OPEN ACCESS

University of BRISTOL

This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Fennell, John

Title:
An expected utility theory that matches human performance

General rights
The copyright of this thesis rests with the author, unless otherwise identified in the body of the thesis, and no quotation from it or information derived from it may be published without proper acknowledgement. It is permitted to use and duplicate this work only for personal and non-commercial research, study or criticism/review. You must obtain prior written consent from the author for any other use. It is not permitted to supply the whole or part of this thesis to any other person or to post the same on any website or other online location without the prior written consent of the author.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to it having been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you believe is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact: open-access@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline of the nature of the complaint

On receipt of your message the Open Access team will immediately investigate your claim, make an initial judgement of the validity of the claim, and withdraw the item in question from public view.

An Expected Utility theory that matches human performance

John Fennell

School of Experimental Psychology

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science.

School of Experimental Psychology

March 2012

Word count: 53,554

I would like to dedicate this thesis to Anna, my wife, and my family without whose indulgence and support my adventure in the world of academia would not have been possible.

Acknowledgements

Solvitur ambulando is a latin phrase that literally means 'the problem is solved by walking' and I want to use it by way of offering my sincere thanks to Dr. Roland Baddeley for his help, guidance and support through this process. *Solvitur ambulando* relates to Roland in several ways; first, the literal meaning relates to the bootleather that we have used doing many circuits around Priory Road discussing problems and solutions; in a second sense *Solvitur ambulando* is a straight forward appeal to practical experience for the solution of a problem, which alludes to a proof of the possibility of motion attributed to Diogenes the Cynic, but which might also be translated as 'you'll find the answer as you go' and sums up the space and freedom I have been allowed in order to develop and complete this thesis. Roland, Thank you.

Abstract

Maximising expected utility has long been accepted as a valid model of rational behaviour, however, it has limited descriptive accuracy simply because, in practice, people do not always behave in the prescribed way. This is considered evidence that either people are not rational, expected utility is not an appropriate characterisation of rationality, or combination of these. This thesis proposes that a modified form of expected utility hypothesis is normative, suggesting how people ought to behave and descriptive of how they actually do behave, provided that: *a*) most utility has no meaning unless it is in the presence of potential competitors; *b*) there is uncertainty in the nature of competitors; *c*) statements of probability are associated with uncertainty; *d*) utility is marginalised over uncertainty, with framing effects providing constraints; and that *e*) utility is sensitive to risk, which, taken with reward and uncertainty suggests a three dimensional representation. The first part of the thesis investigates the nature of reward in four experiments and proposes that a three dimensional reward structure (reward, risk, and uncertainty) provides a better description of utility than reward alone. It also proposes that the semantic differential, a well researched psychological instrument, is a representation or description of the reward structure. The second part of the thesis provides a mathematical model of a value function and a probability weighting function, testing them together against extant problem cases for decision making. It is concluded that utility, perhaps more accurately described as advantage in the present case, when construed as three dimensions and the result of a competition, provides a good explanation of many of the problem cases that are documented in the decision making literature.

Declaration

I declare that the work in this thesis was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the thesis are those of the author.

SIGNED: .... DATE: ....

Contents

List of Figures

List of Tables

Chapter 1

Introduction

It is uncontroversial that some form of preferences must have existed prior to the evolution of intelligence or rationality and that an evolutionary process cannot produce the cognitive machinery to input, reason with or generally use information in a format that was not available from the environment (Cosmides & Tooby, 1996; Robson, 2001). It is reasonable to suppose that one of the simplest forms of information available is and would have been in our evolutionary past, the frequency of good and bad things that occur. Simply put, things that are perceived to make your world better are good and things that are perceived to make your world worse are bad. In order to make choices about what may or may not make our worlds better or worse, a forecast or prediction about what will happen is needed.

It may be thought that reducing prior experience to the apparent use of two outcomes is too limited, however, it is argued that this is not the case. This thesis proceeds on the basis that peoples choices are determined by their beliefs and desires; that is, the things that are likely to be desired and thereby gain them advantage are 'good', otherwise things are 'neutral' or 'bad'. This is not meant to suggest that choices are dominated by something like hedonism, but rather that people try to achieve things that they want (perhaps in our evolutionary past their own or others survival through the benefit of food or escape from danger) and that choices are made in a way that they believe is likely to realise them, in other words, choices are a means to an end, or instrumental. It should also be noted that, in the sense currently used, good and bad should not be taken in a moral or ethical way but rather to indicate that a person feels attraction and aversion or perhaps exhibits assimilation or contrast.

To be able to make accurate predictions and therefore effective choices, implies that certain characteristics are represented for every object, situation and action that we can make choices for, for example, to choose between "do I go for a walk or grab a coffee", I must be able to compare them and to do so requires a 'common currency'. While most of the aspects of things that we know about, for example, a particular concept, situation or object, are specific ("has wings" is an appropriate feature for describing birds, but not situations), arguably only one characteristic of a concept has to be represented in order to make a choice. The quantity used to compare possible options is known in economics as utility and in psychology as predicted reward.

This is taken as the point of departure for the present thesis. It is proposed that a modified form of the expected utility theory is both normative, suggesting how people ought to behave, and is also descriptive of how they do, in fact, behave. Maximising expected utility has long been accepted as a valid model of rational behaviour, however, it has limited predictive and descriptive accuracy simply because, in practice, people do not always behave in the prescribed way.

Accordingly, this is interpreted as evidence that either humans are not rational, expected utility is not an appropriate characterisation of rationality, or some combination of these. The present thesis argues that this observed behaviour can be considered rational and expected utility an appropriate characterisation of rational choice, provided it is accepted that *a*) most utility/reward has no meaning unless it is in the presence of potential competitors; *b*) there is uncertainty in the nature of the competitors; *c*) all statements of probability are also associated with uncertainty; *d*) utility is marginalised over uncertainty, with framing effects providing constraints; and that *e*) utility is also sensitive to risk, which when taken together with reward and uncertainty suggests a three dimensional representation of utility.

Chapter 2 motivates this view based on a survey of the literature that ranges from Damasio's Somatic Marker hypothesis (Damasio, 1994), which argues that rational choices are based on emotions and feelings, through to the expected utility hypothesis and how it has been shown not to be representitive of peoples behaviour, together with how it was modified by Rank Dependent Utility theories and, in particular, Kahneman and Tversky's (1979) Prospect Theory. It seems intuitively right, though, that rational behaviour is to do with choices that are made in a social environment (or social economy as it is sometimes called), which was noted by Neumann and Morgenstern (1944), and must take account of uncertainty and prior experiences, which give rise to beliefs. Our beliefs about things are important because, although they are subjective, we use a process of induction to apply them to novel situations. Indeed, the greater our experience of particular things and situations is, the more sure we are about making choices involving those things and situations.

While choices based on induction can not be easily justified on a rational (philosophical) basis, updating the beliefs that they are based on can be rational, using Bayes' rule. Taken together, the uncertainties and risks that must be inherent in the process of making a choice mean that, consistent with Rushworth and Behrens (2008), choices can not be made based on expected reward alone, but must instead involve two further dimensions, taking the risks and uncertainties of achieving a reward into consideration. It is highlighted that there is a psychological instrument called the semantic differential (Osgood & Suci, 1955; Osgood, Suci, & Tannenbaum, 1957) that has been in use since the 1950's and offers a three dimensional description of concepts in terms of their connotative or emotional content: In particular its reliability and universality are emphasised. The plausibility of thee dimensions of reward is also considered from the perspective of neurophysiology and the neurotransmitters, dopamine, serotonin, acetylcholine and norepinephrine.

Based on these ideas and adopting a very broad question at the outset considering what rules or processes govern people's choices in a social economy, Chapter 3 proposes that there are three dimensions of reward that can be plausibly described using the semantic differential. Implicit in this proposal is that these dimensions of reward must be maintained for all of the concepts that we know. The difficulties of collecting data that are concerned with good and bad experiences covering such a large range of concepts is addressed using internet blogs and an experiment carried out to investigate whether the inferred good and bad experiences are predictive of the dimensions of the semantic differential. It is found that the semantic differential dimensions are predictable and the Evaluation, Potency and Activity dimensions are tentatively relabelled as Reward, Risk, and Uncertainty.

However, the relationship between the Internet data and the semantic differential is correlational and, of course, a causal link can not be inferred from it; what is required is to be able to demonstrate manipulation of the semantic differential, that is, what has been proposed as a description of a three dimensional reward structure, through experiences alone. This is the subject of Chapter 4 which, on the basis of experiences from betting on arbitrary shapes (representing an economic decision), finds that the semantic differential can be manipulated and concludes that, to a first approximation, the semantic differential appears to be a summary of the reward history. Despite the analysis of internet blog data and the AlphaBet shapes experiment the Activity dimension, perhaps unsurprisingly, was not well addressed.

Based on the idea that this latent factor is perhaps more accurately described as control or certainty. Chapter 5, the triangles experiment, successfully focuses on certainty by keeping other variables, such as the shape, static. In addition, because, no explicit rewards are given to participants, unlike the AlphaBet shapes experiment, only the subjective feeling that a good choice has been made is available, this experiment is perhaps more like many everyday choices. A fourth experimental chapter investigates the hypothesised relationship between the semantic differential and reward structure based on the premise that the somatic marker hypothesis, and hence our reward structure, is created through physiological states and learned associations.

Chapter 6 concludes the first part of the thesis with an investigation of the idea that our reward structure will be evident from 'lower level' perceptual information. Since it is known that the mere perception of colour triggers evaluative processes (Elliot & Maier, 2007), perceptual information, in the form of colour, is assumed to be included when viewing a scene briefly, that is to say based on the gist of a scene, and used in its evaluation. It is found that a semantic differential can be produced from ratings of scenes that are viewed briefly and that this relates to the colour composition of those scenes. In addition, during the course of the data analysis, it is shown that a good representation of the basic colours, proposed by Berlin and Kay (1969), can be produced from first principles.

The second part of the thesis is concerned with constructing two functions with broadly the same characteristics as those described by prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), a value function and a probability weighting function. The chapters covering modelling these two functions are preceded by a short chapter that reiterates the importance of uncertainty, introduces Bayesian modelling and discusses a feature of the decision making literature that has received attention recently; decisions from description versus decisions from experience. The modelling then proceeds in Chapter 8 which addresses a function called the value function.

Based around Bayesian techniques, the approach used for the value (or utility) function introduces the novel ideas of utility of advantage and fair judges. Retaining the idea of a personal reference point introduced by Kahneman and Tversky (1979), it is proposed that utility, which can not be measured directly and since Kahneman and Tversky has been considered to be based on changes in wealth, is more correctly concerned with the probability (or how certain) that we are to gain an advantage over real or hypothetical competitors. Furthermore, harking back to the very origins of expected utility, that the winning or gaining advantage over those competitors is based on the idea that the result would be as if it were decided by a fair judge. The resulting value or utility function has all of the key features of the empirically derived value function of prospect theory. Another function, the probability weighting function, modifies the result of the utility function and is generally considered to represent the way that people have been found to overestimate small probability events, but underestimate medium and large probabilities.

Chapter 9 focuses on the probability weighting function, which, evident from the literature, has received much research effort. Since being proposed in prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), the weighting function has been considered to be as described above, underweighting medium and high probabilities and overweighting lower probabilities, however, this has been questioned recently under the heading of decisions from experience. The probability weighting function that is constructed in Chapter 9 tries to account for new and novel experiences using a novel, but straight forward Bayesian approach with a mixture of priors. As with the utility function, the probability function that is constructed has all of the features of the weighting function found in prospect theory. Armed with a model consisting of both of these functions, Chapter 10 carries out tests against the documented problem cases that are found in the decision making literature, based on the data and results from previous studies. The model is found to account well for all of these problem cases.

Chapter 11 draws the thesis to a conclusion, summarising the findings, identifying its implications and suggesting areas of future research.

Chapter 2

Literature Review

Two friends are camping and they are attacked by a bear.
One puts on his running shoes.
"What are you doing? You can't outrun the bear."
"I don't need to out run the bear. I only need to outrun you."
Anon.

2.1 Introduction

This thesis is, to a large extent, concerned with rational choices. "Rationality" in this context is not the same as its common or everyday use, meaning something like "in a clear and considered way", its philosophical uses, or just "sane". According to Herbert Simon (1979, p500), the classical model of rationality requires knowledge of all the relevant alternatives, their consequences and probabilities, and a predictable world without surprises. These conditions, however, are rarely met for the problems that individuals and organizations face, despite being implicitly (and sometimes explicitly) assumed in research. In this thesis, the meaning of "rational" is taken to be something more instrumental, such as, "weighing the presently perceived benefits in order to reach a choice that maximises some sort of personal advantage". This is not too far from the idea of rationality that is now used in some areas of Economics (Blume & Easley, 2008), but, contrary to many philosophers views, it tries to embody the personal and deeply subjective nature of many of our choices, yet retain the notion of reasoning.

The classical model assumes that decision making involves maximization of expected utility, almost as if there were unlimited knowledge, time, and processing power (Bechara & Damasio, 2005), and commonly that rational decisions must be strictly objective, logical and free from emotion (Goldie & Spicer, 2002). Emotions are always partial, arbitrary and just happen to people, whereas rational choices should be impartial, justified and made freely; indeed, a standard instruction given by judges, to members of a jury in a court of law, is that they should not be swayed by emotions, or let emotions influence their judgements (Pizarro, 2000). However, emotions have long been recognised as powerful influences on human behaviour and their function or purpose in our lives has been debated historically. Plato considered emotions and affective reactions in general, to be 'foolish counsellors' and two thousand years later philosophers such as Descartes and Kant continued to view emotions as afflictions that biased and obscured thought and decisions. Psychology and neuroscience, on the other hand, suggest that this view is unrealistic because everything we do is inextricably linked with our emotional and affective systems (Damasio, 1994, 1996, 2000). Unlike the classical model, this thesis considers that our choices are unavoidably coloured by affect and, indeed, that affect is a vital part of making effective choices.

The literature that is concerned with choices is very deep and richly textured, beginning arguably with J. Bernoulli (1713) who was interested in probability, due mainly to gambling, and continuing through to the axiomatisation of Expected Utility theory (Neumann & Morgenstern, 1944) and Prospect Theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Much of the literature, especially in recent years, seems to be characterised by research that concentrates on axiom violation. The intention here, however, is to focus on those parts of the literature that are concerned with affect, context and uncertainty, which are surely very important, especially when making intuitive, everyday, choices and form the foundation for the present thesis; indeed, Davidson and Irwin (1999) conclude that "every region in the brain that has been identified with some aspect of emotion has also been identified with aspects of cognition... The circuitry that supports affect and the circuitry that supports cognition are completely intertwined", a point that is central to Antonio Damasio's Somatic Marker hypothesis.

2.2 The Somatic Marker Hypothesis

The somatic marker hypothesis (Damasio, 1994) offers a neuroanatomical and cognitive framework for decision making and how emotion influences it; building on an idea that was proposed by William James and Carl Lange (James, 1894; Lange, 1885), the somatic marker hypothesis proposes that emotions arise from physiological states of the body, for example, the emotion of fear arises from the physiological state of increased heart rate, sweating, etc. that is associated with some event. Significantly and central to the hypothesis is that marker signals arising in these bioregulatory processes, particularly those that express themselves in emotions and feelings, influence decision making. This influence can be at different levels, some of which occur consciously and others that occur non-consciously

(Bechara, 2000).

The Somatic Marker hypothesis draws a distinction between primary and secondary emotions. Primary emotions, happiness, fear etc. appear to be hard coded and generated by evolutionarily older parts of the brain, particularly in the limbic system (Damasio, 1998). These areas of the brain are responsible for predictable body responses such as increased heart rate and sweating. Secondary emotions or the experience of emotional states (more correctly 'feelings' according to Damasio), on the other hand, are learned and depend on the experiences of an individual; they affect ongoing thinking and accordingly can alter future thinking. Secondary emotions are dependent on the prefrontal cortex, but operate through the older primary network (Damasio, 1994, pp173-183). Indeed, Damasio identifies the prefrontal cortices as one of the very few areas that receive signals about any activity that takes place in the mind and body (Damasio, 1994, p181).

As individuals go about their lives, frontal networks create associations between activity in primary sensory cortices and physiological states of the body; experiences are marked by body states, hence the Somatic Marker hypothesis and it is these that are then used in future thinking. It is uncontroversial that categorisation is one of the most basic and important things that we do in terms of cognition, indeed, there is much evidence, especially from memory research, that memories of experiences and concepts are stored with associated affect (e.g. Squire, 1992; LeDoux, 1993, 1996).

Although there is no agreement on how it might be achieved, despite much research, categorisation of a novel stimulus involves what has been experienced previously, that is, categorisation depends on generalising from particular learned instances to novel situations (Kruschke, 2005). Memory researchers believe that episodic memories (memories for specific experiences) are refined into semantic memories (memory of meaning, understanding, and other concept based knowledge unrelated to specific experiences) over time, along with the attendant emotion (e.g. Squire, 1992). In this process, much of the episodic information about a particular event is generalised and the context of the specific events is lost.

Given this understanding of the way that semantic memories are formed and considering the way that we learn, it can be inferred that the affective component is available from both episodic and semantic forms of memory. Introspectively this seems to be right; we can all remember how it felt to be in a particularly embarrassing situation or the sadness at the loss of someone very close, even if it is a less potent emotion; in fact Damasio et al. (2000) has used just this approach to confirm the hypothesis that there is a close relationship between emotion and homeostasis.

Research from outside of Damasio's lab investigating 'interoception' (our awareness our homestatic state or 'the material me' as Craig (2002) puts it) using imaging techniques (Craig, 2002, 2003; Critchley, Wiens, Rotshtein, Ohman, & Dolan, 2004), also provides strong support for the somatic marker hypothesis. The feelings associated with concepts and experiences through somatic markers may be consciously or unconsciously perceived and replayed, through Damasio's (1994) body-loop or as-if-body-loop, with the result that alternative actions with negative markers are rapidly rejected and those with positive markers receive more attention.

The idea that the evolutionary process has created neocortical systems of regulation on top of more ancient ones is important because, on this basis, judging and decision making serves the same purpose as the more ancient limbic system i.e. survival, reproduction etc.. In Damasio's view, making a good choice is choosing to act so as to be ultimately advantageous to the individual in terms of survival and the quality of that survival (Damasio, 1994, p169). Key here is that there must be correctly operating emotional circuitry and correctly operating cognitive circuitry (including memory) that interacts.

When operating correctly, this interaction serves, amongst other things, to draw attention to salient events (Damasio, 2000), which would seem to be a significant adaptive advantage if those salient events concern things such as danger or reward for example, as is often the case when making choices. People with damage to their prefrontal cortex seem to lose the ability to integrate affect with their choices (Damasio, 1994). They can still carry out reasoning tasks, particularly those that are norm or rule based, but interestingly, having been released from the influence of feelings, these people do not become hyper rational, Mr Spock[1] like, fiercely objective and logical. Rather, they lose the ability to know, quickly and intuitively, that questionable choices should not be made; these individuals do not have intuitive feelings of rightness or wrongness. As Damasio illustrates with the famous case of Phineas Gage, people with frontal cortex damage can not decide which choice to make, often ending up making poor choices or no choices at all (Damasio, 1994). Damasio (1994, pp193-194) also provides an anecdotal illustration of a patient with ventromedial prefrontal damage who, when asked to choose a next appointment date from two alternatives that were provided, took more than half an hour, going through many reasons against each of the dates, and still being unable to come to a conclusion.

[1]Mr Spock is a fictional character in the Star Trek television series. He is one of the three central characters; offering his colleagues a classically rational, emotionally detached and logical perspective.

