5	60	30	40	20
20	5	40	20	20	10
21	5	20	10	18	9
22	5	20	10	10	5
23	5	60	15	40	10
24	5	60	15	16	4
25	5	100	20	50	10
26	5	100	20	80	16
27	5	10	2	5	1
28	5	75	10	60	8
29	5	75	10	45	6
30	5	75	10	30	4
31	5	75	10	15	2
32	5	80	10	64	8
33	5	64	8	40	5
34	5	80	10	32	4
35	5	40	5	16	2

36	5	40	5	32	4
37	5	36	4	18	2
38	5	18	2	10	1
39	4	100	20	110	30
40	4	100	20	130	50
41	4	80	10	140	70
42	4	80	10	90	20
43	4	80	20	100	40
44	4	80	20	85	25
45	4	80	40	85	45
46	4	80	40	100	60
47	4	60	10	70	20
48	4	60	10	100	50
49	4	40	5	42	7
50	4	40	5	80	45
51	4	20	10	25	15
52	4	20	10	70	60
53	4	20	2	22	5
54	4	20	2	40	22
55	4	10	2	11	3
56	4	10	2	15	7
57	4	10	2	20	12
58	4	10	2	30	22
59	4	10	2	50	42
60	4	10	2	100	98

3C Fourfold Pattern of Risk Preferences type of problems.

	Lottery A					Lottery B (Certainty)				
	State 1		State 2			State 1		State 2		
Problem No.	P1	A1	P2	A2	E	P1	B1	P2	B2	E
1	0.01	100	0.99	0.01	1.01	0.01	1	0.99	1	1.00
2	0.01	100	0.99	1	1.99	0.01	2	0.99	2	2.00
3	0.01	200	0.99	0.02	2.02	0.01	2	0.99	2	2.00
4	0.02	100	0.98	1	2.98	0.02	3	0.98	3	3.00
5	0.02	250	0.98	0.1	5.10	0.02	5	0.98	5	5.00
6	0.03	100	0.97	0.09	3.09	0.03	3	0.97	3	3.00
7	0.05	100	0.95	5	9.75	0.05	10	0.95	10	10.00
8	0.05	80	0.95	0.05	4.05	0.05	4	0.95	4	4.00
9	0.1	50	0.9	11	14.90	0.1	15	0.9	15	15.00
10	0.2	45	0.8	5	13.00	0.2	15	0.8	15	15.00
11	0.25	40	0.75	12	19.00	0.25	20	0.75	20	20.00
12	0.3	45	0.7	5	17.00	0.3	15	0.7	15	15.00
13	0.5	20	0.5	10	15.00	0.5	15	0.5	15	15.00
14	0.5	70	0.5	1	35.50	0.5	35.5	0.5	35.5	35.50
15	0.5	50	0.5	2	26.00	0.5	26	0.5	26	26.00
16	0.5	50	0.5	2	26.00	0.5	27	0.5	27	27.00
17	0.6	50	0.4	1	30.40	0.6	30	0.4	30	30.00
18	0.75	55	0.25	0.5	41.38	0.75	50	0.25	50	50.00
19	0.75	48	0.25	1	36.25	0.75	36	0.25	36	36.00
20	0.8	75	0.2	0.1	60.02	0.8	60	0.2	60	60.00
21	0.85	80	0.15	1	68.15	0.85	68	0.15	68	68.00
22	0.9	85	0.1	0.05	76.51	0.9	80	0.1	80	80.00
23	0.95	90	0.05	1	85.55	0.95	85	0.05	85	85.00
24	0.95	100	0.05	0.01	95.00	0.95	95	0.05	95	95.00
25	0.99	95	0.01	1	94.06	0.99	94	0.01	94	94.00
26	0.99	100	0.01	0.01	99.00	0.99	99	0.01	99	99.00

3D Common ratio type of problems.