Somatic markers then, are feelings that have been connected, through experience and learning, to episodes and concepts, which are then used to predict potential outcomes using what Damasio calls body-loops or as-if-body-loops to play out scenarios and experience feelings (Damasio, 1994; Bechara & Damasio, 2005). If a negative marker is associated with a potential outcome it will serve as a warning or danger signal and may be automatically discarded, whereas if a positive marker is associated with a potential outcome attention will be drawn to it.

Importantly, for decision making, the somatic marker hypothesis provides a mechanism that reduces the number of options to be considered for a non trivial decision from indefinitely many for any given alternative to very few, which can and frequently does, occur beyond conscious awareness. Arguably, relying on a reasoning process alone is at best prohibitively time consuming and at worst would not reach a conclusion (as Damasio attempts to show in his anecdote), it is also potentially error prone (Damasio, 1994, p172). This has been a common argument in consideration of the calculations potentially required for Utilitarianism and Consequestialism (e.g. Sinnott-Armstrong, 2011). Damasio (1994, p173) does, however, allow that somatic markers may not be sufficient for all normal decision making and leaves room for a process of reasoning that ultimately selects an alternative in some instances. However, the number of potential alternatives will be narrowed considerably and attention directed to alternatives with the potential for greater advantage.

Support for the somatic marker hypothesis also comes from a different source, moral psychology. Much recent research into moral psychology has resulted in a growing consensus that moral judgement is affectively based and several theories have been proposed to explain the part played by emotion (e.g. Greene, 2002, 2007; Haidt, 2001; Haidt & Bjorklund, 2007; Hauser, 2006; Nichols, 2004; Prinz, 2006). The move towards emotion and affectively laden intuition as the basis of moral choice gathered momentum with the publication of research by Greene, Sommerville, Nystrom, and Darley (2001) and Haidt (2001). Both essentially contend that moral judgements are based on intuition, central to which is affect, arguing that reasoning in moral judgements is for post hoc justification of the intuition that gave rise to the judgement. Characteristic of this is an inability to accurately describe how the moral judgement is arrived at, often resulting in an invented story that Haidt and Bjorklund (2007) refers to as 'moral confabulation'. If moral judgements were the result of a reasoning process then the moral principles that were used to reach the judgement would be evident in the justifications that were given for those judgements.

Greene, Nystrom, Engell, Darley, and Cohen (2004), like Damasio, also leaves room for a reasoned cost benefit analysis when there is no overriding emotion, for example in a moral dilemma. It is interesting to note though, that even after the reasoning process produces a choice, many people would still consider that the result did not feel right. In a moral dilemma[1], the least worst choice (whatever that might be) is still morally wrong, and in the case of ordinary choices we even have an idiom, "letting the heart rule the head" or vice versa, that acknowledges

[1]For example, The crying baby dilemma, Greene et al. (2004): Enemy soldiers have taken over your village. They have orders to kill all remaining civilians. You and some of your friends have sought refuge in the cellar of a large house. Outside, you hear the voices of soldiers who have come to search the house for valuables. Your baby begins to cry loudly. You cover his mouth to block the sound. If you remove your hand from his mouth, his crying will summon the attention of the soldiers who will kill you, your child, and the others hiding in the cellar. To save yourself and the others, you must smother your child to death. *Is it appropriate for you to smother your child in order to save yourself and the others?*

the difficulty.

Nonetheless and somewhat ironically, Greene (2007) argues that the post-hoc justifications and confabulations seen in research are deontological in character and can be plausibly considered to be the basis of philosophical deontology (that is, concerned with obligations, duties and right action). Greene justifies this on the basis that we have strong feelings about what can and cannot be done, but we have no clear idea how to make sense of those feelings. He goes on to suggest that philosophers, in order to explain these feelings, have made up a rational story about rights and duties, concluding that deontology is the cognitive expression of our deepest moral emotions.

Recent research using imaging techniques (Shenhav & Greene, 2010), published during the present project, pulls the findings on moral psychology and decision making together by reporting that both moral and ordinary decisions are based on the same neural mechanisms. Shenhav and Greene (2010) suggest that the prefrontal areas represent both the subjective value of material gains and losses, together with more abstract and hypothetical gains and losses that may have no material effect on the decision maker. It is interesting to note that many of the imaging studies investigating decision making in one form or another and from one standpoint or another, all find activation of the same areas of the brain. Journal article after journal article implicates areas of the prefrontal cortex, more specifically the orbitofrontal cortex and ventromedial prefrontal cortex, as well as the anterior cingulate cortex in decision making and emotion (e.g. in addition to those cited above O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Walton, Devlin, & Rushworth, 2004; Rolls, 2006; Naqvi, Shiv, & Bechara, 2006).

In summary, emotions or feelings are central to good decision making and many of the predictions of the somatic marker hypothesis have been found to be correct (e.g. Bechara & Damasio, 2005). The hypothesis also attracts support from other areas (e.g. Craig, 2002; Critchley et al., 2004; Shenhav & Greene, 2010); indeed, it is easy to see that there are also striking parallels with psychological research more generally, particularly in the areas of affective primacy, (Zajonc, 1980); cognitive impenetrability and confabulation, (Nisbett & Wilson, 1977); and dual process (or system) operation, where an intuitive or automatic affective process can be overridden by (or operate in conjunction with) a conscious deliberative process (Stanovich & West, 2001).

In much of the research already discussed it could be considered that choices are made on the basis of maximisation of simple reward (the maximum good or happiness for example); but this seems to be unsatisfactory as it ignores the risk and uncertainty that is, or might be, associated with obtaining the reward. Rushworth and Behrens (2008) also suggests that models of decision making can be improved by considering a richer conception of reward consisting of reward, risk (or cost) and uncertainty. This is an important part of the present thesis that can also be motivated by considering the decision making environment, in fact, this was considered by Neumann and Morgenstern (1944) under the heading of rational behaviour in *Theory of Games and Economic Behavior*; their classic treatment of expected utility.

2.3 Rational Behaviour

Far from anything to do with neurophysiology, psychology or reinforcement learning, the *Theory of Games and Economic Behavior* (Neumann & Morgenstern, 1944), provided a starting point for the interdisciplinary research field of game theory. Importantly for the present thesis, though, this book also provides a necessary and sufficient set of axioms from which the expected utility hypothesis can be derived (see Appendix A), though it should be noted, as Mongin (1997) does, that the axiomatisation that has become familiar had to wait until the work of economist Marschak (1950) and mathematicians Herstein and Milnor (1953). The expected utility hypothesis is originally attributed to D. Bernoulli (1738) based on his solution to a gambling problem known as the St. Petersburg paradox or prospect (see Appendix B), where rational behaviour can be described as maximizing the expectation of a utility function.

Utility is an abstract concept rather than an observable quantity, which need not be based on money and as such it can account for risk aversion. Because people clearly did not consider the St. Petersburg prospect in terms of expected monetary value, which in this case is infinite, D. Bernoulli argued in effect that people estimate value instead in terms of the utility of money. D. Bernoulli proposed the log function as a description of expected utility because of its property of being able to model decreasing marginal utilities (Mongin, 1997). Moving forward to the twentieth century, the expected utility hypothesis set out in *Theory of Games and Economic Behavior* (Neumann & Morgenstern, 1944) essentially argues that people are rational to the extent that they satisfy the axioms that they set out and accordingly that rationality can be modelled as maximising an expected utility or reward. This has been taken to be what people ought to do and remains the normative theory of choice in economics.

However, although it was hailed as a book of outstanding importance (Hurwicz, 1945), very soon after the Theory of Games and Economic Behavior was published, research became available that showed that human behaviour violated the normative principles that it set out (e.g. Mosteller & Nogee, 1951). As has been shown with many experiments that result in violation of the axioms (see below), the expected utility hypothesis has limited predictive and descriptive accuracy simply because, in practice, people do not always behave in the appropriate way, rationally trying to maximise wealth or some fixed non-linear function of it. Nevertheless there are considerations identified by Neumann and Morgenstern that are important to the present thesis.

Early in the Theory of Games and Economic Behavior, Neumann and Morgenstern (1944, §2.1, p8) discuss the problem of rational behaviour; beginning with the sorts of choices that might face the only member of a very simple economy, Robinson Crusoe. Whereas Robinson Crusoe's problem amounts to simple maximisation, in the sense that external conditions are given and he has to make choices in order that his position is as good as it can be, a member of a social economy, while having preferences, also has to enter into competitive relationships with others. Importantly, in a social economy, a choice will depend not simply on an individuals own actions but also on those of others, with each attempting to maximise a function where they do not control all the variables. In this situation people are forced to deal with risk and uncertainty. In much of the decision making literature it is often difficult to know whether a discussion is based on risk or uncertainty, but the difference seems to be quite intuitive and was captured by Knight (1921). According to Knight, risk is about an event or situation where the probability of an outcome can be determined, while uncertainty, in contrast, refers to an event where probability cannot be determined. Neumann and Morgenstern conclude that the problem in the social economy, unlike Robinson Crusoe's problem, is a "disconcerting mixture of several conflicting maximum problems"; accordingly, it can be seen that rational behaviour is considered to be a complex maximising process in a social economy and that, due to the complexities identified, the members of that economy must deal with uncertainty. Research with the Ulimatum game potentially supports this view (e.g. Hagen & Hammerstein, 2006; Henrich et al., 2005).

The Ultimatum game has variations that can be played by two or more participants. In its simplest form, a proposer offers a responder part of a fixed amount given by the experimenter. If the responder accepts the offer, he/she gets the amount, the proposer keeping the remainder, but if the offer is rejected neither proposer nor responder get anything. The expected utility hypothesis suggests that participants ought to be utility maximisers with the proposer offering a minimal amount and the responder accepting it. However, in anonymous one-off experiments, conducted in different cultures, proposers commonly offer considerably more than a minimal amount, and in some cultures responders reject relatively generous offers (Henrich et al., 2005).

However, despite steps taken to ensure that participants remain anonymous and will not meet again, players tend not to be given explicit details of the game context, or framing, beyond instructions on how the game works (Hagen & Hammerstein, 2006). In view of this, participants are likely to take cues, which can be subtle and occur automatically, from the environment in which they find themselves, constructing their own frame of reference (Haley & Fessler, 2005). For example, Bateson, Nettle, and Roberts (2006) have shown that in a naturalistic setting people will contribute significantly more to an honesty box for drinks when a picture of eyes is placed above the honesty box and drinks equipment, rather than a picture of flowers. It is thought that the eyes motivate cooperative behaviour, because they create a perception of being watched. The authors argue that the results confirm the hypothesis that concern for reputation is a powerful force motivates cooperation. A similar result was found by Haley and Fessler (2005) in a game played anonymously with either a bland background on the computer screen they used or one showing a drawing of eyes. These examples are compelling evidence that humans are attuned to cues that may affect the judgements they make and actions they perform.

This may also explain variations in the levels of generosity or cooperation that is observed when the Ultimatum game is played in different cultures, for example, a series of studies undertaken across fifteen small scale cultures based on the Ultimatum game (Henrich et al., 2004). In western cultures offers tend to be around 45%. In other cultures, people offer more, and in some they offer less. Among the Machiguenga of Peru, the average sum offered was 26% and the most frequent was 15%, an offer that most western players would consider unfair and likely reject. The New Guinean Au and Gnau consider accepting a gift creates a strong obligation to reciprocate, which the receiver often finds difficult. If the receiver does not reciprocate there are serious social consequences. This may be the reason that these people tend to reject offers exceeding 50% in the Ultimatum game (Tracer, 2003). These examples show that players draw upon information that is not apparent in the formal structure of the game or the experimenters instructions and is presumably based on prior experience of social factors.

Nonetheless and despite their own comments, Neumann and Morgenstern state that they will ignore social considerations on the basis that they do not consider that they change the formal process of maximising (1944, §2.2.1, p10), indeed, Marschak (1950, p113) also explicitly excludes basing rational behaviour on incomplete information, opting instead for a theory based on complete information. However, while social considerations may have little effect on the maximisation process itself, the same may not be true of processes surrounding value and independence. Uncertainty and incomplete information must necessarily be considered. While this point is revisited later in this thesis it is clear that many of the unknown and uncertain variables that people have to handle must be inferred from prior knowledge and experience. In order to do this, people must reason inductively, that is, to derive general principles from particular facts or instances, and while it is an obvious point for some, it is one that is important, forming part of the background to this thesis and motivating the use of Bayes' rule.

2.4 New knowledge

The problem of induction is a philosophical question about whether inferences based on a series of observations lead to new knowledge. It is a problem that has exercised philosophers since at least the time of Sextus Empiricus (c. 160-210 CE) and in its current form since Hume (Hume, 1739). Although the arguments surrounding the problem of induction are not covered in detail here, a description is provided of how people are thought to gain knowledge that is sufficient to make choices.

We can know some things with certainty even though they have not been, and in many cases never will be, observed, for example, a common knowledge statement that 'I have not observed every triangle that there could be, but I know in advance and with certainty that they will all have three sides' is, while elementary, a valid deductive argument. A valid deductive argument is one that is true by definition, where the conclusion follows from the premises and true premises guarantee a true conclusion; indeed, to try to deny that a triangle has three sides is clearly incoherent. However, while this type of argument may reorder or rearrange what is known, it does not add anything to it and it is not the sort of knowledge that is of interest here, rather, what is of interest are what Hume (1748, §IV, Part I) calls matters of fact.

Matters of fact claim to report the nature of things in the world and a proposition that is a matter of fact has the distinctive feature that both the fact and its contradiction are conceivable. There is no doubt that we have opinions and beliefs that involve unobserved matters of fact; indeed it is probably fair to say that practical life would be impossible without them; it has already been mentioned in §2.2 that inference is used in categorisation, albeit unconsciously. Knowledge of unobserved matters of fact can not be derived *a priori* (prior to experience). Rather, knowledge of matters of fact must in some way result *a posteriori* (from experience).

To paraphrase Hume (1748, §IV, Part I); someone would have to be very smart to discover, just by thinking, that ice is the effect of cold, without previously having experienced water turning to ice in the cold. Having experienced English winters, for example, we would obviously predict that water will turn to ice given sufficiently low temperatures, and clearly be right. However, the water could stay liquid or turn to ice; neither conjecture involves any kind of contradiction and as long as no investigation is allowed, neither proposition can be ruled out. So *a priori* there are no grounds for an opinion or belief one way or the other. The difference between us and our potentially smart friend, is simply a difference in experience; which seems to be completely general. While it may not be obvious how our experience is relevant to our prediction, the fact that it is relevant, is obvious.

How then is our prediction derived from our experience? Hume (1748, §VI & VII) asserts that all reasoning concerning matters of fact seems to be based on the idea of cause and effect since it is the only way that we can go beyond our experience. However, there is no necessary connection and any attempt to logically prove a connection results in an unwarranted assertion or assuming induction (the conclusion) in the argument, which is begging the question; the belief that two events are causally related must therefore be a habit that is acquired through experience because of the constant conjunction of those events. In other words, because every time a particular event is seen it is always preceded by the same thing it is inferred that the preceding event caused the succeeding event. This simple form of argument is known as enumerative induction and goes from a number of particular observations to a general principle (or law); historically induction was understood as this sort of enumerative induction, however, thinking about induction has become more sophisticated since Hume. A weaker form of enumerative induction, singular predictive inference, leads not to a generalisation but to a singular prediction:

1	a_1, a_2, ... , a_n are all Fs that are also G
2	a_{n+1} is also F
$\therefore$	a_{n+1} is also G

Singular predictive inference also has a more general probabilistic form:

1	The proportion p of observed F's have also been G's
2	a, not yet observed, is an F
$\therefore$	It is probable that a is a G

Hume (1748) asserts that all beliefs and opinions about unobserved matters of fact are derived from experience by induction or equivalently, our beliefs in matters of fact arise from a sentiment or feeling rather than from reason or the rational application of the formal rules of logic or probability (Morris, 2011); note here the similarity with the Somatic Marker hypothesis (Damasio, 1994). Interestingly, research has highlighted that people are generally poor at applying the formal rules of logic and probability theory to everyday problems and choices (e.g. Tversky & Kahneman, 1983, and for a review of research in this area ; Oaksford, Chater, and Stewart (2009)). The problem is though, that because induction is contingent it brings with it the risk of error and of not rationally knowing which belief to hold; if that is the case, then surely there is a serious problem with induction?

At this point it is perhaps sufficient to note that the descriptive problem of how we can know about unobserved matters of fact is resolved if they are derived from experience by induction, as Ramsey (1926) states, echoing Hume, "We are all convinced by inductive arguments and our conviction is reasonable because the world is so constituted that inductive arguments lead on the whole to true opinions. We are not, therefore, able to help trusting induction, nor, if we could help it, do we see any reason why we should". Nonetheless, there is not a sufficient justification of induction from a reasoned logical point of view: This is the normative problem of induction, but, further consideration of the problem is not part of the present thesis except to the extent that it might be possible to avoid the problem.

We can accept, as Ramsey (1926) does in the quotation above, that Hume is correct and that any opinion can be formed with varying degrees of belief from an inductive argument, however, whether those opinions are rational or not is not the question of interest. The question of interest is really whether beliefs and opinions can be modified rationally based on additional experience or evidence. Induction is the reasoning we do every day while operating in the real world, that is, what we do when making choices about the world. We can think of it as learning from experience and applying our prior experiences to new, but similar, situations. There is a uniquely reasonable way to learn from experience, using Bayes' rule (Hacking, 2001). Using Bayes' rule avoids the normative problem of induction by accepting the descriptive explanation, whether it is reasonable or not, and identifying a model of reasonable change in belief that is sufficient for being rational in a changing world.

Assuming the somatic marker hypothesis and maintenance of markers on (or at least that can be modelled on) a Bayesian basis, implies that there is a marker for every concept that we know, and consistent with Rushworth and Behrens (2008), it is reasonable to consider that it will be based on reward, cost (or risk) and uncertainty. It is worthy of note that there is a (social) psychological instrument, that has been around since the 1950's, has been extensively tested, and which purports to measure the connotative (or emotional) meaning of concepts, in three dimensions: The Semantic Differential.

2.5 The Semantic Differential

The literal or primary meaning of a concept is its denotation, in the case of a word, its dictionary definition; however, in addition to this meaning, there is another way to understand a concept, its connotation. Connotation is the idea or feelings that a concept invokes. Consider, as an example, that the dictionary definition of 'pub' is *(noun): A place of business where alcoholic beverages are sold and drunk [Short for public house]*, whereas, 'pub' evokes connotations such as merriment, pleasure, cheerfulness, perhaps some sadness, etc. Similarly, words such as summer, love, and melody tend to carry positive connotative associations for most people, while words like cancer, fight and homeless have negative connotations. The Semantic Differential (Osgood, 1952; Osgood et al., 1957) is thought to measure the connotative meaning of concepts; indeed, it is probably the most successful empirical method that has been devised for studying the nature of connotative meaning.

Establishing a semantic differential is procedurally straight forward: commonly, a large number of concepts (e.g. objects, actions and settings) are presented to participants who are asked to rate them on perhaps as many as 40-50 scales, although this can be varied with potentially as few as one concept being rated on just three scales. Each scale is typically a seven point Likert type scale based on contrasting adjectives (e.g. clean vs dirty; fast vs slow etc). Data from Semantic Differential scales are coded numerically on the basis that the scales are polar opposites and that they are equal interval scales passing through zero. While the use of Semantic Differential scales on this basis involves some assumptions that may not be perfectly accurate, violation of these assumptions is not considered serious enough to interfere with its application (Heise, 1969). Generalisations are supported by many research studies on the use of Semantic Differential rating scales for measuring affective meanings (Snider & Osgood, 1969), the scales are a straight forward and economical method for collecting data on reactions to stimuli, and they are easy to use for adults and for children from any culture (Osgood, May, & Miron, 1975; DiVesta, 1966). Once the scale data have been collected, factor analysis is then used to analyse the ratings, resulting in three robust observations.