	Lottery A						Lottery B					
	State 1		State 2				State 1		State 2			
Problem No.	P1	A1	P2	A2	E		P1	B1	P2	B2	E	Ratio
27	0.9	9	0.1	0	8.1		0.45	18	0.55	0	8.1	2
28	0.8	9	0.2	0	7.2		0.4	18	0.6	0	7.2	2
29	0.6	9	0.4	0	5.4		0.3	18	0.7	0	5.4	2
30	0.4	9	0.6	0	3.6		0.2	18	0.8	0	3.6	2
31	0.2	9	0.8	0	1.8		0.1	18	0.9	0	1.8	2
32	0.02	9	0.98	0	0.18		0.01	18	0.99	0	0.18	2
33	0.9	6	0.1	0	5.4		0.3	18	0.7	0	5.4	3
34	0.6	6	0.4	0	3.6		0.2	18	0.8	0	3.6	3
35	0.8	6	0.2	0	4.8		0.27	18	0.73	0	4.86	3
36	0.4	6	0.6	0	2.4		0.13	18	0.87	0	2.34	3

3E The percentage of subjects who choose the two lotteries in the Fourfold Pattern of Risk Preferences type of problems.

Problem No.	A	B
1	0.538	0.463
2	0.738	0.263
3	0.488	0.513
4	0.675	0.325
5	0.538	0.463
6	0.513	0.488
7	0.550	0.450
8	0.388	0.613
9	0.700	0.300
10	0.375	0.625
11	0.500	0.500
12	0.475	0.525
13	0.638	0.363
14	0.200	0.800
15	0.225	0.775
16	0.200	0.800
17	0.288	0.713
18	0.013	0.988
19	0.250	0.750
20	0.200	0.800
21	0.275	0.725
22	0.113	0.888
23	0.250	0.750
24	0.075	0.925
25	0.200	0.800
26	0.200	0.800

3F Instructions for the first section.

centre for experimental economics

Instructions

Preamble

Welcome to this experiment. Thank you for coming. Please read very carefully these instructions. They are to help you to understand what you will be asked to do and how will you get paid. The experiment is simple and gives you the chance to earn money. You will be paid in cash immediately after the experiment is completed.

The Experiment

The experiment is interested in how you take decisions. This is an individual decision making experiment. Your decision will not affect the payoff of the others nor do their decisions affect yours. The experiment is separated into two sections. For the first section, you will be presented with a series of 120 problems, while in the second section, you will be presented with a series of 36 problems. Details of the first section are given in these instructions. Details of the second section will be presented to you after you finish the first section.

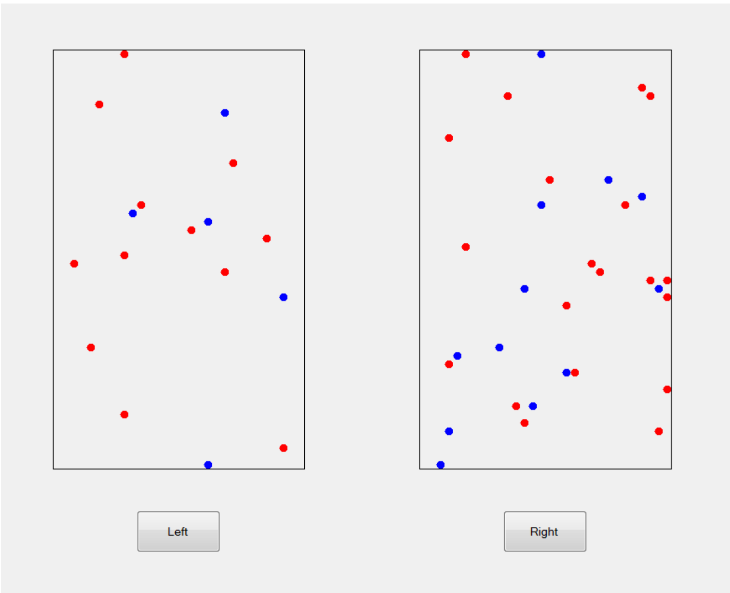
Your payment for the first section will be the *average payoff* from the 120 problems in this section.

Your total payment for this experiment will be the payment from this first section plus the payment from the second section, rounded up to the nearest 10p. In addition, you will also be given a £2.50 show-up fee.

The first section

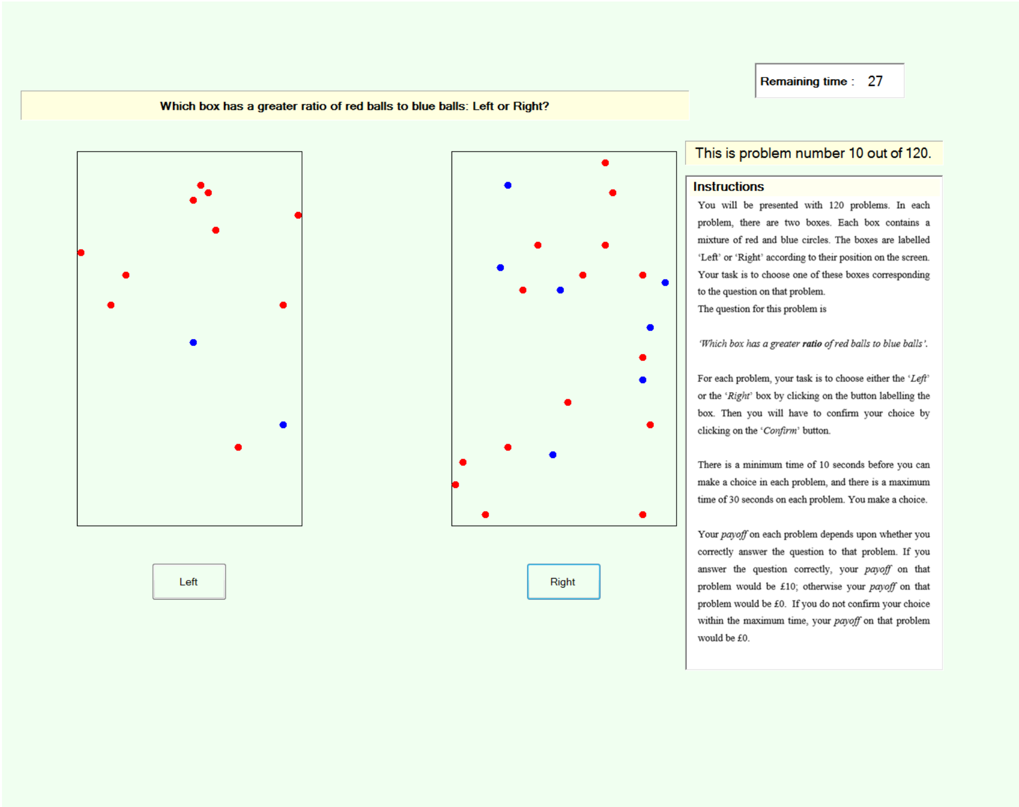
In each problem, there are two boxes. Each box contains a mixture of red and blue circles. Each box is labelled ‘Left’ or ‘Right’ according to its position on the screen. Your task is to choose one of these boxes corresponding to the question in that problem. There are two types of question, namely, ‘Which box has a greater ratio of red balls to blue balls’ and ‘Which box has a greater difference between the number of red balls and the number of blue balls’. **Your *payoff* on each problem depends upon whether you correctly answer the question to that problem. If you answer the question correctly, your *payoff* on that problem is £10; otherwise your *payoff* on that problem is £0.**

The visual representation of the box will be like the two examples below. Examples of a screenshot of problems will be given next in the ‘Example’ section.