First, approximately 50% of the variance in the rating data can be captured by just three dimensions (Evaluation, Potency and Activity). Second, the most important of these dimensions, known as Evaluation, almost always corresponds to whether the concept is 'good' or 'bad'. Third, the two other dimensions, Potency and Activity, each account for about the same amount of variance, with Potency capturing the extent the concept is 'strong' or 'weak', and Activity whether the concept is 'calm' or 'chaotic'. Cross cultural research has shown that these Evaluation, Potency and Activity dimensions are clearly common across cultures and languages (Osgood et al., 1975).

Although both this technique and these results are over fifty years old, it has stood the test of time and has been found to be robust across domains (Dalton, Maute, Oshida, Hikichi, & Izumi, 2008; Kim & Kang, 2009), languages and cultures (Osgood et al., 1975; Heise, 2001). A review of the literature that has been published during the last fifty years suggests that there are few psychological principles that have received such cross group and cross cultural verification, and there are few approaches that are associated with such applicability and breadth of findings as those that are found in Semantic Differential applications. Indeed, substantial numbers of papers are published each year using or referring to the Semantic Differential in an enormous range of journals (Google scholar in December 2011 returned $> 4,440$ hits since 2010 for the term "semantic differential").

The Semantic Differential is clearly significant as a multivariate approach to affective meaning, especially when compared with, for example, attitude measurement or expected utility which deal with the single dimension of Evaluation (Heise, 2001). However, Osgood et al. (1975) was hard pressed to explain the universal patterns of affective meaning, but points to the centrality of emotion in human affairs: We can imagine the situation for our ancestors when coming across a bear, for example. Three things had to be dealt with rapidly: *a*) Is it good or bad for me? (evaluation); *b*) Is it stronger or weaker than me? (potency); and *c*) Is it faster or slower than me? (activity). Regardless of their origin, these emotional reactions seem to be universally held and the semantic differential has shown itself to be a useful tool in their investigation. Theoretically though, the construct is less satisfactory and has been sharply criticised as a measure of meaning (Weinreich, 1958).

Besides the rather vague idea that the semantic differential measures connotative or affective meaning, perhaps better referred to as sentiment, it is still far from clear what is actually being measured and why such a robust, reproducible finding is found over so many domains and cultures (Miron, 1969; Osgood, 1969). It is proposed as part of the present thesis that theoretical insight can be gained from one defining feature of the semantic differential, its generality and that the semantic differential actually represents or describes a three dimensional reward structure.

Considering the semantic differential as a representation of our reward structure implies that we use it to make choices, as in the example above, to evaluate and then choose a course of action. Kahneman and Tversky's Prospect Theory (Kahneman & Tversky, 1979) provides a descriptive theory of how people make choices for certain types of problem.

2.6 Axiom Violation and Prospect Theory

The expected utility hypothesis tells us that given an option consisting of either a reward or a penalty that occurs with a known probability that the rational thing to do is to choose the prospect that maximises the expected utility associated with it (Davis, Hands, & Maki, 1998). Although the expected utility hypothesis is simple to understand and, on an abstract level, seems to be a good characterisation of what people should do, it provides a poor description of the choices that people actually make (e.g. Kahneman & Tversky, 1979). As already discussed in §2.3 the expected utility hypothesis is due to D. Bernoulli (1738) and was axiomatised two centuries later by Neumann and Morgenstern (1944), however, very soon afterwards, research began to be published that showed that human behaviour did not seem to conform with the hypothesis.

A great deal of research work has been carried out on axiom violation since the publication of The theory of Games and Economic Behavior, much of which has led to alternative theories, including prospect theory. The small industry of axiom violation experiments can be summed up by the following quotation: "Give me an axiom and I'll design the experiment that refutes it" (attributed to Amos Tversky in Gilboa, 2010). Details of much of this research work can be found in a meta analysis of axiom violation research by Yaqub, Saz, and Hussain

(2009, although it concentrates on the Behavioural Economics era from 1990).

Prospect theory greatly improved the match between theory and human performance by making two modifications to the expected utility hypothesis for which Kahneman received the Nobel Prize for Economic Sciences in 2002. The first modification is to replace the absolute utility function with a nonlinear value function that is expressed relative to a persons current position, convex (risk averse) for gains, concave (risk seeking) for losses and the curve for losses is steeper than that for gains; in other words, things that make your world better are good, and things that make your world worse are bad. The second modification is that when people are presented with probabilities, they behave as if they are not using the given probability, but a version of it transformed by a probability weighting function. This function has four robust characteristics; it increases the effective probability of improbable events; decreases the probability of probable events; crosses the line of equality at a probability less than 0.5; and is systematically different for the assessment of positive and negative prospects.

Kahneman (2011, Part IV, Chapter 25 & 26) cites inclusion of a reference point, utility (value) being based on changes in wealth, and losses being treated different from gains as the significant contributions that prospect theory made. In fact, Kahneman (2011) expresses considerable surprise that what, on reflection, appear to be rather obvious anomalies in expected utility theory, had not been spotted for the two hundred years of its existence, despite the powerful intellects of the people involved, putting it down to 'being in the grip' of a theory, essentially accepting that the theory is correct despite empirical data that could indicate otherwise.

To illustrate the importance of the changes to the utility function in how a choice might be perceived consider two people, $p1$ and $p2$, whose current wealth is 1 million and 4 million respectively. Both are offered the following gamble:

1. equal probability to end up owning 1 million or 4 million; or

2. own 2 million for certain

It is easy to imagine yourself in each position, making it clear that each person is likely to make a different choice; for $p1$, the certain option is very attractive, doubling wealth, while the other option contains no risk as wealth will either remain the same or quadruple. This, however, is very different for $p2$, choosing the certain option reduces wealth by half, while the best outcome from the risky option is to lose nothing but potentially lose three quarters of wealth. From the point of view of expected utility theory, though, both people face the same choice, expected wealth will be 2.5 million and 2 million respectively and the predicted choice would be the same for both people. However, this prediction is wrong because their current wealth or reference point is important and must be considered (a fuller description of these points can be found in Kahneman, 2011, Part IV, Chapter 25).

The principle of loss aversion, that is, that losses loom larger than gains when they are weighted against each other, is interesting as it adds to what is becoming a common thread in the research that has been reviewed; that emotions or affect is important to ordinary decision making. Loss aversion would seem to have evolutionary benefit, as the philosopher Quine (1970) put it "Creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing their kind" or in other words, since creatures that treat threats as more urgent than opportunities have a better chance to survive and reproduce; making a good decision is rewarding, but there may be severe punishments for getting it wrong. Identifying choices with reward in this way allows some speculation about the underlying neurophysiology that might support these decision making mechanisms.

2.7 Reinforcement and Neurophysiology

Clearly, the ability to make accurate predictions is of great importance in making decisions, whatever the choice or decision is; in providing information about possible outcomes the more accurate the prediction and the better the decision or choice. This is important for learning and is consistent with reinforcement learning models such as temporal difference learning (Sutton & Barto, 1998) and classical conditioning models such as the Rescorla-Wagner model (Rescorla & Wagner, 1972), where, in order to learn effectively, a prediction error is one of the key variables. The finding that the same brain areas always seem to be implicated in imaging studies of decision making has been highlighted in §2.2 and are often associated with the neurotransmitter, dopamine.

2.7.1 Dompamine

Reinforcement learning based models, particularly temporal difference models (Sutton & Barto, 1998), have had great success in developing an understanding of how we learn to make choices in an uncertain world. The standard model uses a one dimensional representation of reward, associated with the action of dopamine (Schultz, Dayan, & Montague, 1997; Schultz, 1998, 2010). The expected values for rewards that originate from prior experience and used for predicition, are thought to be represented in the prefrontal cortex, with strong evidence suggesting more specifically the orbitofrontal cortex (Kringelbach, 2005, 2005). Evidence for how these priors are maintained is provided from neurophysiology where recent research has shown that dopamine neurons, which project along the mesolimbic and mesocortical pathways, code the subjective value of rewards (Schultz, 2010). Both of these pathways originate in the ventral tegmental area, which is located in the midbrain, projecting to the nucleus accumbens, in the case of the mesolimbic pathway, and frontal cortex, in the case of the mesocortical pathway.

Dopamine neurons respond to rewards to the extent that a reward differs from what was expected. If a reward is better than expected it produces activation of dopamine neurons, but if the reward is worse than expected it produces a depression. Accurately predicted rewards produce no response and is perhaps the reason why rewards that do not change lose their influence and why there is a drive for greater reward (Schultz, 2010). In other words, dopamine neurons provide a reward prediction error, which maintains the prior values represented in the prefrontal cortex (Tobler, Fiorillo, & Schultz, 2005). There is much evidence to support the part played by dopamine in the reward system, however, a full range from reward to punishment must be managed, which is thought to involve another neurotransmitter, Serotonin.

2.7.2 Serotonin

Serotonin pathways originate in the raphe nuclei of the brain stem and project both to the spinal cord and very widely to cortical and sub cortical areas; and results of research suggest that the general function of serotonin in motivation is to encode aversive outcomes. For example, serotonin releasing neurons are activated by aversive events such as inescapable shocks (Takase et al., 2004), and animals with low serotonin levels show less behavioural suppression to cues and contexts predictive of punishment (Soubrié, 1986). Very often in the brain, continua, such as a range of rewards and punishments, are managed by pairs of systems through opponency: Somewhat controversially, recent research has also proposed that prediction error, across a full range of rewards and punishments, is managed through opponency between the dopamine and serotonin systems (Huys & Dayan, 2009; Crockett, Clark, & Robbins, 2009; Boureau & Dayan, 2010; Cools & Nakamura, 2010).

However, while serotonin has for a long time been implicated commonly with affective processing of punishments and threats, these sorts of events have been found to covary both positively and negatively with serotonin levels (Boureau & Dayan, 2010), making the relationship between the serotonin system and the dopamine system potentially complex.

Dayan and colleagues analysis of extant literature (Boureau & Dayan, 2010; Huys & Dayan, 2009) suggests that the contrary findings for serotonin levels relates to differences between threats and punishments attributable to classical type conditioning and those attributable to instrumental type learning. The negative covariance between serotonin and threat being the result of the refusal to engage with potential or actual threats and punishments in the case of classical type conditioning. Behavioural evidence that supports this hypothesis has been provided by Crockett et al. (2009) using tryptophan depletion to demonstrate a difference between sensitivity to aversive outcomes (attributable to classical conditioning) and punishment induced inhibition (attributable to instrumental learning). In support of Dayan's hypothesis (Boureau & Dayan, 2010), the authors claim that they have shown that serotonin is critical for linking behavioural inhibition with predictions of aversive outcomes rather than performing inhibitory or aversive processing alone (Crockett et al., 2009).

In order to make good inferences and predictions about the world requires that, as well as being able to predict on the basis of reward and punishment, we have some idea about how sure the inference or prediction is; presumably, doing so is an important function that our brains must contend with.

2.7.3 Acetylcholine and Norepinephrine

Dayan and Yu (e.g Dayan & Yu, 2002, 2006; Yu & Dayan, 2002, 2003; Yu, 2003; Yu & Dayan, 2005) propose that what they refer to as expected uncertainty is associated with acetylcholine (ACh) and unexpected uncertainty with norepinephrine (NE, alternatively noradrenaline, NA). An example of these different sorts of uncertainty might be the choice of wearing a raincoat when leaving for work: this straightforward choice involves considering potentially conflicting information such as the forecast on the radio and the large black cloud overhead. For an individual who notes the weather forecast, the chance of incorrect forecast is a type of expected uncertainty, whereas a significant change in the reliability of forecasts would be unexpected uncertainty, presumably resulting in the individual to note other weather information (Yu & Dayan, 2005).

Research suggests that higher acetylcholine and noradrenaline leads to the reduced influence of prior, experience based information in the integration of experience based and sensory based information. Acetylcholine and noradrenaline also play a role in experience based placiticity in the cortex (Gu, 2002), which provides for maintenance of representations based on new experience. Acetylcholine and noradrenaline depletion has been shown to suppress experience based plasticity (e.g. Baskerville, Schweitzer, & Herron, 1997) and increases of Acetylcholine and noradrenaline shown to raise cortical reorganization when paired with sensory stimulation (e.g. Kilgard & Merzenich, 1998).

The evidence from reinforcement learning and neurophysiology supports the idea that somatic markers represent our reward structure, are three dimensional and are maintained in the prefrontal cortex, further, they are maintained through the action of neurotransmitters, dopamine for reward and serotonin for cost with acetylcholine and noradrenaline involved with uncertainty.

2.8 Summary

In a natural environment the only sources of information that can be used to inductively reason are an individuals own observations and experiences, supplemented by those signalled by others. This would seem to be as true today as it was for our ancestors, although in the modern world we are presumably faced with a greater quantity of information, presented in a greater variety of formats, from wider social groups and media. It is uncontroversial that some form of preferences must have existed prior to the evolution of intelligence or rationality and that an evolutionary process cannot produce the cognitive machinery to input, reason with or generally use, information in a format that was not available from the environment (Cosmides & Tooby, 1996; Robson, 2001). It is reasonable to suppose that one of the simplest forms of information available is and would have been in our evolutionary past, the frequency of good and bad things that occur and that this information is readily available through our emotional systems. Simply put, things that make your world better are good and things that make your world worse are bad.

While other theories propose that affect (feelings, emotion) provides information that can then be used when making choices (e.g. Schwarz & Clore, 2007; Schwarz, 2010) and are complementary, the somatic marker hypothesis provides a framework where choices are intimately tied with affect (Damasio, 1994, 1998, 2000). The somatic marker hypothesis is consistent with, and supported by, recent research in the areas of interoception (e.g. Craig, 2002) and moral psychology (e.g. Greene, 2007) in particular, and also from psychology more generally (e.g. Duckworth, Bargh, Garcia, & Chaiken, 2002; Zajonc, 1980). When the areas of the brain that are associated with affect, such as the pre-frontal cortices, are damaged, people do not turn into hyper rational, decisive individuals, but rather are unable to reach quick and effective choices. This suggests that feelings, far from obscuring good decision making, are an important and integral part of the process.

When faced with an uncertain or novel situation, the somatic markers are used, often beyond conciousness, to reduce the potential options to perhaps an obvious choice that can be used automatically, or to a manageable few, that can then be considered on the basis of a cost benefit analysis, with one that is appropriate, chosen. These processes, one which is fast and intuitive and the other that is slow as calculating, are what Stanovich and West (2001) and Kahneman (2011) refer to as system1 and system2 respectively.

Information for an uncertain or novel situation must be inferred or arrived at through a process of induction, reaching a general principle from particular instances; whether this process can be rationally (philosophically) justified in itself seems to be rather less important than whether somatic markers can be rationally maintained. It is argued that they can be maintained and revised in the light of experience which, in turn, can be well described using Bayes' rule. Taking this view and to base decision making on the somatic marker hypothesis implies that certain characteristics are represented for every object, situation and action we can make choices for. Somatic markers also imply that a one dimensional reward structure, as used in temporal difference learning for example (Sutton & Barto, 1998), is unsatisfactory because, consistent with Rushworth and Behrens (2008), it ignores the risk or cost and uncertainty that is associated with the reward.

In order to flourish in a particular environment then, our experiences can (must) be used as part of the process of making advantageous choices, where advantageous choices are generally rewarding. Indeed, it would seem strange if our prior experiences were not used; all the more so in a social economy where there are competitors and complicated situations that have to be navigated. As highlighted by Neumann and Morgenstern, rationally gaining an advantage over competitors, for whatever reason, in a social economy turns what should be a simple maximisation problem into a "disconcerting mixture of several conflicting maximum problems".

Based on the ideas above, the broad question for the present thesis is: what rules or processes govern people's choices in a social economy? In the next chapter it is proposed that relying on a single dimension, consisting simply of reward to base choices on, is inadequate; this is because, as well as the expected reward, it is also important to know what the potential cost (risk/danger) might be in obtaining that reward and how sure estimates or predictions are.

Chapter 3

Reward is assessed in three dimensions

3.1 Introduction

While most aspects of meaning are specific to a particular concept (for example, "has wings" is an appropriate feature for describing birds, but not for describing situations), possibly only one characteristic has to be represented for all situations, actions and objects. If we are to make choices in everyday life and compare different options, such as shall I go for a walk, or grab a coffee; or do I look at my book rather than the person that just walked into the room, a 'common currency' is needed. This quantity, used to compare possible alternatives, is generally known in economics as utility and in psychology as predicted reward or payoff. Utility or reward is usually conceptualised as a single dimension, however, it seems obvious that to make a rational choice, it is also important to know what the potential cost (risk/danger) might be in taking a particular choice is and how sure estimates or predictions are (Rushworth & Behrens, 2008). Here, the possibility that people actually have a richer, multi dimensional, representation of utility or reward is explored.

Chapter 2 (§2.5) identified the semantic differential as a very robust, widely used and widely applicable instrument, however, besides the rather vague idea that the semantic differential measures the connotative or affective meaning of a concept, it is still far from clear what is actually being measured and why such a robust, reproducible finding is found in so many domains and cultures. It is proposed here that theoretical insight can be gained from one defining feature of the semantic differential, its generality.

Understanding the benefit of more than one dimension on which to base choices is straight forward and can be illustrated using the example of the n-armed bandit (Sutton & Barto, 1998, §2.1). The n-armed bandit problem is to maximise the expected reward when faced with multiple, n, one armed bandits[1], without knowing the distribution of rewards that is applicable to each of them. If the distribution of rewards were known it would be easy to maximise them, always choosing the one most likely to pay out. However, the distribution of rewards for each of the bandits (or more generally concepts in the present experiment) has to be learned. With a single dimension of reward and adopting what Sutton and Barto (1998) call a greedy strategy, the rewards that are known can be exploited, but no exploration (sampling) of alternative, albeit (apparently) inferior, choices is made. While this might maximise immediate reward, it is unlikely to maximise reward in the longer term or if the distribution of reward changes.

[1] A one armed bandit is a gambling machine, otherwise known as a slot or fruit machine, so called because they were operated by a lever on the side (the arm) and the potential of leaving the player with no money (bandit).

Maintaining further quantities that represent risk and uncertainty means that the currently known maximum reward can be chosen most of the time and exploring actions can be made periodically, governed by risk and uncertainty, allowing the known rewards, risks and uncertainties to be maintained. In this way, over time, the reward structure is learned through experience of good and bad choices, which will support the maximisation of reward in the longer term. Clearly, if the distributions of reward are non-stationary, that is, the reward distributions change over time as they do in the real world, the need for exploration in order to maintain potential reward profiles becomes increasingly important.

The benefit of having more than one dimension on which to base choices therefore seems obvious, particularly in the non-stationary real world; for example, it allows assessment of the possible reward (or punishment) that might be attained, and of the potential cost of being involved in an activity, to be made independently of one another. This might have the result, say, that involvement in some activity or making a particular choice could be vetoed if there were too great a cost or risk. In addition a further uncertainty dimension will provide a basis on which to explore and thereby gain better knowledge of potential rewards and risks.

It has already been discussed that information from the environment must be used to establish preferences (Cosmides & Tooby, 1996; Robson, 2001) and that it is reasonable to suppose that one of the forms that this information takes is, and would have been in our evolutionary past, the frequency of good and bad experiences that an individual has. It is proposed that, rather than the one dimensional representation of utility or reward that is implicitly assumed in most theories, at least a two and probably a three dimensional representation of reward is built up through experiences of good and bad things happening and that this is what is being used to make choices. It is also proposed that these dimensions are what is represented by the semantic differential.