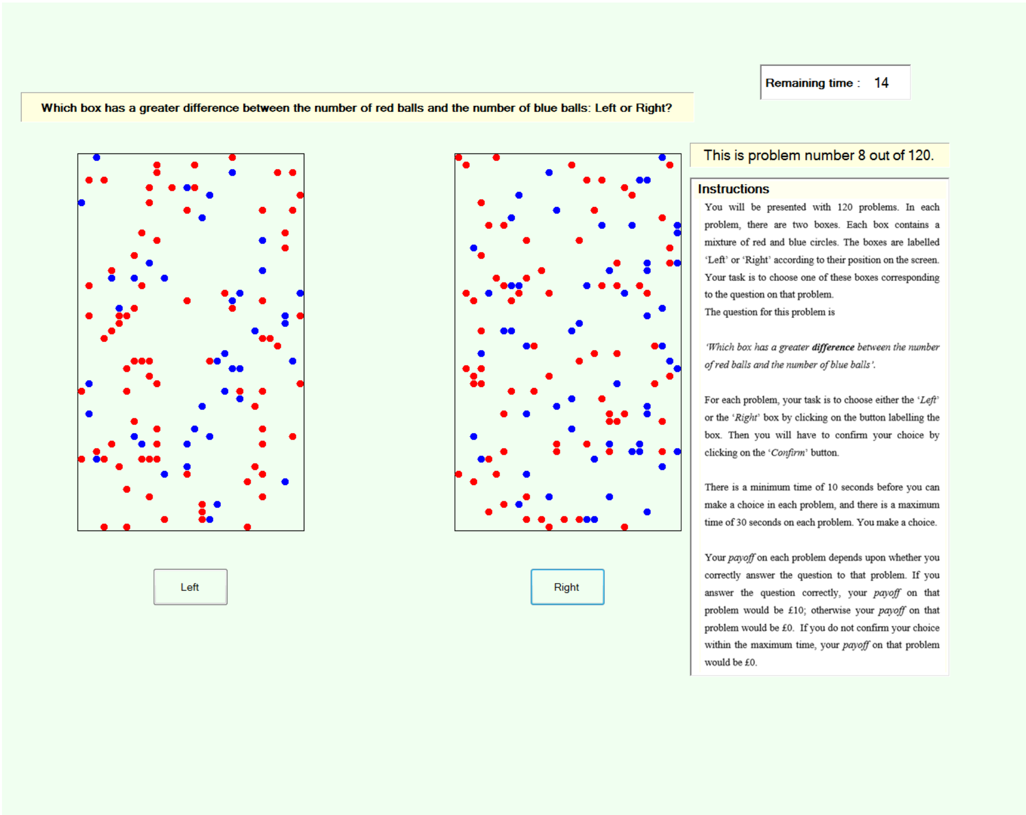
Example

For each problem, your task is to choose either the ‘*Left*’ or the ‘*Right*’ box according to the question in that problem. There are two types of questions. The first type is *‘Which box has a greater ratio of red balls to blue balls’*. A screenshot of a problem of this type is shown below.



Your task for this type of question is to identify which box has a greater ratio of red balls to blue balls. The ratio of red balls to blue balls is defined as *the number of red balls in that box divided by the number of blue balls in that box.* So, for example, in this particular problem, the '*Left*' box has a greater ratio. If you choose the '*Left*' box, your *payoff* on this problem would be £10. If you choose the '*Right*' box, your *payoff* on this problem would be £0.

The second type of question is *'Which box has a greater difference between the number of red balls and the number of blue balls'*. A screenshot of a problem of this kind is shown below.



Your task for this type of question is to identify which box has a greater difference between the number of red balls and the number of blue balls. The difference between the number of red balls and the number of blue balls is simply defined by the number of red balls in that box minus the number of blue balls in that box. So, for example, in this particular problem, the '*Left*' box has a greater difference. If you choose the '*Left*' box, your *payoff* on this problem would be £10. If you choose the '*Right*' box, your *payoff* on this problem would be £0.

<u>Choices</u>

For each problem, your task is to choose either the '*Left*' or the '*Right*' box corresponding to that problem's question. There is a minimum time of 10 seconds before you can make a choice in each problem. There is also a maximum time of 30 seconds that you can make a choice in each problem. You can make a choice by clicking at the button labelling the box. Then you will have to confirm you choice by clicking the '*Confirm*' button. If you do not make a choice within the maximum time, your *payoff* on that problem would be £0.

What happens next

When we finish reading these Instructions and have answered any questions that you may have, we will start the first section of the experiment. Each problem screen has a countdown timer at the top right corner of the screen. You cannot confirm your choice any until 10 seconds have passed from the start of that problem. There is a time limit of 30 seconds to make a decision on any problem in this section. You can change your decision as many times as you want during this time period before clicking the '*Confirm*' button. You can click the *'Confirm*' button before the time limit is reached. Once you click *'Confirm*' button, that problem is over and you will be immediately led to the next problem. **If you do not click the *'Confirm*' button before the time limit is over, your *payoff* on that problem will be £0.**

When you finish this section, please click the *'Continue to the second section'* button at the bottom of your screen, it will lead you to the instructions for the second section of the experiment. Please read the instructions very carefully as it will affect your income from the experiment. If you have any questions, please raise your hand and an experimenter will come to you.

If you have any questions at any stage of the experiment, please raise your hand and an experimenter will come to you.

Thank you for your participation.

John Hey
Nuttaporn Rochanahastin

November 2017

3G Instructions for the second section.