It is hypothesised that for some measure of good and bad experience there will be a relationship with the scores for the factors in a typical semantic differential (see Chapter 2, §2.5). To test this hypothesis two things are needed; the measured location in the semantic differential space of a large number of concepts, together with a means to estimate the probability of the reward associated with them. Obtaining the semantic differential scores is straight forward as there are readily accessible and publicly available semantic differential dictionaries (e.g. Francis & Heise, 2006). However, more difficult to estimate are the rewards and punishments associated with these concepts in everyday life.

At first sight, the problem of estimating the distribution of the probability of good or bad things happening across a wide range of contexts and concepts seems impossible. Ideally, but somewhat impractically, someone would be observed throughout their lifetime and for each of a large range of contexts and concepts, the number of times good and bad things that happened, together with the number of times that something good or bad could have happened, would be recorded. Fortunately, a more practical solution to this problem is provided by the recent phenomenon of the internet weblog or blog. Blogs are short descriptions of peoples' life experiences (good, bad and indifferent). They are also searchable.

3.2 Method

3.2.1 Materials

The semantic differential dictionary used here consists of 1500 concepts grouped under the broad headings of behaviour (naming actions that one person can perform on another person), identity (naming different kinds of individual), setting (naming places or times where social interactions might take place) and modifier (naming emotions, traits, and statuses that might characterise people), offering a broad selection for analysis. This dictionary, which was compiled during 2002/3 at Indiana University (Francis & Heise, 2006) was chosen because of *a*) accessibility; it was straight forward to download the dictionary required *b*) scope; it was the largest single dictionary that could be found *c*) age; it was the latest large dictionary that could be found and *d*) pedigree; the principal researcher, David Heise, has extensive experience with the semantic differential technique in general and has also published extensively on the subject since the mid 1960's.

To use the semantic differential data to establish counts of good and bad experiences required a mechanism for searching blog space as widely as possible. Initially use of the Google blog search engine was planned because a check of the online documentation, while non specific, suggested extensive coverage of blogs in much the same way as the Google web search engine covers web pages. It was quickly found, though, that attempting to run automated search scripts using the Google blog search engine was detected and blocked. Alternative search engines were found (Technorati and Blogscope) that provided extensive coverage of blogs, provided structured searching facilities and did not limit the use of automated search scripts. Technorati (technorati.com) and BlogScope (blogscope.com) are both free at the point of use and provide extensive coverage[1]. Blogscope is being developed as part of a research project at the University of Toronto, while Technorati is a commercial search engine funded by advertising. Since it was used for the data gathering described here, Technorati has undergone extensive restructuring.

MATLAB (2008) scripts were written to use the semantic differential data to build queries and submit them to both of the blog search engines; Using this method it is straightforward to find how many blogs, out of the millions indexed, contain a given concept (say, knife).

3.2.2 Procedure

To approximately classify each blog post containing a concept as being associated with positive, negative or neutral situations, a count was made of the number of posts that contained the concept, N_t; the number of posts containing the concept in combination with any of ten unambiguously good words, N_g, and the number in combination with ten unambiguously bad words, N_b. The good and bad word lists, shown in Table 3.1, were the highest and lowest evaluated in the modifiers subgroup of the semantic differential dictionary that was used (Francis & Heise, 2006). The distribution of the frequency of occurrence in spoken and written English of the good and bad modifiers (taken from the British National corpus using their simple search (www.natcorp.ox.ac.uk)) was not significantly different ($t(10) = 0.55, p = .60$). The presence of at least one of these modifiers was taken to indicate generally positive or negative situations or experiences (but see

[1]In June 2008 Technorati claimed to index 112.8 million blogs, however, at the time of writing no up to date figure was available. At the time of writing Blogscope claimed to be monitoring over 52.50 million blogs with 1.3 billion posts.

discussion).

Table 3.1: Good and bad modifier words used for the internet blog search.

Good	Bad
good	bad
amused	suicidal
polite	evil
relaxed	abusive
pleased	cruel
helpful	depressed
delighted	miserable
friendly	rude
generous	hurt
honest	mean
happy	unhappy

In order to carry out the blog searches, three statements were constructed for each concept as follows:

The concept on its own (N_t)e.g.

bully

The concept and a disjunctive list of good words (N_g) e.g.

bully and (good or amused or polite or relaxed or pleased or helpful or delighted or friendly or generous or honest or happy)

The concept and a disjunctive list of bad words (N_b)e.g.

bully and (bad or suicidal or evil or abusive or cruel or depressed or miserable or rude or hurt or mean or unhappy)

in this way, a blog post containing a concept with a modifier (i.e. good or bad word) appearing more than once was only counted once and the presence of one or more modifiers was taken to indicate that the post described generally positive or negative situations. From an original 1500 items in the semantic differential dictionary that was used (Francis & Heise, 2006), two data sets were compiled each consisting of single word concepts that occurred at least once using each search engine. Technorati provided a data set consisting of 972 concepts and Blogscope provided a data set consisting of 1071 concepts (the concepts found for Technorati and Blogscope can be found in Appendix C and Appendix D respectively).

3.3 Results

In order to investigate the relationship between the reward associated with a concept and its position in the semantic differential space, first six derived measures of the positive and negative rewards associated with a concept were constructed. The first two measures simply measure the proportion of times the concept was associated with good and bad contexts, $\frac{N_g}{N_t}$, and $\frac{N_b}{N_t}$ (absolute positive and negative reward). The third measure quantifies the proportion of rewarded situations where this reward was positive, $\frac{N_g}{(N_g+N_b}$ (relative reward). The fourth measure, included since it is known that preference (Evaluation) can be caused simply by exposure (Zajonc, 1980), was a measure of frequency, $log(N_t)$. Lastly to investigate any interactions between the reward measures, the interaction between 'negative and relative reward', $\frac{N_b}{N_t} \times \frac{N_g}{N_g+N_b}$, and 'positive and relative reward', $\frac{N_g}{N_t} \times \frac{N_g}{N_g+N_b}$, was also included. Using these as the independent measures stepwise regressions were performed against the three dimensions of the semantic differential (Evaluation, Potency and then Activity).

The full results for the analyses are given in Tables 3.2 and 3.3 where the beta coefficient, b, standard error of the beta value, SEb, standardised coefficient, β, result of the t-test for the beta, t, multiple correlation coefficient, R, variance explained, R^2, and the variance explained adjusted for the number of terms in the model, aR^2, are given. It should be noted though that with the blogscope data, using Evaluation as the dependent variable, there was a third step that entered $\frac{N_b}{N_t}$ into the model, however, this was disregarded as the R^2 change from the previous step was only .002 and considered too small to include for such a noisy data set. The analysis shows a similar pattern of results for both of the data sets, with the same independent variables being included for each dependent variable, with a similar amount of variance explained; a number of robust effects were found:

1. Evaluation was strongly related to the relative reward (technorati $R^2 = .35$, $F(2, 969) = 256.07$, $p < .001$; blogscope $R^2 = .35$, $F(2, 1068) = 284.36$, $p < .001$); but perhaps surprisingly, not to absolute reward (technorati $R^2 = .01$, $F(1, 970) = 5.72$, $p = .017$; blogscope $R^2 = .01$, $F(1, 1069) = 9.923$, $p = .002$), Figure 3.1.

2. For Potency the strongest relationship is with the absolute negative reward (technorati $R^2 = .20, F(2, 969) = 108.02, p < .001$; blogscope $R^2 = .23, F(2, 1068) = 138.97, p < .001$). Figure 3.2.

3. The best predictor for Activity was the 'negative x relative reward' interaction term, $\frac{N_b}{N_t} \times \frac{N_g}{N_g+N_b}$. Though this was highly significant, the level of correlation was small (technorati $R^2 = .09$, $F(2, 969) = 44.80$, $p < .001$; blogscope $R^2 = .08$, $F(2, 1068) = 48.10$, $p < .001$), Figure 3.3.

Table 3.2: Technorati stepwise multiple regression for each factor using absolute positive and negative reward, $\frac{N_g}{N_t}$ and $\frac{N_b}{N_t}$; relative reward $\frac{N_g}{N_g+N_b}$; exposure $log(N_t)$; and interactions 'negative x relative reward' $\frac{N_g}{N_t} \times \frac{N_g}{N_g+N_b}$ and 'positive x relative reward' $\frac{N_b}{N_t} \times \frac{N_g}{N_g+N_b}$ as independent variables

DV	Step	IV	b	SEb	β	t	R	R^2	aR^2
E	1	(Constant)	-14.21	0.66					
		Relative reward	23.78	1.09	0.57	21.86	0.57	0.33	0.33
	2	(Constant)	-14.24	0.65					
		Relative reward	22.03	1.14	0.53	19.40			
		Exposure	0.12	0.02	0.13	4.83	0.59	0.35	0.34
P	1	(Constant)	2.08	0.13					
		Negative reward	-5.06	0.40	-0.37	-12.55	0.37	0.14	0.14
	2	(Constant)	0.45	0.26					
		Negative reward	-3.70	0.44	-0.27	-8.47			
		Exposure	0.13	0.02	0.23	7.10	0.43	0.18	0.18
	3	(Constant)	-0.10	0.30					
		Negative reward	-4.83	0.49	-0.36	-9.89			
		Exposure	0.13	0.02	0.23	7.08			
		Positive x Relative reward	3.32	0.67	0.16	4.96	0.45	0.20	0.20
A	1	(Constant)	1.67	0.14					
		Negative x Relative reward	-6.76	0.78	-0.27	-8.72	0.27	0.07	0.07
	2	(Constant)	0.85	0.27					
		Negative x Relative reward	-5.51	0.85	-0.22	-6.51			
		Exposure	0.06	0.02	0.12	3.55	0.29	0.09	0.08

All $p <= .001$, $n = 972$

4. In each case the effect of the best predictor seems to be mediated by the amount of exposure that there has been to the concept, making good things slightly better and bad things slightly worse.

Given the high level of noise associated with our method of evaluating reward and the inherent noisiness of the semantic differential, the correlations could, perhaps, be considered surprisingly strong. In order to gain further insight into the relationship between the semantic differential factors and the measures of reward, the data were averaged across a range of factor values, in other words the data were binned, and further regressions carried out. The advantage of binning

Table 3.3: Blogscope stepwise multiple regression for each factor using absolute positive and negative reward, $\frac{N_g}{N_t}$ and $\frac{N_b}{N_t}$; relative reward $\frac{N_g}{N_g+N_b}$; exposure $log(N_t)$; and interactions 'negative x relative reward' $\frac{N_g}{N_t} \times \frac{N_g}{N_g+N_b}$ and 'positive x relative reward' $\frac{N_b}{N_t} \times \frac{N_g}{N_g+N_b}$ as independent variables

DV	Step	IV	b	SEb	β	t	R	R^2	aR^2
E	1	(Constant)	-9.06	0.42					
		Relative reward	15.88	0.72	0.56	22.05	0.56	0.31	0.31
	2	(Constant)	-10.11	0.43					
		Relative reward	14.47	0.73	0.51	19.93			
		Exposure	0.16	0.02	0.19	7.54	0.59	0.35	0.35
P	1	(Constant)	1.81	0.09					
		Negative reward	-3.62	0.25	-0.40	-14.36	0.4	0.16	0.16
	2	(Constant)	0.17	0.23					
		Negative reward	-2.90	0.26	-0.32	-11.04			
		Exposure	0.12	0.02	0.23	7.77	0.45	0.21	0.21
	3	(Constant)	-0.37	0.25					
		Negative reward	-3.40	0.28	-0.38	-12.35			
		Exposure	0.12	0.02	0.23	8.03			
		Positive x Relative reward	2.57	0.48	0.16	5.38	0.48	0.23	0.23
A	1	(Constant)	1.39	0.11					
		Negative x Relative reward	-4.81	0.55	-0.26	-8.81	0.26	0.07	0.07
	2	(Constant)	0.50	0.24					
		Negative x Relative reward	-40	0.58	-0.22	-6.92			
		Exposure	0.06	0.02	0.13	4.06	0.29	0.08	0.08

All $p <= .001$, $n = 1071$

is that there is a reduction in noise, which may help in gaining insight into average behaviour. In 'real' data there is always some amount of noise associated with a measurement or signal, which may be random or systematic. In the present case noise is assumed random and unavoidable. By binning the data i.e. placing multiple measurements into a bin, the ratio of signal to noise can be made larger. For example if four measurements are placed in a bin, each with an amount of signal and noise, the ratio will be $\frac{4 \times Signal}{\sqrt{4 \times Noise}}$; because it is random, noise adds as the square root, making the signal to noise ratio bigger.

One obvious drawback to this approach is the reduction in the number data

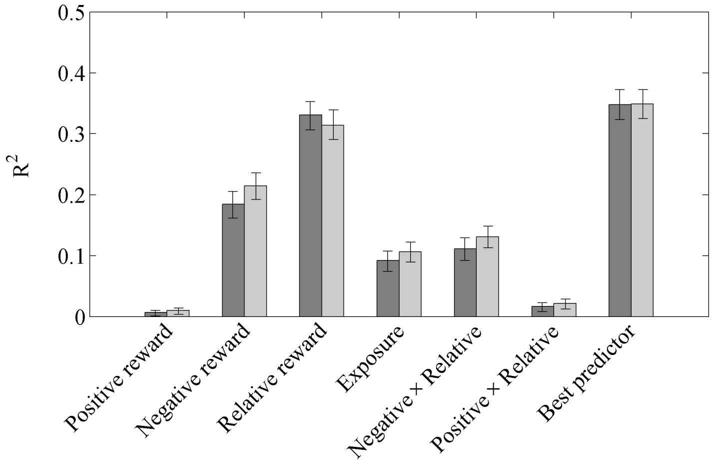


Figure 3.1: Correlations (R^2) between evaluation and each of the derived measures of reward, which clearly shows that relative reward is the best single predictor of evaluation. Dark bars represent the technorati search engine, and light bars blogscope. The bars labelled "best predictor" represent the adjusted R^2 for the best fitting regression model.

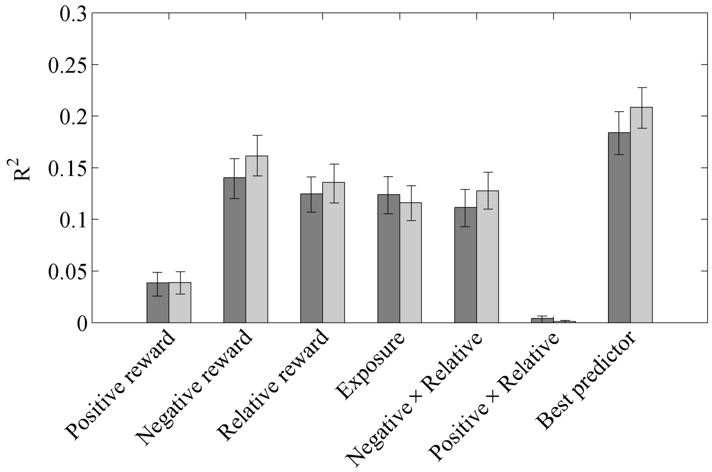


Figure 3.2: Correlations (R^2) between potency and each of the derived measures of reward showing that the best predictor of Potency is the probability of bad events (risk), though a number of other predictors are of reasonable size on their own. Dark bars represent the technorati search engine, and light bars blogscope. The bars labelled "best predictor" represent the adjusted R^2 for the best fitting regression model.

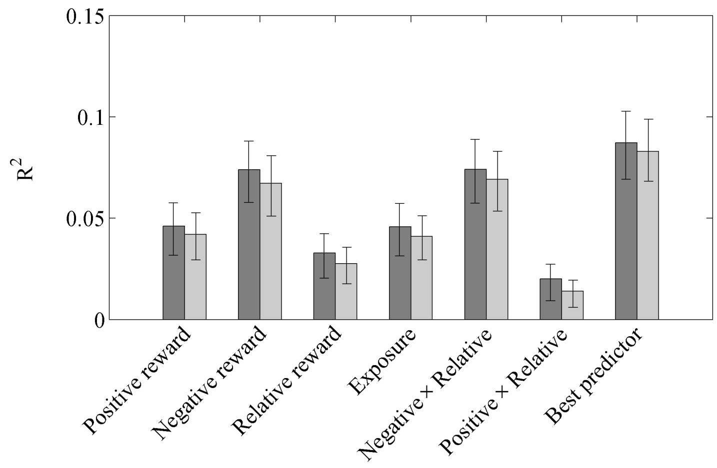


Figure 3.3: Correlations (R^2) between Activity and each of the derived measures of reward. The main observation for this dimension is that though many of the measures of reward are significantly correlated with activity, the absolute level of correlation is small. Dark bars represent the technorati search engine, and light bars blogscope. The bars labelled "best predictor" represent the adjusted R^2 for the best fitting regression model.

points that are available for further analysis. However, in the present case, due to the initial number of data points this was not considered to be a particular problem. The width of each bin was 0.2 for each of the semantic differential factors and regression analysis on the 'new' binned data points revealed:

1. for Evaluation predicting relative reward, $\frac{N_g}{N_g+N_b}$, $b = 0.01$, $t(47) = 14.86$, $p < .001$ $R^2 = 0.82$, $adjusted\ R^2 = 0.82$, $F(1, 47) = 220.73$, $p < .001$;

2. for Potency predicting the probability of bad events (risk), $\frac{N_b}{N_t}$, $b = -0.02$ $t(36) = -6.80$ $p < .001$ $R^2 = 0.56$, $adjusted\ R^2 = 0.55$, $F(1, 36) = 46.23$, $p < .001$;

3. and for Activity predicting 'negative x relative reward' interaction term, $\frac{N_b}{N_t} \times \frac{N_g}{N_g+N_b}$, $b = -0.01$, $t(35) = -7.99$, $p < .001$ $R^2 = 0.65$, $adjusted\ R^2 =$

0.64, $F(1, 35) = 63.77$, $p < .001$.

Figure 3.4 shows the form of the relationships with each of the semantic differential factors, and as can be seen, the relationships are all very close to linear.

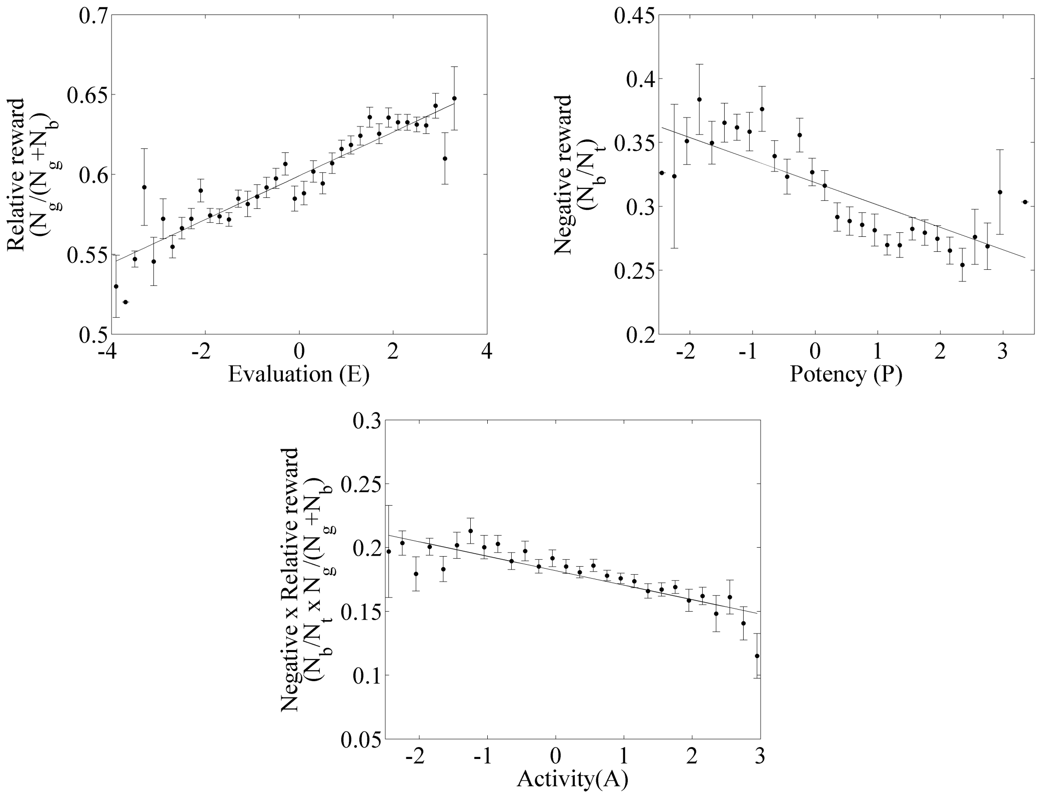


Figure 3.4: The relationships between the semantic differential factors and their respective experienced rewards with mean data calculated in bins of width 0.2. Each of the resulting data points is shown with error bars representing the standard error of the mean of the original values that were used to calculate it.

Using the predictors identified in the stepwise multiple regressions for each data set, further multiple regressions were carried out for each sub group of concepts (described in §3.2.1), the results of these analyses are shown in Tables 3.4 and 3.5 for the Technorati and Blogscope data respectively. It can be seen that the predictability of the three dimensions of the semantic differential varies considerably between these categories. For instance, for concepts within the settings category for the Technorati concepts, the correlation coefficient reached 0.72 ($p < .001$) for predicting Evaluation.

Table 3.4: The amount of variance accounted (R^2) for in the Technorati data from multiple regression for each factor for each subgroup of concepts using the predictors identified in the stepwise multiple regression.

Group	Dim	r	R^2	F	p
	E	.46	.21	33.54	<.001
Behaviour $n = 261$	P	.36	.13	12.90	<.001
	A	.14	.02	2.78	=.063
	E	.63	.40	118.12	<.001
Identity $n = 360$	P	.40	.16	23.29	<.001
	A	.32	.10	19.71	<.001
	E	.66	.44	102.28	<.001
Modifier $n = 262$	P	.54	.29	34.32	<.001
	A	.28	.08	10.60	<.001
	E	.72	.52	46.13	<.001
Setting $n = 89$	P	.48	.23	8.22	<.001
	A	.10	.01	0.47	*n/s*

Exploring the relationship of the interaction for the Activity factor, as shown in Figure 3.5, suggests that concepts that are considered passive are ones that are *a*) associated with a larger probability of something bad happening; and *b*) the probability of something good is smaller. This suggests that, in general, experience of higher probabilities of bad things occurring may be associated with lower potency scores and lower evaluation scores.

Table 3.5: The amount of variance accounted (R^2) for in the Blogscope data from multiple regression for each factor for each subgroup of concepts using the predictors identified in the stepwise multiple regression.

Group	Dim	r	R^2	F	p
	E	.47	.22	42.10	<.001
Behaviour $n = 305$	P	.36	.13	15.04	<.001
	A	.14	.02	2.86	=.05
	E	.61	.37	116.02	<.001
Identity $n = 392$	P	.40	.16	23.77	<.001
	A	.30	.09	18.62	<.001
	E	.71	.50	130.36	<.001
Modifier $n = 263$	P	.58	.34	45.20	<.001
	A	.30	.09	13.00	<.001
	E	.66	.44	42.76	<.001
Setting $n = 111$	P	.45	.20	8.82	<.001
	A	.10	.01	0.47	*n/s*

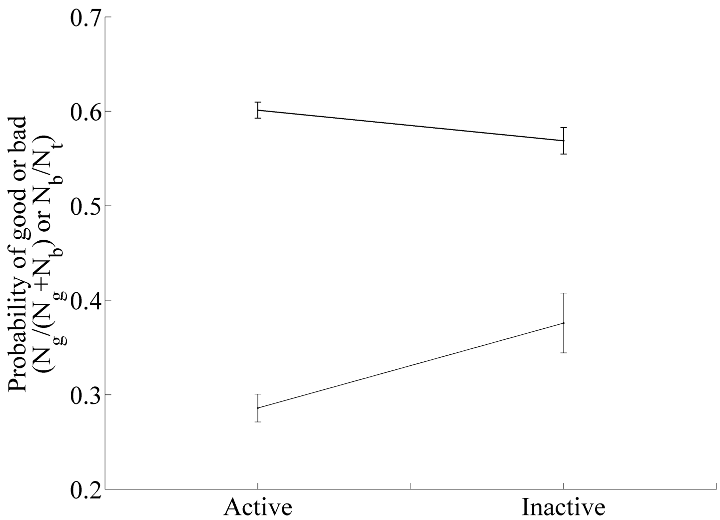


Figure 3.5: Form of the interaction probability of a good event occurring (thick line) and probability of a bad event occurring (thin line) for active and inactive concepts.

To explore the gross properties of the data and whether this pattern of positive and negative rewards was associated with the different levels of the semantic differential, the data were simply split into positive and negative values, based on the Evaluation and Potency dimensions, and the means calculated for good and bad things occuring as illustrated in Tables 3.6 and 3.7.

Table 3.6: Mean probability of a good event occurring for the semantic differential space bounded by positive and negative values of Evaluation and Potency.

P \ E	-	+
-	.54	.60
+	.56	.61

Table 3.7: Mean probability of a bad event occurring for the semantic differential space bounded by positive and negative values of Evaluation and Potency.

P \ E	-	+
-	.44	.30
+	.36	.29

Two analyses of variance, one for good probabilities and the other for bad, comparing the probabilities described above were carried out for each of the Blogscope and Technorati data sets. For Blogscope there was a main effect of quadrant on probability of bad $F(3, 1067) = 117.297$, $p < .001$ and post-hoc analysis using Tukey's test showed a significant difference between the mean probabilities of all of the quadrants ($p < .05$) with the exception of E+P- and E+P+

where there was no significant difference. There was also a main effect of quadrant on probability of good $F(3, 1067) = 125.801$, $p < .001$. Post-hoc analysis, again using Tukey's test, revealed a significant difference between the mean probabilities of all of the quadrants ($p < .05$) with the exception of E+P- and E+P+ where there was no significant difference.

For Technorati there was a main effect of quadrant on probability of bad $F(3, 968) = 83.932$, $p < .001$. Post-hoc analysis using Tukey's test again revealed a significant difference between the mean probabilities of all of the quadrants ($p < .05$) with the exception of E+P- and E+P+ where there was no significant difference. There was also a main effect of quadrant on Probability of good $F(3, 968) = 117.252$, $p < .001$. Post-hoc analysis, again using Tukey's test, revealed a significant difference between the mean probabilities of all of the quadrants ($p < .05$) with the exception of E-P- and E-P+ where there was no significant difference.

As can be seen in Table 3.6, where the probability of something being good increases, it is evaluated more highly, but, the effect on Potency seems negligible. This however, is not the same for the probability of something being bad. In this case, as can be seen in Table 3.7, the experience of higher probabilities of bad things attracts lower Potency and Evaluation, while lower probabilities of bad things occurring attracts higher Potency and Evaluation. This interaction is shown graphically in Figure 3.6.

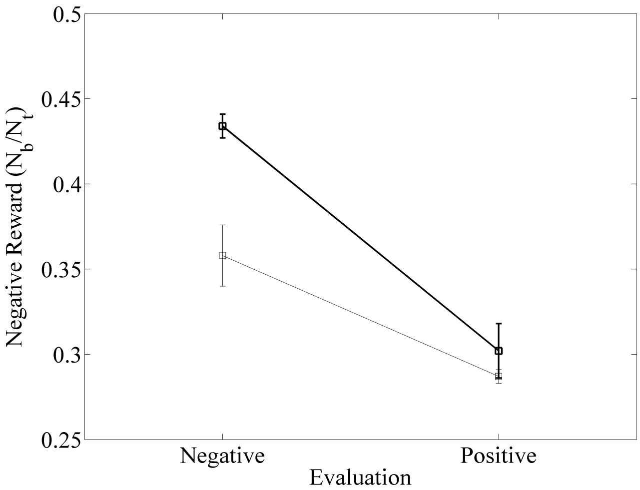


Figure 3.6: The interaction between potency and evaluation. Shown is the risk (probability of bad context) for positively evaluated (Evaluation > 0) and negatively evaluated concepts (Evaluation < 0). These are subdivided into concepts that are considered potent (Potency > 0; light line) and not potent (Potency < 0; dark line). In general more potent concepts seem to be associated with lower probabilities of risk. Positively evaluated concepts, whether of high potency or not seem to be associated with lower and similar probabilities of risk, whereas, for concepts evaluated negatively there seems to be a significant difference in risk between potent and impotent things, with potent things associated with lower risk and impotent things with higher risk..

3.4 Discussion

This chapter started out with the observation that when factor analysis is applied to rating scales, in a very large number of domains, three factors emerge: Evaluation, Potency, and Activity. Whilst this fact has been known for over 50 years, why it happens has been less clear. Here it is proposed that these three factors are, in fact, a representation of the history of reward associated with a concept and that this representation of reward is required whenever we need to compare and decide between alternatives. What does this representation look like?

The factor that almost always captures the most amount of variance in the semantic differential is Evaluation. This we found was very strongly correlated with the proportion of rewarded events that were positively rewarded (technorati $r = .57$, blogscope $r = .56$): Evaluation is to first approximation simply relative reward. Perhaps the most surprising thing about this is that there is essentially no correlation with the absolute proportion of positively rewarded events (technorati $r = -.08$, blogscope $r = -.10$): Evaluation does not measure the probability of good events happening, but the ratio of good to bad. The second characteristic associated with Evaluation is the (log) frequency of occurrence; things that happen more often are preferred to those that are infrequent. This is simply the well known Mere Exposure effect (Zajonc, 1980) and once the semantic differential is identified with reward, it is not surprising to find this, things that are commonly encountered (and therefore well understood) are preferred to things that are only rarely evaluated.

The second dimension, Potency, was most strongly related to the absolute probability of negative reward (technorati $r = -.37$, blogscope $r = -.40$). Potency essentially measures the risk (of bad things happening), making it clear why Potency needs to be represented for every object we can make decisions about; it is important to know not only the average reward associated with an option, but what the cost (risk/danger) might be in obtaining it.

Taking the findings for Evaluation and Potency together suggests that, in general, experience of higher probabilities of bad things occurring may be associated with lower potency scores and lower evaluation scores. Calculating the mean probability of a bad experience based on the gross positive and negative values for Evaluation and Potency (as described on page 57) shows that this is the case.

The data analysis has less to say about the third dimension of the semantic differential, Activity. Despite this, by identifying Activity with part of a representation of reward, one potential role is suggested. As well as a representation of the average relative reward and risk associated with an option, in order to make effective decisions, we also need to know how certain we are of this assessment. Uncertainty can come in two forms. Either we have limited and variable experience of a concept or, often more importantly, we have little control over the concept e.g. an object or a situation. Concepts of high Activity (e.g. ones that are fast, noisy and active) are very often associated with less certainty and less controllability than ones that are slow, quiet and inactive. This is not directly measurable in the data analysis, but to make effective decisions, we need to know the level of controllability/certainty associated with an option: Activity is proposed as that measure.

In conclusion, it is proposed that two representations from very different disciplines; reward (or utility), and the semantic differential, are in fact the same thing. Identifying the semantic differential as a characterisation of reward offers a solution to the main theoretical issue with the semantic differential (what it is), and suggests why it is ubiquitous across domains, languages, and cultures. Almost all objects, actions and contexts need at some time to be compared with others (do I attend to this object or that, do I perform this action or another...). To do this one needs to have an estimate of the reward associated with each alternative. It is proposed that the semantic differential is this representation and "connotative meaning" is to first approximation a summary of reward history. It also tells us potentially why three dimensions are needed: To make a choice, we need not only to know how rewarding an alternative is, but also how potentially dangerous it is and how sure we are of this. A single dimension will be blind to risk and uncertainty, unable to efficiently balance exploration and exploitation, and choose options that, whilst of very high average reward, could be associated with high levels of uncertainty and risk. The perils of making decisions that simply maximise reward, while ignoring risk and uncertainty, have been amply demonstrated by many of the transactions made before the credit crunch.

Finally, identification of the semantic differential with reward allows us to draw relationships with the underlying neurophysiology. As discussed in §2.7, reinforcement learning based models, particularly the Temporal Difference models (Sutton & Barto, 1998), have helped in developing an understanding of how we learn to make choices in an unknown world. The standard model uses a one dimensional representation of reward, associated with the action of dopamine (Schultz, 1998, 2010; Schultz et al., 1997). This dopamine associated dimension clearly corresponds most closely to the Evaluation dimension of the semantic differential. The strong and significant relationship between the probability of a good experience in a rewarded situation, $\frac{N_g}{(N_g+N_b)}$, and Evaluation is particularly interesting because it accords very well with findings in the literature that dopamine neurons respond only to rewards, providing a reward prediction error that scales to the relevant range of magnitudes (Tobler et al., 2005) maintaining the prior values represented in the prefrontal cortex. Accordingly, accurately predicted rewards and unrewarded events would be of no consequence and have no part to play in maintaining these values. Though most computational models only have a single dimension of reward, more recent work has begun to look at cost or punishment (or negative reward) and associated it with serotonin (Boureau & Dayan, 2010; Cools & Nakamura, 2010; Crockett et al., 2009) and corresponding to the Potency signal; and uncertainty, the Activity signal and associated it with acetylcholine and noradrenaline (Dayan & Yu, 2006; Yu & Dayan, 2005).

The correlations between the semantic differential and internet blog data do not, however, show that this relationship is causal. In order to establish this, an experiment where arbitrary shapes were associated with different distributions of positive, negative and neutral events was conducted. According to the present hypothesis, providing participants with a sufficient number of random trials of these shapes should be enough to affect the reward summary for a given shape and be capable of changing the semantic differential associated with it. This is the subject of Chapter 4.

Chapter 4

AlphaBet

4.1 Introduction

If the semantic differential represents our reward structure as hypothesised, it is reasonable to expect that the reward structure and thereby the relevant semantic differential, can be manipulated by changing the rewards that are applicable to a particular concept. Indeed, it was proposed at the end of Chapter 3 that any novel object can gain connotative meaning simply by being associated with a reward history. In other words, the connotative meaning or sentiment towards something can be learned through good and bad experiences. This chapter describes an attempt to manipulate the reward structure applicable to simple coloured shapes (see Figure 4.1) based on an experiment reminiscent of the Iowa gambling task and the use of different distributions of rewarding events.

The Iowa gambling task (Bechara, Damasio, Damasio, & Anderson, 1994) was introduced as an instrument for investigating patients with ventromedial prefrontal cortex damage and it is used to provide supporting evidence for the

Somatic marker hypothesis (Damasio, 1994). The Iowa gambling task is considered to simulate real-life decision making as it involves uncertainty as well as reward and punishment; it is also considered robust (for example, Dunn, Dalgleish, and Lawrence (2006) identifies more than 100 papers that use the Iowa gambling task). The task is simple, presenting participants with four decks of cards, typically on a computer screen, with each card indicating an amount that can be won or lost, however, the distribution of winning and losing amounts is different in each of the decks.

Participants are informed that they should attempt to maximise the amount of money they win in a long series of selections, but unknown to the participants, because of the different distribution of winning (good) and losing (bad) amounts, two decks are 'good' and will win in the long term and the other two are 'bad' and will lose in the long term (winning and losing amounts are only available after a card has been selected). Participants are though allowed to switch from any deck to another freely whenever and as often as they want to, but they are not told the number of card selections to make (in fact, the task is stopped after 100 trials) (Bechara et al., 1994; Damasio, 1994).

Typically, most healthy participants are good at sticking to the more advantageous decks after forty or fifty trials, suggesting that they have overcome the uncertainty and learned the distributions of reward. This is contrasted with patients that have prefrontal cortex damage, who do not appear to know which decks are disadvantageous, often continuing to lose overall. Interestingly, measurements taken using galvanic skin response shows that healthy participants exhibit a reaction to the bad decks after as few as ten trials and before conscious awareness the decks are bad (Bechara, Tranel, Damasio, & Damasio, 1996). In contrast, prefrontal patients do not exhibit the same reaction to potential losses. Bechara et al. (1994) and Damasio (1994) use these results in support of the somatic marker hypothesis.

It seems obvious to say that good or bad things occurring is dependent on the context or concept in question; to consider a few extremes for example, birthdays, weddings and Christmas are associated with good things happening, while house fires, and earthquakes are associated with bad things happening. It is also worth noting that these concepts are not exclusively associated with good or bad things, we all know of fights at weddings and arguments at Christmas, while disasters such as earthquakes can be the scenes of good acts and heroism. There are also those occasions when an outcome is neither particularly good or bad, but indifferent. What is clear is, that given certain contexts or concepts, there is a higher or lower probability of good or bad things happening; in other words, there is a distribution of good and bad events and these distributions can be different for different concepts. Accordingly, the purpose of the experiment below is to influence participants reward structure (and therefore semantic differential) by changing the rewards associated with novel concepts: In this case, after exposure to shapes with differing reward statistics, do the shapes gain a semantic differential like factor structure and are the characteristics of these dimensions predictable from their reward histories?

The approach used in the experiment is reminiscent of the Iowa gambling task to the extent that the objective is to maximise reward, based on repeated trials where gambles are made against six initially unknown distributions. However, the reward distributions were created so as to reflect the sort of thing that occurs in the world i.e. that expected events do not always occur and even when they do, a concept that is associated with predominantly good outcome can turn out to have bad things that occur and vice versa.

Two experiments were produced and based on pages presented via a 'web browser', the first was a rating experiment and the second a betting experiment. In experiment one it was assumed, consistent with the mere exposure literature (e.g. Zajonc, 1980; Duckworth et al., 2002), that even novel things can attract at least some sort of evaluation: Participants were therefore asked to carry out a semantic differential on the shapes only, in order to capture any pre-existing sentiment that the shapes might have and which is assumed to be similar for everyone.

4.2 Experiment 1 - Ratings

The mere exposure effect is the idea that people tend to prefer things that they have been repeatedly exposed to, presumably because they are more familiar (Zajonc, 1980). The mere exposure effect has been demonstrated with words, faces, paintings etc. It has also been shown that novel things are (or can be) evaluated after being displayed very briefly (250ms in the case of Duckworth et al. (2002)). In this initial experiment it is assumed that the coloured shapes are likely to be familiar to participants in one way or another and/or will be evaluated rapidly; in other words, the shapes were expected to have pre-existing sentiment, accessible in the form of a semantic differential, associated with them. It was reasoned that any manipulation of semantic differential values should be with reference to this base line.

4.2.1 Method

4.2.1.1 Participants

Forty participants volunteered to take part in the experiment. Two participants declined to provide details of both age and sex with a further three participants declining to provide details of sex. Of the thirty eight participants providing their age, the mean was 27.74 years, SD=8.45 (female M=27 years, SD=9.38, N=21; male M=27.43 years, SD=6.85, N=14). All participants reported normal or corrected to normal vision and all participants provided their informed consent prior to commencement of the experiment.

4.2.1.2 Materials

The experiment was produced as a PHP[1] script. This provided 'web browser' based access to the experiment on demand, from any location with access to the server where the script was located. The shapes that participants were asked to rate are shown in Figure 4.1 below.