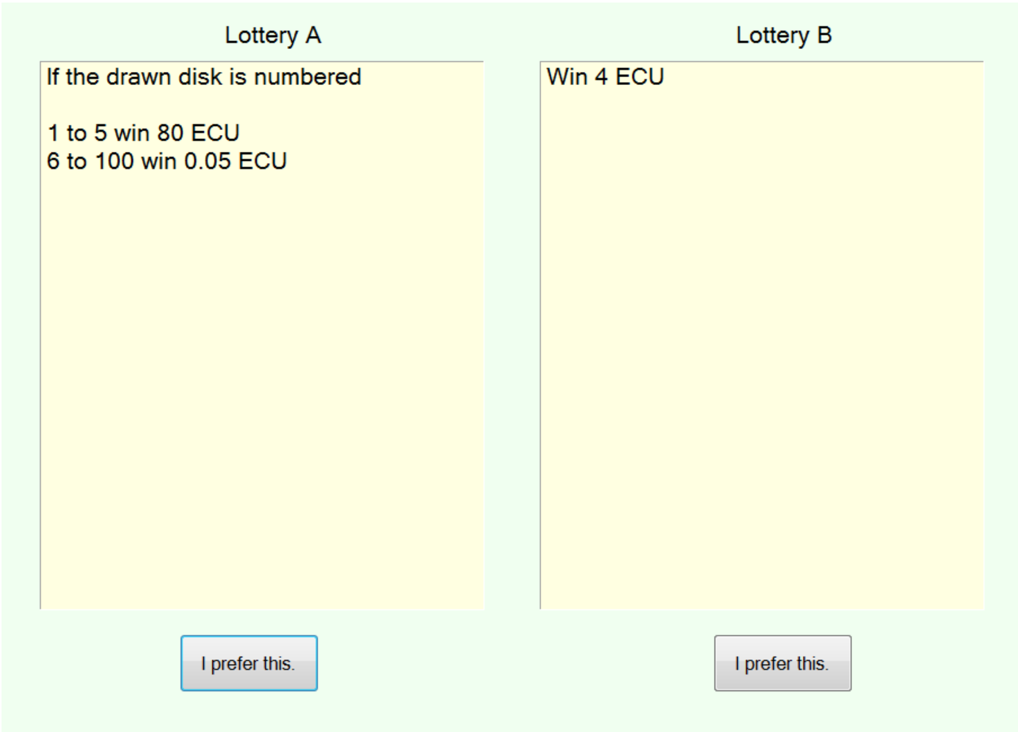
Instructions

The second section

This section of the experiment involves pairwise choices. A pairwise choice is a choice between two lotteries. There are 36 pairwise problems, all of the same type. In each problem, you have to decide which of two lotteries you prefer. A lottery in a problem will be presented in terms of a written description. Each lottery involves either one or two possible outcomes. All outcomes are either zero or positive amounts. The payment for this section will be implemented by you randomly selecting one problem. Then the lottery that you chose in that problem will be played out for real.