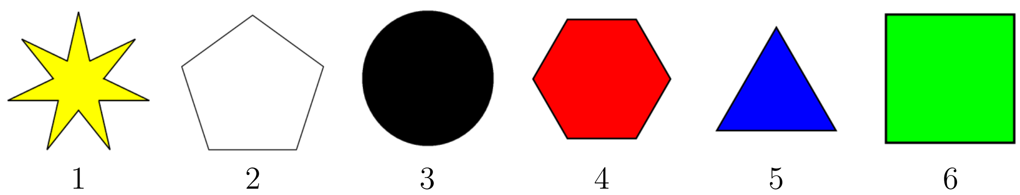


Figure 4.1: The shapes used for the rating and AlphaBet experiments together with the numbers used to identify each shape.

Nine scales, based on slider bars, were used to rate each shape (see Table

[1]PHP is a powerful, fast and widely used general purpose server side scripting language that is particularly suited to Web development. PHP is available from www.php.net and is distributed under an open source licence.

4.1), these scales were the three identified as the most heavily loaded for each of Evaluation, Potency and Activity in a factor analysis carried out by Osgood and Suci (1955, p336) and were typical of reliable scales. Each slider was initialised to the middle of the scale and produced values between 0 and 100. Participants moved the slider towards one extreme or the other, as indicated by the noun, in order to record their rating.

Table 4.1: The nine scales used for rating the shapes

Evaluation	Potency	Activity
Awful - Nice	Light - Heavy	Slow - Fast
Ugly - Beautiful	Weak - Strong	Passive - Active
Dirty - Clean	Small - Large	Dull - Sharp

4.2.1.3 Design

This experiment used a repeated measures design. Independent variables were the shapes presented for rating and the dependent variables were the rating scores given by the participants using the sliders. The rating page was a simple, one screen design that collected brief demographic information and presented all of the shapes to be rated. For each shape nine sliders were presented to the participant with each slider presented between two nouns of opposite meaning. A button was included at the bottom of the screen in order to allow results to be submitted. The rating page is shown in Figure 4.2.

4.2.1.4 Procedure

On entering the URL for the experiment, participants were presented with brief instructions as follows: "*On the following page you will be presented with six different shapes. Please take as long as you need to provide a considered rating for each of the shapes based on each pair of words, using the sliders provided. It is important to provide a rating for ALL of the shapes using ALL of the sliders.*". Participants were asked to rate every shape by using every slider bar. Once all of the ratings were completed, participants were asked to press the submit button at the bottom of the page and were then presented with a thank you page, from which they could navigate away from the experiment.

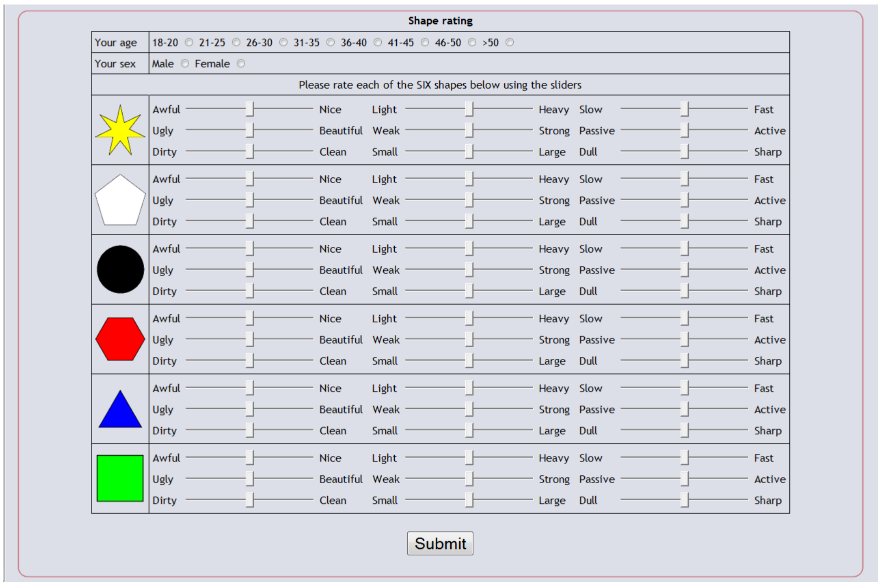


Figure 4.2: The shape rating page consisting of two questions collecting demographic information, then the six experimental shapes each with 9 sliders for participants to indicate their ratings between the extremes indicated by the two opposite words.

4.2.2 Results

The rating data for one participant was removed due to incomplete use of all of the rating scales, leaving ratings for thirty nine participants for the analysis.

4.2.2.1 Factor analysis

The data collected from participants for each of the rating scales were first checked for 'factorability' using the Kaiser-Meyer-Olkin measure of sampling adequacy (.68) and with Bartletts test of sphericity ($\chi^2(36)$=674.26, p <.001).

On the basis of these checks, factor analysis was carried out on the data. Three principal components with eigenvalues greater than 1 were revealed, accounting for 67.36% of the variance in the ratings. Table 4.2 shows loadings of the component matrix following a varimax rotation.

Table 4.2: Rotated Component Matrix showing factor loadings greater than .5. Extracted using principal component analysis, with varimax rotation which converged in five iterations.

Scale	Factor 1	Factor 2	Factor 3
Awful - Nice		0.867	
Ugly - Beautiful		0.878	
Dirty - Clean		0.518	
Light - Heavy			0.736
Weak - Strong			0.802
Small - Large			0.579
Slow - Fast	0.867		
Passive - Active	0.856		
Dull - Sharp	0.743		

Though the factors were in a slightly different order from previous semantic differentials, three factors, consistent with previous findings, were revealed by the factor analysis, each containing three variables. Factor one related to rating scales indicating Activity, Factor 2 to rating scales indicating Evaluation and Factor 3 to rating scales indicating Potency. The factor score coefficients calculated for each scale are shown in Table 4.3.

Table 4.3: Factor score coefficients.

Scale	Factor 1	Factor 2	Factor 3
Awful - Nice	0.004	0.399	0.076
Ugly - Beautiful	-0.067	0.59	-0.007
Dirty - Clean	-0.046	0.05	-0.113
Light - Heavy	-0.023	-0.014	0.451
Weak - Strong	0.057	0.029	0.352
Small - Large	-0.074	0.012	0.225
Slow - Fast	0.371	-0.074	0.079
Passive - Active	0.425	-0.052	0.127
Dull - Sharp	0.219	-0.019	-0.052

If the connotative meaning of these shapes were to be altered, then presumably the results found here are the initial values that need to be changed. Manipulating the connotative meaning of the shapes was attempted in Experiment 2.

4.3 Experiment 2

If, as proposed, any novel object can gain connotative meaning simply by being associated with a reward history, then a semantic differential carried out after participants have attempted to maximise their reward based on arbitrary shapes that were paired with differing reward statistics, should reflect this. The Iowa gambling task offers some evidence for this alteration of reward structure, in that it shows that participants can recognise 'good' decks after reasonably modest numbers of samples. In experiment 2 the experimental shapes were each assigned a different reward structure, but, unlike the Iowa gambling task, a win or a loss did not occur on every gamble.

In order to try and better reflect real world choices, that is, sometimes when we make a choice there is an indifferent outcome, a win or a loss only occurred based on certain probabilities (see Table 4.4). Of the different reward structures two shapes had distributions where the probability of something happening was high and the probability of winning was low, two shapes had distributions where the probability of something happening was high and the probability of winning was high and two shapes had distributions where the probability of something happening was neither low nor high and the probability of winning was also neither low nor high. The result of these different distributions is that two shapes were 'bad', two shapes were 'good' and two shapes were indifferent,.

Participants were repeatedly exposed to the distribution of rewards assigned to the experimental shapes on the basis of a betting game and since they would become equally familiar with each shape it was reasoned that the reward associated with each shape and which they would experience, would determine the changes in sentiments towards the shapes.

4.3.1 Method

4.3.1.1 Participants

A total of sixty five participants volunteered to take part in the experiment. Four participants declined to provide details of both age and sex with a further three participants declining to provide details of sex. Of the sixty one participants providing their age, the mean was 26.4 years, SD=8.5 (male M=26.4 years, SD=7.2, N=20; female M=26.1 years, SD=8.9, N=38). All participants reported normal or corrected to normal vision and all participants provided their informed consent prior to commencement of the experiment.

4.3.1.2 Materials

As with the rating experiment above, this experiment consisted of a set of web pages produced as PHP scripts, which provided 'web browser' based access to the experiment on demand, from any location with access to the server where the script was located. The web pages are shown in Appendix E and provide a means for participants to place bets on the shapes shown in Figure 4.1 and receive feedback on those bets. At the end of the experiment participants were also resented with a ratings page consisting of same nine rating scales as in the initial experiment above and shown in Table 4.1.

4.3.1.3 Design

This experiment used a repeated measures design. Independent variables were the shapes and the position that they were presented on the screen. The dependent variables were the amounts of the bets that were placed by the participants during the experiment and the rating scores provided by the participants at the end of the experiment using the sliders.

The shapes were presented in distinct horizontal positions across the screen. The presentation locations were randomised for each participant but maintained for an experimental session, as it was reasoned that the position of the shape might influence betting amounts or ratings. The order that the shapes were presented in was randomised. For each random presentation of a shape whether a win or loss would happen was calculated using the distributions shown in Table 4.4.

Table 4.4: Means for distributions governing the occurrence of either a win or loss. Each distribution has a standard deviation of 0.05.

Shape	Probability of win or loss occurring	Probability of win	Probability of loss	Distribution type
	.85	.10	.90	Bad
	.65	.25	.75	Bad
	.55	.50	.50	Indifferent
	.55	.50	.50	Indifferent
	.65	.75	.25	Good
	.85	.90	.10	Good

First, whether an event would happen was calculated and then, if an event was to happen, whether that event would be a reward or a punishment was calculated. Distributions were created so that shapes with the highest probabilities of an event happening attracted the highest probability of reward or punishment.

4.3.1.4 Procedure

The experiment was accessible via the internet so that participants could choose to undertake it at a time and location that was convenient to them. Once the URL for the experiment was entered and consent provided, participants were presented with a page of instructions (see Figure E.1 in Appendix E). The instructions page explained that the experiment was concerned with making choices when some information was unknown, but could be learned. It was further explained that choices would be associated with six shapes and that each shape could generate a win, a loss, or nothing may happen. It was emphasised that there were no right or wrong answers and that what was of interest was 'gut reactions', in order to encourage quick responses.

For each presentation of a shape the participant was asked to risk a proportion of the banked total on the outcome, with the aim of building a bank total as large as possible. Along with the first randomly presented shape, participants were informed that they had been given an initial banked total of 100 (see Figure E.2).

If a win (good outcome) occurred with the shape the participant was returned the stake and won the equivalent amount, increasing the banked total; if a loss (bad outcome) occurred with the shape the participant lost the amount bet from the banked total; and if the event was indifferent the participant was returned the stake and the banked total was unaltered. After specifying the amount to bet on the currently displayed shape and pressing the next button, the participant was given feedback about the outcome of the event involving the shape using smiley faces displayed for a short duration (1000ms).

For a win a conventional smiley face was displayed (with the mouth turned up), for a loss a sad 'smiley' face was displayed (with the mouth turned down) and for an indifferent outcome a neutral 'smiley' face was displayed (with the mouth displayed as a horizontal straight line). An example of this feedback is shown in Figure E.3. After the feedback had been provided the experiment continued with the next randomly generated shape (see Figure E.4).

The experiment was self paced with no limit to the number of bets that participants were able to make and the experiment could be terminated at any time by pressing the 'Finish experiment' button; participants were, however, asked to make at least fifty bets as this quantity was thought sufficient to ensure that they experienced the full range of the experimental shapes and their distributions. It was possible for participants to go 'bust' if a stake of 100% of the banked total was placed on what turned out to be a loss, in which case the information gathering screen was displayed automatically, as if the 'Finish experiment' button had been pressed.

On finishing the experiment (or going bust), participants were presented with a rating page requesting brief demographic information and presenting all of the shapes to be rated based on the same rating scales as the initial ratings experiment (see Figure E.5 and Table 4.1). As with the initial experiment, participants were asked to rate each shape by using a slider bar situated between two nouns of opposite meaning. Each slider was initialised to the middle of the scale and the participants moved the slider towards one end (extreme) or the other to indicate their rating.

Once all of the ratings were completed, participants were asked to press the submit button at the bottom of the page and were then presented with a thank you page, from which they could navigate away from the experiment.

4.3.2 Results

Of the sixty five participants that completed the experiment a total of thirty eight either failed to complete the requested fifty bets or provided fewer than fifty percent of the ratings that were required, leaving twenty seven for the remainder of the analysis. Eighteen out of the thirty eight excluded participants chose not to complete the fifty bets requested, while eight were unable to complete enough bets because were they went bankrupt and taken automatically to the rating page; the remainder provided fewer than fifty percent of the ratings that were requested.

More than 57% of the excluded participants had been presented with every shape fewer than five times, seven participants having had zero or only one presentation of at least one of the shapes. Overall, the mean number of trials undertaken was M=108.56, SD=46 with each shape presented for the mean numbers of trials shown in Table 4.5.

4.3.2.1 Factor scores

Factor scores were calculated using the factor score coefficients established in the factor analysis for the initial ratings experiment (see Table 4.3).

To make the results more comprehensible the reward distributions were reduced from six to three by combining the good, bad and indifferent distributions that were identified in Table 4.4; the mean reward values of the combined distributions are shown in Table 4.6. Accordingly the results for the initial ratings (experiment 1) and the present experimental ratings were combined to provide results sets for shapes  & ,  &  and  .

Independent samples analyses of variance, with dependent variables Evaluation, Potency and Activity, comparing the differences between each distribution of the initial semantic differential and the experimental semantic differential, were carried out on the data. The analyses revealed significant interactions for the Evaluation ($F(2,390) = 3.83, p = .022$) and Potency ($F(2,390) = 3.10, p = .046$) dimensions, but not for the Activity dimension ($F(2,390) = 0.48, p = n/s$). Post-

Table 4.5: Descriptive statistics for the presentations of shapes.

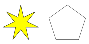
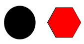
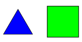
Shape	Mean presentations	SD
✴	18.74	7.29
⬠	17.41	8.89
●	17.67	8.19
⬢	19.59	6.72
▲	16.56	7.59
■	18.59	8.13

Table 4.6: Combined means for distributions governing the occurrence of either a reward or punishment. Each distribution has a standard deviation of 0.1.

Shape	Probability of win or loss occurring	Probability of win	Probability of loss
✴ ⬠	.750	.175	.825
● ⬢	.550	.500	.500
▲ ■	.750	.825	.175

hoc analysis, carried out using the Tukey HSD test, for each of the combined shape distributions revealed that for the Evaluation dimension the experimental semantic differential was significantly lower than the initial semantic differential for the combined distributions of shapes ✴ & ⬠, but not the other combined distributions (results of the analysis and descriptive statistics are given in Table 4.7). For the Potency dimension the experimental semantic differential was significantly lower than the initial semantic differential for the combined distributions of shapes ✴ & ⬠, ● & ⬢ but not ▲ & ■. Both semantic differentials are shown in Figure 4.3, where the differences can be seen, however, perhaps more illustrative are just the differences shown for each factor of the semantic differential in Figure 4.4. Note that whether the pairwise differences between the combined shapes are or are not significant is not of interest here, only that the series of betting trials has influenced the semantic differential.

Table 4.7: Descriptive statistics and results of Tukey's test for comparisons of initial and experimental ratings for each of the semantic differential factors for the combined distributions.

		Initial		Experimental			
Distribution	Factor	Mean	SD	Mean	SD	Diff	p
	E	50.60	21.74	35.49	21.45	-15.10	$< .001$
	P	55.50	18.64	51.52	19.10	-3.99	n/s
	A	47.32	27.92	45.20	18.66	-2.11	n/s
	E	52.69	18.54	42.10	14.68	-10.59	$= .015$
	P	72.99	15.65	63.75	13.17	-9.24	$= .017$
	A	42.16	18.43	40.34	11.58	-1.82	n/s
	E	60.54	15.30	57.97	16.49	-2.56	n/s
	P	60.43	16.77	61.27	10.72	0.83	n/s
	A	46.76	20.25	48.95	11.24	2.19	n/s

4.3.2.2 Reward histories

The numbers of good (wins), bad (losses) and indifferent outcomes were also recorded for the experiment and these were collapsed across participants and shapes in order to provide a basis to investigate the relationship between the reward histories for each of the shapes and the semantic differential for the shapes. In order to do this the same measures of reward as for the blog search data discussed in Chapter 3 §3.3, were calculated and correlated with Evaluation, Potency and Activity.

Although they were modest, significant correlations suggest that both Eval-

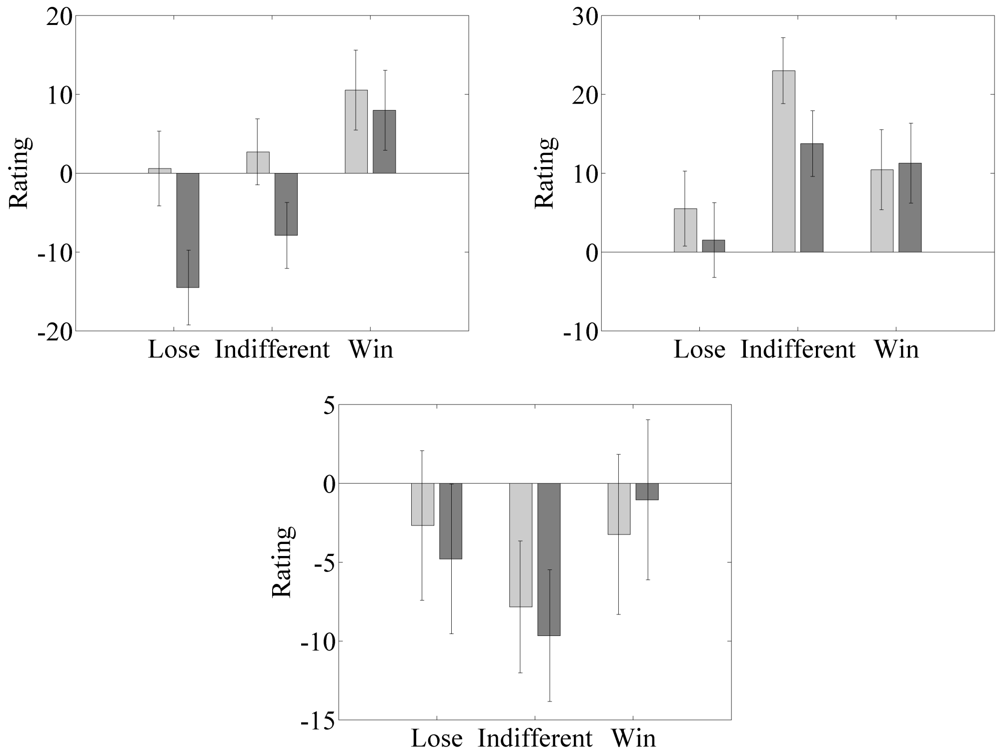


Figure 4.3: Ratings for the initial (light bars) and experimental (dark bars) semantic differentials. The left panel shows ratings for the Evaluation dimension for each combined distribution of shapes, the right panel ratings for the Potency dimension and the bottom panel the activity dimension. Ratings were between -50 to 50 with 0 indicating neutral. Error bars are standard error of the mean.

uation and Potency were changeable by being exposed to the different reward distributions in the shapes experiment. In particular relative reward, $\frac{N_g}{N_g+N_b}$, was correlated with the change in Evaluation, $r(159) = 0.43, R^2 = 0.18, F(2, 159) = 18.01, p < 001$; and negative reward, $\frac{Nb}{Nt}$, was correlated with Potency, $r(159) = -0.19, R^2 = 0.03, F(2, 159) = 2.84, p = .049$. There was no significant correlation for Activity, $r(159) = 0.11, R^2 = 0.01, F(2, 159) = 0.89, n/s$

Taken together, these results suggest that, at least for the Evaluation and

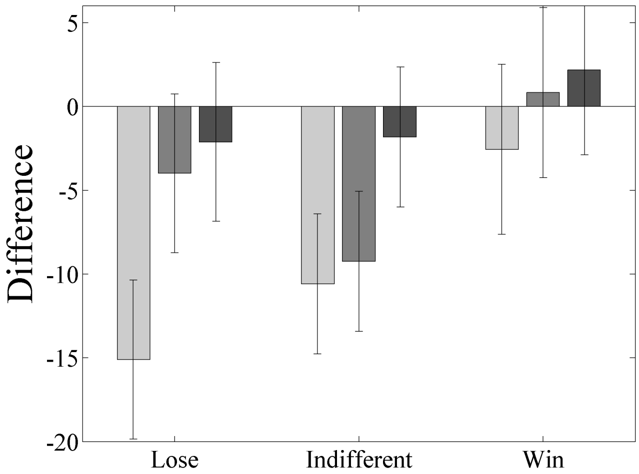


Figure 4.4: The differences between the initial and experimental semantic differential for Evaluation (lighter left bars), Potency (darker middle bars) and Activity (darkest right bars) for each combined shape distribution. Error bars are standard error of the mean.

Potency dimensions, the semantic differential has been significantly influenced by the series of bets carried out by participants.

4.4 Discussion

This chapter addresses the idea that, if the semantic differential represents our three dimensional reward structure, as proposed in Chapter 3, it will be possible to influence that reward structure and thereby the associated semantic differential through exposure to stimuli associated with varying rewards. Based on an approach reminiscent of the Iowa gambling task, which has been used extensively for investigating the Somatic Marker hypothesis, the AlphaBet experiment achieved this using arbitrary coloured shapes that were associated with differing reward distributions, in a betting game. It was hypothesised that providing participants with a sufficient number of random trials of these shapes should be enough to affect the reward summary for a given shape and be capable of changing the semantic differential associated with it.

An initial rating experiment, using the semantic differential technique, captured pre-existing connotative meaning for each of six arbitrary coloured shapes. The results of analysing the ratings provided by participants showed that, a semantic differential was formed with the expected dimensions of Evaluation, Potency and Activity. Although there is no discernible pattern across the shapes in the context of a simple rating experiment, participants were shown to have preconceived ideas about such simple things as coloured shapes. It might be, of course, that the shape/colour combinations used were novel in which case presumably prior associations were used to assess the shape and arrive at a rating. For example the shape/colour combination of yellow star was associated with good things, which would seem to be reflected in everyday meaning (he/she is a star, pop star, sports star, have a gold star etc.). Alternatively, it might also be that, for example, something green would be rated positively, as it was, since green is most commonly associated with nature. Accordingly, by expressing any post-exposure ratings relative to this baseline the effects of any pre-existing colour or shape biases associated with these shapes be minimised or eliminated from experiment 2.

Experiment 2 exposed participants to the different rewards statistics of the shapes through a betting experiment. Unfortunately a sizeable proportion of participants' data was eliminated from the analysis, due to too few bets, as there was concern that some shapes may have been insufficiently sampled, and insufficient use of the rating scales. Nonetheless, even though statistics confirmed insufficient sampling by some, the remaining number of participants was sufficient for the analysis. A further semantic differential was calculated and found to be significantly different from the semantic differential produced from the initial ratings; this can only be attributable to the participants' reward structure and hence semantic differential, having been influenced by the reward statistics of the experimental shapes. Perhaps more interesting are the differences between the combined shape distributions for each of the semantic differential dimensions, shown in Figure 4.4.

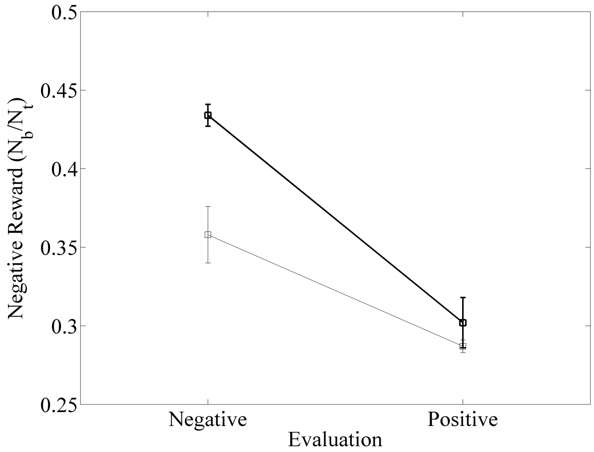
The dimension labelled as reward in Chapter 3 (Evaluation) showed the greatest change between the experiments, with the 'losing' shape combination attracting the largest difference and, as can be seen in Figure 4.4, being significantly different to the difference in reward for the 'winning' shape combination. Given that the participants were exposed to approximately the same numbers of trials for every shape, suggests that 'more notice' was taken of the more negative distributions, which might be considered to be consistent with loss aversion (Kahneman & Tversky, 1979), where sensitivity to a loss is more acute than an equivalent win.

Sensitivity in the 'indifferent' shapes combination included the risk dimension (again as labelled in Chapter 3) (Potency), which showed a significant difference between the experiments. Considering this distribution of shapes as risk or danger is interesting because while these shapes had equal probability for a good or bad outcome, they also had equal probability of something happening or nothing happening. The result of this is that, although taken together these shapes are not as 'objectively bad' as those shapes in the losing distribution, they are somewhat tedious and considerably more difficult to predict; consistent with the argument put forward in Chapter 3 this constitutes greater risk and is presumably the reason for the worse ratings. As already observed in Chapter 2, it is important to be more sensitive to threats and potentially dangerous things since doing so provides better opportunities to survive and reproduce. This seems to be evident even in the context of a simple betting experiment.

The comparison above raises an interesting question of whether the overall pattern of reward/Evaluation and risk/danger/Potency is similar for the internet blog data and for the shapes experiment data. The present data set obviously has many fewer data points than the Internet data set, therefore, to see if the same pattern of positive and negative rewards was associated with the different levels of the semantic differential, the data were simply split into 4 bins, calculating mean probabilities for bad things occurring for each quadrant described by axes representing Potency on the vertical 'y' axis and Evaluation on the horizontal 'x' axis. The results of this binning process are shown in Figure 4.5, which shows this relationship for the internet blog data and shapes experiment data respectively.

As can be seen, there is a correspondence between the rewards associated with certain locations in the semantic differential, both when induced experimentally and when induced by the nature of our environment: The semantic differential appears simply to be a summary of the reward history.

However, although the winning shape distributions are the only ones to have a reduced level of activity (and as argued here uncertainty) it is not significant; neither the analysis of internet blog data, nor the AlphaBet shapes experiment, seem to directly address anything to do with Activity. Based on the idea that this latent factor is more accurately described as control or certainty, Chapter 5 attempts to focus on this dimension, by keeping other variables, such as the



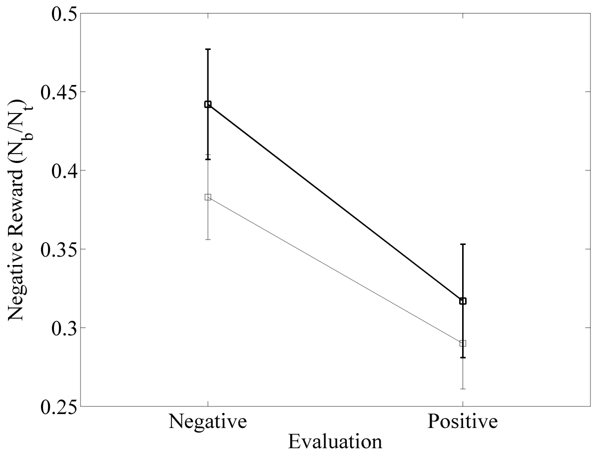


Figure 4.5: The relationship between potency and evaluation for Technorati data set (top panel) and AlphaBet data set (bottom panel). Shown is the risk (the probability of a bad event) for positively evaluated (Evaluation > 0) and negatively evaluated concepts (Evaluation < 0). These are subdivided into concepts that are considered risky or potent (Potency > 0; light line), and not risky (Potency < 0; dark line).

shape, static.

Chapter 5

Triangles Experiment

5.1 Introduction

The AlphaBet shapes experiment in Chapter 4 attempted to show that the relationship hypothesised in Chapter 3 between connotative meaning and the reward structure of individuals, as represented by arbitrary shapes, can be manipulated. The AlphaBet experiment, reminiscent of the Iowa gambling task (Bechara et al., 1994; Damasio, 1994), required participants to use the experience that they gained during the course of the experiment to maximise the amount of money/points that could be won from betting on the appearance of six coloured shapes; the experiment provided explicit rewards for the participants following a bet on a shape; where either nothing happened (according to different probability distributions for each shape) or if it did participants could lose their stake or win the equivalent amount (also according to different probability distributions for each shape).

This scenario, however, is not representative of many of the situations where we make choices, especially in everyday life: Often there is no explicit reward or feedback. We may perhaps have only vague feedback information or even just a subjective belief that a correct choice has been made. Consider, for example, many of the choices that are made when deciding and buying clothes or footwear; a new pair of shoes may feel reasonably comfortable and look nice, but this is a feeling which is subjective and not like winning a bet. In addition, the Alphabet experiment did not help a great deal in gaining a better understanding of the Activity or uncertainty dimension.

In order to attempt to address both of these issues in one experiment, and to remove any pre-conceived ideas that participants might have about the stimuli (it was clear, for example, that the initial ratings taken in the AlphaBet experiment showed that participants did have pre-conceived ideas about the shape and colour combinations that were used), the same shape and colour was used, a black triangle, but displayed in different orientations depending on a pre-determined distribution. In order to try and manipulate Activity the triangles experiment required participants to predict the size that a shape would be the next time it was displayed in the same position. Everything was held constant apart from the way that the size of each presentation was manipulated. In addition, for the predictions that the participants made, no feedback was provided either when the prediction was made or when the shape was displayed the next time.

Accordingly, the purpose of the experiment below is to influence participants reward structure (and therefore semantic differential) by changing the rewards associated with the concepts, that is, the shapes. However, in this experiment reward is considered to be associated with predictability since there are no explicit rewards. In this case, after exposure to shapes with differing predictability, do the shapes gain a Semantic Differential like factor structure and are the characteristics of, particularly the Activity dimension, predictable from their reward histories?

5.2 Method

5.2.1 Participants

Nineteen participants volunteered to take part in the experiment and were recruited from amongst the post graduates and staff in the School of Experimental Psychology. There were seven male and twelve female participants with mean age 27.53 years *SD*=6.85. All participants reported normal or corrected to normal vision and all participants provided their informed consent prior to commencement of the experiment.

5.2.2 Materials

Because it has been reported that people have better memory for content that is shown on a larger screen (e.g. Detenber & Reeves, 1996) it was considered that participants would have the best chance of learning the distributions of triangle shapes if they were projected onto a large screen. The equipment used to display each experimental trial consisted of a standard desktop computer connected to a high resolution Canon XEED SX6 projector with the user interface presented as an interactive computer program written using the MATLAB Graphical User Interface, shown in Figure 5.1. The area where the trials were displayed was 700 millimetres by 700 millimetres, giving visual subtense of 20.05 degrees for both width and height.

Within a large (white) square presented on a mid grey background, a black

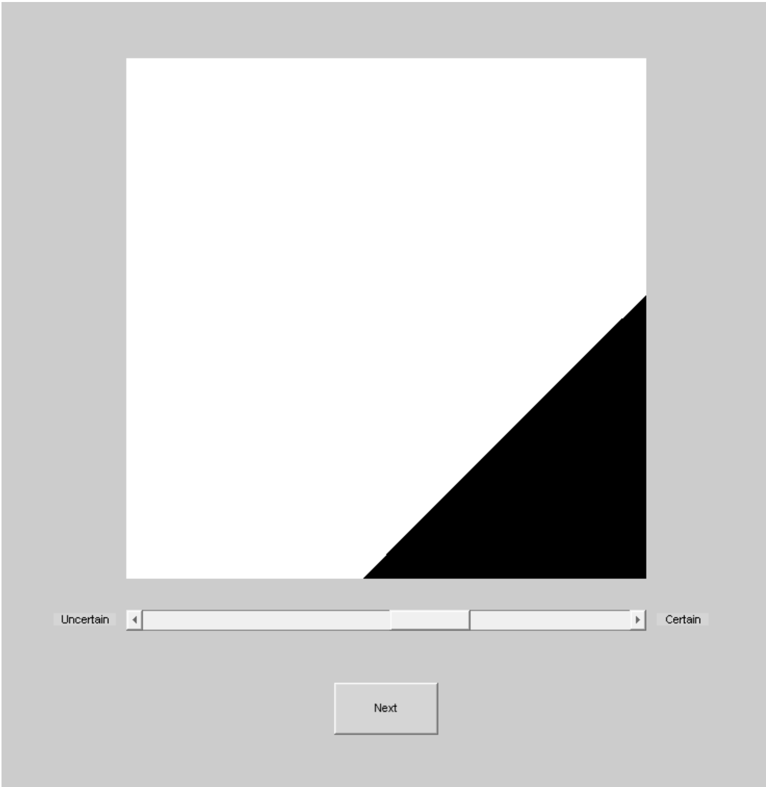


Figure 5.1: Illustration of how the triangles were presented.

triangle was displayed randomly (one at a time) at every corner. The size of the triangle that was displayed in each trial was governed by four, predefined, distributions. The distributions assigned to the triangles varied in size as follows *a*) random; *b*) probabilistically increasing; *c*) uniform; *d*) increasing in equal increments for first 50% of trials then decreasing in equal increments. The distributions are shown graphically in Figure 5.2:

5.2.3 Design

In a repeated measures design, independent variables were the four triangles that were presented in each corner of the display, each of which had a different

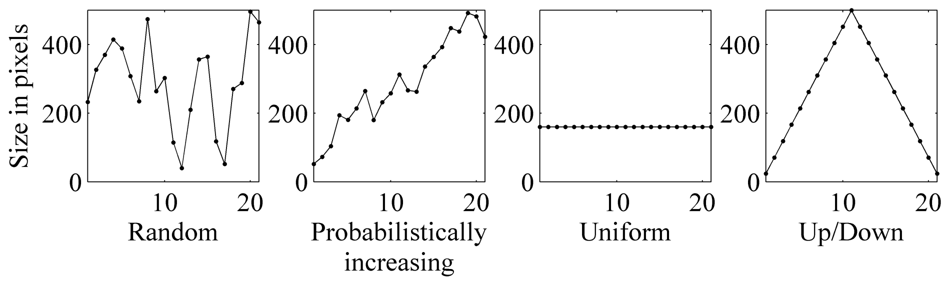


Figure 5.2: The size changes that were assigned to each of the four triangles. The x-axis represents the trial number and the y-axis the size of the triangle in pixels for each trial.

distribution of change in size (see Figure 5.2). Dependent variables were size predictions for the next triangle, presented in the same position and made graphically using the mouse wheel, the confidence levels for each prediction entered using a slider and Semantic Differential rating scores given by the participants at the end of the experiment. For each triangle participants completed a twenty scale Semantic Differential with each of the scales presented between two nouns of opposite meaning. Participants were asked to complete the ratings for each triangle on a separate sheet of pre formatted paper (shown in Appendix F) by marking a seven point Likert type scale. The order and direction of the scales were counterbalanced.

5.2.4 Procedure

Participants were asked to make themselves comfortable and after completion of an informed consent form, requested to read some brief instructions for the experiment. The instructions were also reiterated verbally. Participants were informed that they would be presented with a triangle in one of the four corners of a (white) square, as shown in Figure 5.1, and that the experimental task was to predict the size of the triangle the NEXT time it is presented in that position. Predictions were made by using the mouse wheel to make the triangle, that was currently displayed, larger or smaller. Once a prediction had been made, participants were also asked to indicate how confident they were of their prediction using a slider bar towards the bottom of the screen. Once the participant was happy with the prediction and confidence rating the 'Next' button at the bottom of the screen could be pressed to display the next trial. The experiment was self paced with no time pressure, however, participants were asked not to spend too long thinking about their predictions, but rather to go on 'gut' feel.

The experiment consisted of twenty presentations of a triangle from each distribution making eighty presentations in total. The presentations for each distribution were displayed in order, but which distribution was displayed for a particular trial was randomised. The overall task, therefore, consisted not only of learning the distributions in order to be able to make predictions for the next size to be displayed for a particular corner, but also remembering the last size of triangle that was displayed for a particular corner. Once all of the presentations had been completed participants were asked to complete the Semantic differential ratings for each triangle. Following completion of the ratings, participants were debriefed and thanked for their participation.

5.3 Results

The scale data were collapsed across participants providing eighty mean scale values and a factor analysis carried out. The analysis revealed the three components usually associated with the semantic differential, although they were found in the order Activity, Potency and Evaluation; Table 5.1 shows loadings of the component matrix following a varimax rotation.

Table 5.1: Rotated Component Matrix showing factor loadings greater than .8. Extracted using principal components, with varimax rotation which converged in five iterations.

	Factor 1	Factor 2	Factor 3
Active-Passive	0.988		
Calm-Excitable	0.984		
False-True	-0.983		
New-Old	0.932		
Usual-Unusual	0.844		
Slow-Fast	0.843		
Masculine-Feminine		0.977	
Weak-Strong		0.966	
Colorless-Colorful		0.893	
Angular-Rounded		0.892	
Savory-Tasteless			0.993
Important-Unimportant			-0.900
Wise-Foolish			0.855
Unsuccessful-Successful			0.822

This analysis may not seem entirely appropriate given the reduction in sample size through using mean values across participants, however, previous research regarding sample sizes and factor/principal components analysis (e.g. Guadagnoli & Velicer, 1988; Osborne & Costello, 2004) has established that the validity of the analysis is most importantly dependant on component loadings and absolute sample size. The analysis above appears to be valid based on previously published data, most particularly that each of the latent factors had loadings of $> .8$ for the first four components for each factor, and also to the extent that, in common with previous semantic differentials, the Evaluation, Potency and Activity factors were found.

The factor scores were then calculated, using the factor score coefficients calculated above, for each distribution for every participant. A repeated measures analysis of variance was then carried out on the scores for each latent factor (Activity, Potency and Evaluation) with distribution as the factor for the analysis. For the Activity factor, Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2(5) = 11.45, p = .04$, so degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .68$, which revealed a main effect of distribution, $F(2.03, 36.50) = 4.56, p = 0.017$, see Figure 5.3. Post-hoc analysis, using Tukey's test, revealed that the uniform distribution was rated significantly lower (or less active) than the other distributions (Random difference = $3.60, p = .002$; Probabilistically increasing difference = $4.05, p < .001$; Up/down difference = $6.73, p < .001$).

For Potency, Mauchly's test indicated that the assumption of sphericity had also been violated, $\chi^2(5) = 20.26, p < .001$, so degrees of freedom were again corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .63$, which revealed a main effect of distribution, $F(1.88, 33.89) = 4.43, p = 0.021$, see Figure 5.4. Post-hoc analysis, using the Tukey test, showed that the random distribution was rated significantly higher, in other words more potent, than the other distributions (Probabilistically increasing difference = $3.60, p = .002$; Uniform difference = $2.60, p = .02$; Up/down difference = $4.56, p < .001$).