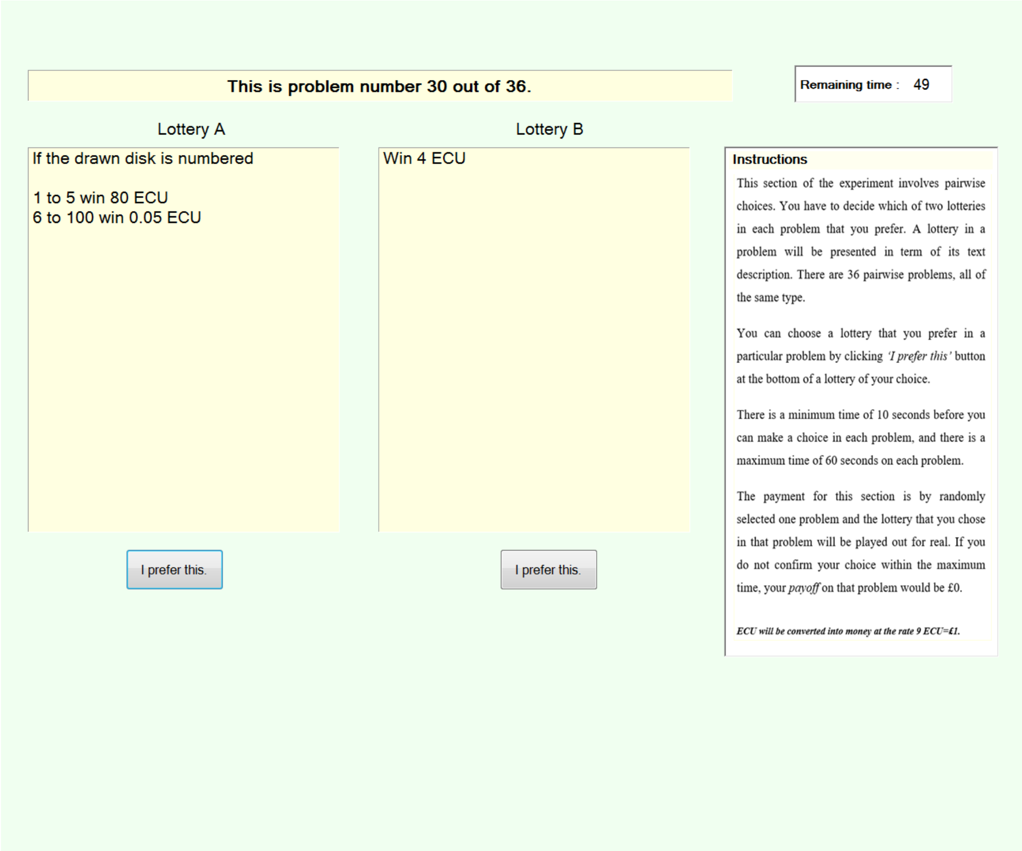
Example

The visual representation of a pairwise choice will be like an example below.



There will be two boxes. Outcomes are presented in Experimental Currency Units (ECU); we will tell you the exchange rate between ECU and real money at the end of these Instructions. Each box represents a lottery. In this example, the left lottery gives you a 5 per cent chance of a payoff of 80 ECU and a 95 per cent chance of a payoff of 0.05 ECU; it means that your payoff would be either 80 ECU or 0.05 ECU. For the right 'lottery', it leads to a payoff of 4 ECU with certainty. You can choose a lottery that you prefer in a particular problem by clicking *'I prefer this'* button at the bottom of a lottery of your choice.

A screenshot of a problem of this type is shown below.



<u>Payment</u>

When you complete the 36 problems in this section, please raise your hand and an experimenter will come to you. You will be led to a separate room where the payment will take place.

Your choice in a *randomly selected* problem will determine your payment for this section. You will be presented with a closed bag containing the numbered disks from 1 to 36. You will draw a disk to determine a randomly selected problem. We will recall your choice on that problem. If your choice was a lottery you will play out the lottery for real; if your choice was a certainty, you will be paid that certainty.

<u>How Is a Lottery Played Out?</u>

A lottery has two outcomes *X* ECU and *Y* ECU with respective probabilities *p* and *1-p.* You will play it out by drawing one disk at random out of a bag containing disks numbered from 1 to 100.

Let us give an example. Suppose the lottery gives you a 50 percent chance of a payoff of 20 ECU and a 50 percent chance of a payoff of 10 ECU. If the disk you draw from the bag is numbered from 1 to 50 then your payoff would be 20 ECU; if the disk you draw from the bag is numbered from 51 to 100 then your payoff would be 10 ECU.

ECU will be converted into money at the rate 9 ECU=£1.

As we have already noted, your payment for the experiment as a whole will be the sum of the payments from each section plus a £2.50 show-up fee.

<u>What you should do next</u>

When you finish reading these Instructions, you should click on the '*start*' button at the bottom of the screen (you will not be able to click this button until at least 5 minutes have passed). This will lead you to the experimental problems; you will then be starting the second section of the experiment. Each problem screen has a countdown timer at the top right corner of the screen. You cannot confirm your choice any until 10 seconds have passed from the start of that problem. There is a time limit of 60 seconds to make a decision on any problem in this section. You can change your decision as many times as you want during this time period before clicking the '*Confirm*' button. You can click the *'Confirm*' button before the time limit is reached. Once you click *'Confirm*' button, that problem is over and you will be immediately lead to the next problem. If you do not click the *'Confirm*' button before the time limit is over, your payoff on that problem will be 0 ECU.

If you have any questions at any stage of the experiment, please raise your hand and an experimenter will come to you.

Thank you for your participation.

John Hey

Nuttaporn Rochanahastin

November 2017

References

Abdellaoui, M., Baillon, A., Placido, L., & Wakker, P. (2011). The rich domain of uncertainty: source functions and their experimental implementation. *American Economic Review*, *101*, 695–723.

Ahn, D.S., Choi S., Gale, D., & Kariv, S. (2010). Estimating ambiguity aversion in a portfolio choice experiment. Working Paper.

Barberis, N., & Huang, M. (2008). Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review, 98*(5), 2066-2100.

Bone, J., Hey, J. D., & Suckling, J. (2009). Do people plan?. *Experimental Economics*, *12*(1), 12-25.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience Theory of Choice Under Risk. *Quarterly Journal of Economics, 127*(3), 1243-1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013a). Salience and Asset Prices. *American Economic Review, 103*(3), 623-628.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013b). Salience and Consumer Choice. *Journal of Political Economy, 121*(5), 803-843.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2015). Salience Theory of Judicial Decisions. *Journal of Legal Studies, 44*(S1), S7-S33.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2016). Competition for Attention. *The Review of Economic Studies, 83*(2), 481-513.

Bouacida, E., & Martin, D. (2017). Predictive Power in Behavioral Welfare Economics. Working paper.

Caplin, A. (2016). Measuring and modeling attention. *Annual Review of Economics*, *8*, 379-403.

Caplin, A., & Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review, 105*, 2183-2203.

Caplin, A., & Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics, 6*(1), 19-48.

Caplin, A., Dean, M. & Martin, D. (2011). Search and Satisficing. *American Economic Review, 107*, 2899-2922.

Cettolin, E., & Riedl, A. (2016). Revealed Incomplete Preferences under Uncertainty: Evidence for Bewley Preferences. Working Paper.

Cherepanov, V., Feddersen, T., & Sandroni, A. (2013). Rationalization. *Theoretical Economics, 8*(3), 775-800.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*, 1145-1177.

Conlisk, J. (1996). Why Bounded Rationality?. *Journal of Economic Literature, 34*, 669-700.

Costa-Gomes, M., Cueva, C., Gerasimou, G., &Tejiscak, M. (2014). Choice, Deferral and Consistency. School of Economics and Finance Discussion Paper No. 1416, University of St Andrews.

Danan, E., & Ziegelmeyer, A. (2006). Are Preferences Complete? An Experimental Measurement of Indecisiveness under Risk. Working Paper, 0601, Max Planck Institute of Economics, Strategic Interaction Group.

De Clippel, G., & Rozen, K. (2014). Bounded Rationality and Limited Datasets. Cowles Foundation Discussion Paper No. 1853.

De Los Santos, B., Hortaçsu, A., & Wildenbeest, M.R. (2012). Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. *American Economic Review, 102*, 2955-2980.

Dehaene, S. (2011). *The Number Sense*: Oxford University Press.

Dhami, S. (2016). *The Foundations of Behavioral Economic Analysis*: Oxford University Press.

Dubra, J., Maccheroni, F., & Ok, E.A. (2004). Expected Utility Theory without the Completeness Axiom. *Journal of Economic Theory, 115,* 118-133.

Etner, J., Jeleva, M., & Tallon, J.M. (2012). Decision theory under ambiguity. *Journal of Economic Surveys*,*26,* 234–270.

Friedman, M., & Savage, L. J. (1948). The Utility Analysis of Choices Involving Risk. *Journal of Political Economy, 56*(4), 279-304.

Gajdos, T., Hayashi, T., Tallon, J.M., & Vergnaud, J.C. (2008). Attitude toward imprecise information. *Journal of Economic Theory*, *140,* 27–65.

Ghirardato, P., Maccheroni, F., & Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, *118*, 133–173.

Grether, D. M. (1978). Recent psychological studies of behavior under uncertainty. *The American Economic Review*, *68*(2), 70-74.

Grether, D. M., & Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, *69*(4), 623-638.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665.

Halevy, Y. (2007). Ellsberg revisited: an experimental study. *Econometrica*, *75*, 503–536.

Hayashi, T., & Wada, R. (2010). Choice with imprecise information: an experimental approach. *Theory and Decision, 69,* 355-373.

Hey, J.D., & Pace, N. (2014). The explanatory and predictive power of non two-stage-probability theories of decision making under ambiguity. *Journal of Risk and Uncertainty*, *49*, 1-29.

Hey, J.D., Lotito, G., & Maffioletti, A. (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty*, *41,* 81–111.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*(2), 263-292.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A Smooth Model of Decision Making Under Ambiguity. *Econometrica*, *73*, 1849-1892.

Kőszegi, B., & Szeidl, A. (2013). A Model of Focusing in Economic Choice. *The Quarterly Journal of Economics, 128*(1), 53-104.

Leland, J. W., & Schneider, M. (2015). Salience and Strategy Choice in 2× 2 Games. *Games, 6*(4), 521-559.

Leland, J. W., & Schneider, M. (2016a). *Axioms for Salience Perception*. Chapman University. Working Paper.

Leland, J. W., & Schneider, M. (2016b). *Salience, Framing, and Decisions under Risk, Uncertainty, and Time*. Chapman University. Working Paper.

Leland, J. W., & Schneider, M. (2018). A Theory of Focal Points in 2x2 Games. *Journal of Economic Psychology, 65*, 75-89.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory, 170*, 70-85.

Mandler, M. (2005). Incomplete Preferences and Rational Intransitivity of Choice. *Games and Economic Behavior, 50,* 255-277.

Manski, C.F. (2017). Optimize, Satisfice, or Choose without deliberation? A Simple Minimax-Regret Assessment. *Theory and Decision, 83*(2), 155-173.

Manzini, P., & Mariotti, M. (2007). Sequentially rationalizable choice. *American Economic Review, 97*(5), 1824-1839.

Manzini, P., & Mariotti, M. (2010). Revealed preferences and boundedly rational choice procedures: an experiment. *Unpublished Paper*.

Manzini, P., & Mariotti, M. (2012). Categorize then choose: boundedly rational choice and welfare. *Journal of the European Economic Association, 10*(5), 1141-1165.

Manzini,P., & Mariotti, M. (2014). Stochastic choice and consideration sets", *Econometrica*, *82*, 1153-1176

Manzini, P., Mariotti, M., & Tyson, C. J. (2013). Two-stage threshold representations. *Theoretical Economics, 8*(3), 875-882.

Masatlioglu, Y., & Nakajima, D. (2013). Choice by iterative search. *Theoretical Economics*, *8*, 701-728.

Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed Attention. *American Economic Review, 102*(5), 2183-2205.

Matejka, F., & McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, *105*, 272-298.

McCall, J.J. (1970). Economics of Information and Job Search. *Quarterly Journal of Economics 84*, 113–26.

Morgan, P., & Manning, R. (1985). Optimal Search. *Econometrica 53*, 923–44.

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature, 215*, 1519.

Nau, R. (2006). The Shape of Incomplete Preferences. *The Annals of Statistics, 34*, 2430-2448.

Ok, E.A., Ortoleva, P., & Riella, G. (2012). Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs versus Tastes. *Econometrica, 80*, 1791-1808.

Papi**,** M. (2012). Satisficing choice procedures. *Journal of Economic Behavior & Organization, 84*, 451-462.

Prelec, D., & Loewenstein, G. (1991). Decision Making over Time and under Uncertainty: A Common Approach. *Management Science, 37*(7), 770-786.

Reutskaja, E., Nagel ,R., Camerer, C.F., & Rangel, A. (2011). Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study. *American Economic Review, 101*, 900-926.

Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, *5*(17), 61-71.

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

Schley, D. R., & Peters, E. (2014). Assessing "economic value" symbolic-number mappings predict risky and riskless valuations. *Psychological science, 25*(3), 753-761.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57,* 571–587.

Schneider, M., Leland, J. W., & Wilcox, N. (2016). *Ambiguity framed*. Chapman University. Working Paper.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences, 21*(2), 153-167.

Selten, R. (1998). Aspiration Adaption Theory. *Journal of Mathematical Psychology*, *42*, 191-214.

Simon, H.A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics, 69*, 99-118.

Sims, C. A. (1998). Stickiness. Paper presented at the Carnegie-Rochester Conference Series on Public Policy.

Sims, C.A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*, 665-690.

Sims, C. A. (2010). Rational inattention and monetary economics. *Handbook of Monetary Economics* (Vol. 3, pp. 155-181): Elsevier.

Siniscalchi, M. (2009). Vector expected utility and attitudes toward variation. *Econometrica*, *77,* 801–855.

Stecher, J., Shields, T., & Dickhaut, J. (2011). Generating Ambiguity in the Laboratory. *Management Science, 57*, 705-712.

Stigler, G.J. (1961). The Economics of Information. *Journal of Political Economy*, *69*, 213–25.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297-323.

Tyson, C. J. (2008). Cognitive constraints, contraction consistency, and the satisficing criterion. *Journal of Economic Theory, 138*(1), 51-70.

Tyson, C. J. (2013). Behavioral implications of shortlisting procedures. *Social Choice and Welfare, 41*(4), 941-963.

Tyson, C. J. (2015). Satisficing behavior with a secondary criterion. *Social Choice and Welfare, 44*(3), 639-661.

Wakker, P. P. (2010). *Prospect Theory: For Risk and Ambiguity*. Cambridge: Cambridge University Press.

Weitzman, M. (1965). Utility Analysis and Group Behavior: An Empirical Study. *Journal of Political Economy, February*.