For Evaluation, Mauchly's test again indicated that the assumption of sphericity had been violated, $\chi^2(5) = 19.00, p = .002$, so degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .59$, which showed that there was no significant effect of distribution on Evaluation, $F(1.76, 31.76) =$

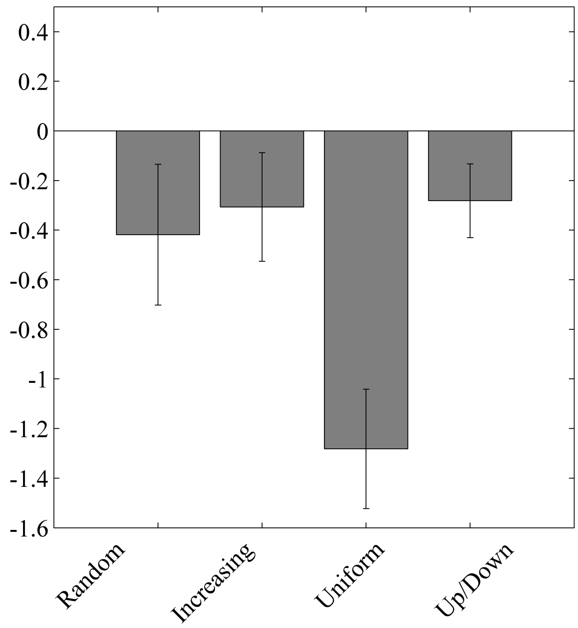


Figure 5.3: The mean factor scores for Activity for the distributions of size changes. The uniform distribution was found to be significantly lower than each of the other distributions. Error bars are standard error of the mean.

$3.18, p = n/s$, see Figure 5.5.

Further repeated measures analyses of variance were also carried out on the probability of a correct prediction, the certainty associated with predictions and the squared prediction error, each with distribution as the factor for the analysis. For the probability of correct predictions Mauchly's test of sphericity was significant, $\chi^2(5) = 11.70, p = 0.039$, and as for the previous tests above, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .98$. Using the corrected degrees of freedom revealed a significant effect of distribution, $F(2.94, 1113.78) = 16.53, p < .001$. The sphericity assumption

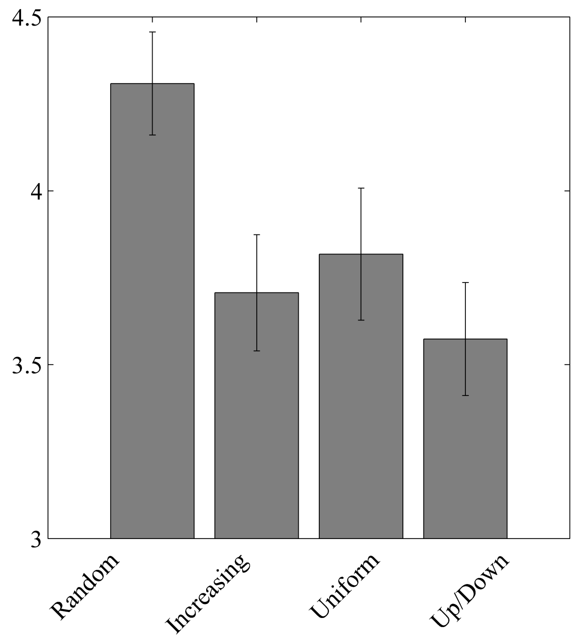


Figure 5.4: The mean Potency factor scores for distributions of size changes. The random distribution was found to have the highest level of potency and was significantly different from each of the other distributions. Error bars are standard error of the mean.

was also tested for the certainty given for predictions and Mauchly's test was significant, $\chi^2(5) = 92.36, p < .001$; degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .87$, and the resulting values showed a significant effect of distribution, $F(2.60, 984.70) = 32.64, p < .001$. For squared prediction error in pixels, again, sphericity was tested using Mauchly's test and found to be significant, $\chi^2(5) = 198.05, p < .001$, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity, $\varepsilon = .74$, which revealed a main effect of distribution, $F(2.21, 839.16) = 67.83, p < .001$.

Post-hoc analyses using Tukey's test revealed that, for the probability of cor-

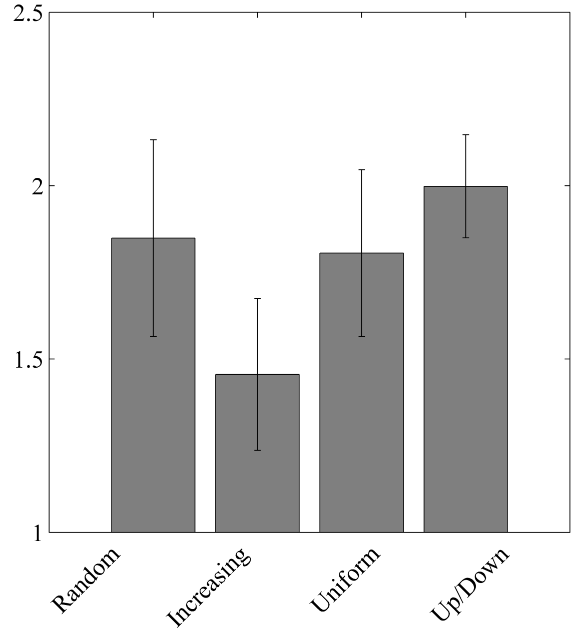


Figure 5.5: The mean Evaluation latent factor scores for each of the four distributions of size changes. Error bars are standard error of the mean.

rect prediction, all of the pairwise comparisons were significantly greater than the critical difference, this is shown in Table 5.2. As can be seen in the table, this was also the case for certainty of prediction and prediction error, with the exception of comparison of the errors for the probabilistically increasing and Up/down distributions.

These analyses suggest that the different distributions of change of size significantly effect the probability of predicting the correct outcome, confidence in the prediction and the accuracy of the prediction. The differences for each distribution are shown graphically in Figure 5.6.

Table 5.2: Results of Tukey's test for pairwise comparisons of distribution for each of probability of correct prediction, certainty of prediction and squared error of prediction.

Distribution comparison		Correct		Certainty		Error	
		Diff.	p	Diff.	p	Diff.	p
Random	Increasing	2.437	= .021	2.695	= .011	15.317	< .001
Random	Uniform	10.133	< .001	9.821	< .001	31.406	< .001
Random	Up/down	6.078	< .001	5.961	< .001	14.146	< .001
Increasing	Uniform	7.238	< .001	7.516	< .001	8.401	< .001
Increasing	Up/down	3.473	= .001	3.451	= .001	0.776	= .030
Uniform	Up/down	3.039	= .004	4.737	< .001	4.107	< .001

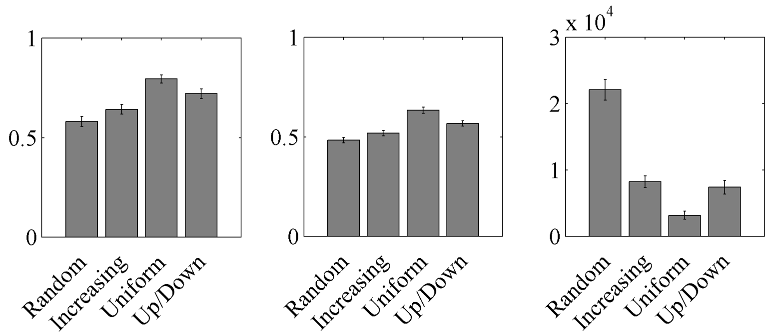


Figure 5.6: Measures of prediction accuracy for each distribution. Left panel: Probability of correctly predicting the change in size; Centre panel: How certain the prediction was; Right panel: The squared value of the difference (error) between the prediction and the actual size of the triangle. Error bars are standard error of the mean.

5.4 Discussion

The present experiment successfully attempted to both *a*) present a scenario to participants that was perhaps more like real life choices, to the extent that there was no explicit feedback and success in the task was due to experience alone; and *b*) to investigate the Activity/uncertainty dimension by varying the distribution of changes to the size of shapes while maintaining their colour and shape (only changing their orientation).

For the Activity/uncertainty dimension, this can be seen clearly and directly by comparing Figures 5.3 and 5.6 where the least active distribution in Figure 5.3 is the same as the distribution that is most certain in the centre panel of Figure 5.6; this distribution is the uniform distribution and by any judge is the least active and should be the easiest to see and so be certain of. Perhaps unsurprisingly, this is also supported by the numbers of correct predictions and squared prediction errors for the uniform distribution shown in the left and right panels of Figure 5.6 respectively.

It is also evident from a comparison of the two figures that the random distribution is rated significantly higher for the latent Potency/risk factor, that is to say it is more risky, than the other distributions. Again, this is supported by the objective measure of the lower number of correct predictions and the significantly higher squared prediction error for the random distribution, shown in the left and right hand panels of Figure 5.6 respectively.

The Evaluation dimension was not significantly different across the distributions, which was expected since there were no explicit rewards and other attributes of shape and colour did not change. Indeed the lower value for the probabilistically increasing distribution, which was approaching significance, was unexpected. Perhaps the reason for this difference for the probabilistically increasing distribution in evaluation is that the final two presentations for that distribution were reductions in size (as can be seen in Figure 5.2) where participants were more likely, based on their previous experience with the distribution, to expect a size increase. In retrospect it was probably an error for this distribution to finish with two reductions in size.

It may reasonably be asked whether the experimental task was too easy for the participants to recognise the distributions. It is, though, believed that this is not the case for several reasons. The experimental task is, in fact, quite difficult; remembering some representation of the last presented size for an orientation and learning the pattern of the distribution challenged all of the participants. While it seems clear that some of the participants must have recognised one or more of the patterns in the changes due to the different distributions, the majority claimed not to have recognised them (even the uniform distribution) in an informal discussion during the debrief. Also, if the distributions did not evoke some sort of emotional reaction or feeling, consistent with the Somatic marker hypothesis (Damasio, 1994), that could be differentiated, then a three factor semantic differential would presumably not result.

The results of the present experiment suggest that, unlike the Iowa gambling task (Bechara et al., 1994) and the AlphaBet shapes experiment described in Chapter 4, that although explicit rewards (or reward histories) may be sufficient to learn which are the good things to choose and which are bad, they are not a necessary requirement; our learning, in the present case making good predictions, appears to be sensitive to rather small and quite subtle differences. The histories associated with these predictions have been shown to affect the latent factors of a semantic differential, as hypothesised.

The present experiment adds to the evidence presented in the Alphabet shapes experiment that our reward structure is three dimensional and is represented by the semantic differential. Each of the dimensions of the semantic differential does seem to be capable of being relabelled reward, risk/danger and uncertainty respectively and this has been shown using an approach based on explicit rewards, mimicking, to some extent, an economic decision, and one based on subjective rewards and perhaps representative of a more everyday type of choice.

If, as hypothesised in the somatic marker hypothesis and described in Chapter 2, our choices are intimately related to feelings that come about through 'bodily' changes, then it is reasonable to consider that this will be found with choices that are made based on lower level perceptual information. The following Chapter investigates whether choices are influenced by perceptual information by relating reward structure (through the semantic differential) with scene gist through colour.

Chapter 6

Perceptual Information

6.1 Introduction

It is known that the mere perception of colour triggers evaluative processes (Elliot & Maier, 2007) and that perceptual information, in the form of colour, is included in the gist of scene. This chapter uses this idea to investigate the premise that the somatic marker hypothesis, and hence our reward structure, is created through physiological states and learned associations, and that our reward structure will be evident from this 'lower level' perceptual information. A number of images are rated and the semantic differential that is produced with factor analysis then used, with the probability of particular colours appearing in the image, to carry out regression analysis. Two different models of colour are used, the first based on the eleven basic colours proposed by Berlin and Kay (1969) and the second based on eleven colours found from a set of images representing 'the world'. The colour models were produced using a Gaussian mixture model, which makes it straight forward to calculate the probability of colours appearing in an image.

It is concluded that our reward structure is evident from low level perceptual information in the form of colour and in addition, a proposal for how the Berlin and Kay (1969) basic colours arise is made. First, however, the basic colour opponency theory is described.

One of the mainstays of vision research is the opponent process theory of colour, originally proposed by Hering in 1878 (Hering, 1964). Herings theory, largely arrived at through introspection, for example, that certain colour combinations, such as a bluey yellow or greeny red, cannot be imagined or produced, suggested that there are four primary colours and that they are detected by pairs of opponent processes, one for red/green and one for blue/yellow. However, compelling empirical evidence had to wait until Hurvich and Jameson (1957) and an experiment based on hue cancellation. Hue cancellation experiments start with a colour (e.g. yellow) and attempt to determine how much of the opponent colour (e.g. blue) must be added to eliminate any of that component from the starting colour, for example, Hurvich and Jameson (1957) asked their participants to move a control backwards and forwards until what they saw was neither yellowish or bluish. Nevertheless, questions do remain, not over an opponency process *per se*, but rather, whether there are exactly two pairs of opponent processes (e.g. Saunders & Brakel, 1997, p173).

Based on the opponency theory, a unique green is neither yellowish nor bluish and a unique blue is neither greenish nor reddish; effects that were considered by Berlin and Kay (1969) in their influential colour naming theory. The original Berlin and Kay (1969) analysis was based on a comparison of colour words in 20 languages from around the world that were used to choose exemplars from 320 Munsell colour chips. Most research into colour naming since Berlin and

Kay (1969) has investigated saliency using an approach based on the frequency of description used by participants for a range of colours that are presented to them (e.g. Hays, Margolis, Naroll, & Perkins, 1972; Boynton & Olson, 1990; Taft & Sivik, 1997). Boynton and Olson (1990), for example, presented participants with two trials of 424 colours, asking them to name the colours with a single term of their choice, and found that the basic colour terms have greater agreement between participants and are used more consistently within participants than any other terms. McManus (1983, 1997) on the other hand, used frequency of occurrence of colour terms in poetry and literature, finding that the colour terms and their order of evolution correlates very strongly with Berlin and Kay (1969).

However, despite the general acclaim for the theory, many detailed reviews of Berlin and Kay (1969) were critical of their methods of gathering and presenting data, few more so than Saunders and Brakel (1997, 2002) who are severely critical of many of the aspects of the original, as well as much of the subsequent, research. Saunders and Brakel (2002, pp335-336) argues in particular that all that has been accomplished is to confirm what was already known, because the system and methods used by Berlin and Kay (1969) and subsequent research, determine in advance how facts relate to each other, to participants and to the colour stimuli used and that the scientific evidence for the universality of the basic colour terms is flimsy at best. For example, Saunders and Brakel (2002, p336) argues the Munsell colour chips that were used as a basis to gather data were assumed to be exhaustive of colour, but all that was established was that participants could discriminate the most saturated, salient, Munsell chips in the set that was used. Whatever the eventual outcome of this debate, the Berlin and Kay (1969) basic colour terms are controversial; however, from the perspective of the present thesis, the original and subsequent research does not address how the colour that is gleaned from a scene may be utilised.

Gist normally refers to the substance or essence of something, for example, understanding the main point or essence of an argument is getting the gist of that argument, alternatively a statement such as *Anna didn't catch every word between them, but she heard enough to get the gist of the conversation* also illustrates the meaning. In vision research, gist is commonly used in a very similar way to describe the information that can be gleaned from a scene in a brief glance (e.g. Rousselet, Joubert, & Fabre-Thorpe, 2005) and also more broadly from a longer look, as it is here, to allow more complete descriptions of a scene (e.g. Fei-Fei, Iyer, Koch, & Perona, 2007; Potter, 1976). It has long been known that people are exceptionally good at getting the gist of a scene, especially in terms of scene classification and recognition (e.g. Fei-Fei et al., 2007; Potter, 1975, 1976; Rousselet et al., 2005); presumably, being able to act on the information gained from the gist of a scene is an especially useful evolutionary adaptation, both for prey and predator alike.

Colour vision evolved because it provided a means to gain better information about the world and to better recognise and detect things that contributed to survival (Cornelissen, Brenner, & Smeets, 2003), it therefore seems intuitively right that colour has more than aesthetic value; colour provides information and has specific meaning according to context and, consistent with a model by Elliot and Maier (2007), influences behaviour through learned associations that may be the cognitive expression of emotions and feelings. It has been found, for example, that avoidance behaviour is promoted by red (compared with blue for approach behaviour) (Mehta & Zhu, 2009), without conscious awareness. Interestingly, the same research also found that performance on a creative task was improved by blue, while red improved performance on detailed task.

These colour associations are learned from infancy and some may be developed from evolutionarily embedded dispositions, with the result that simple perception of a particular colour in a particular context will influence cognition and behaviour accordingly (Jacobs, 1981; Mollon, 1989); in this way colour acts as an unconscious prime. That is to say that the colour association acts on behaviour unintentionally and without conscious awareness. There is compelling evidence for just this sort of priming effect in the automaticity literature (e.g. Bargh & Chartrand, 1999; Bargh, Gollwitzer, Lee-chai, Barndollar, & Trotschel, 2001). But what of colour preferences?

Colour preferences are complicated and the literature that discusses them has been described as "bewildering, confused and contradictory" (McManus, Jones, & Cottrell, 1982). As an illustration of this complexity, it has been argued that red is particularly powerful as a colour because of associations with blood and also because of its unusualness, for example, in sunsets (e.g. Humphrey, 1983); Elliot and Niesta (2008) suggests that red has a positive effect and enhances men's attraction to women, but Elliot, Maier, Moller, Friedman, and Meinhardt (2007) suggests that prior exposure to red has a detrimental effect on performance (in an IQ test). In general though, blue is preferred over reds and yellows of equivalent hue and chroma (McManus et al., 1982), but, people tend to have preferences for high chroma colours and when they are presented with them, red is preferred to blue; however, it is difficult to reproduce high value/chroma blues and the equivalent reds/yellows often look brownish or washed out, even with digital media (I.C. McManus, personal communication, October 15, 2009).

On the basis that colour has the sort of priming effect on behaviour described above, it is hypothesised that the colour composition of images, depicting differing content and contexts, will influence the semantic differential (reward structure) for those images in a similar way to the influence that winning and losing had on the reward structure for the shapes in the AlphaBet experiment. The semantic differential has been successfully used with images previously, both for objective and non-objective art (non-objective art is art that has no recognisable subject) (e.g. Springbett, 1960; Hagtvedt, Hagtvedt, & Patrick, 2008) as well as for aesthetic subjects that are not art, such as coastal landscape (Kim & Kang, 2009).

Assuming that 'our world' consists of a limited set of digital images also provides the opportunity, using an approach based on a Gaussian mixture model, to test whether a reasonable approximation of the 11 basic colours identified by Berlin and Kay (1969) is produced. With the colour information contained in the mixture model and based on the idea that colour has meaning according to context, the influence of colour on the semantic differential ratings produced from short duration viewing of the images, will be apparent. Nevertheless, given that colours and combinations of colours will provide different information in different contexts, it should perhaps not be expected that there will be a conventional preference order to the extent that each colour contributes, though the extent of the contribution of each colour is expected to be different.

The remainder of this Chapter falls into four distinct areas, contained in the next two sections: *a*) establish a test set of digital images and carry out a rating experiment with that image set; *b*) calculate a semantic differential from the experimental ratings data; *c*) build a Gaussian mixture model from the Berlin and

Kay (1969) focal colours and analyse the model and semantic differential in order to investigate the contribution of colour to the ratings; and *d*) for comparison, carry out the same analysis but with a Gaussian mixture model built from samples of images depicting everyday scenes.

6.2 Method

6.2.1 Participants

Three participants volunteered to take part in the rating experiment and were recruited from amongst the post graduates from the School of Experimental Psychology. All participants were male and reported normal or corrected to normal vision and all of the participants provided their informed consent prior to commencement of the experiment.

6.2.2 Materials

A set of 235 images was collected from digital media, including the internet, that were judged to be representative exemplars of action, classical art, surreal art, nature and romance genres (an example of each of these genres is given in Appendix H, Figures H.1, H.2, H.3, H.4 and H.5 respectively).

The images were displayed one at a time on a web page written in PHP. Below each image nine scales, based on slider bars, were used to rate the image (see Table 6.1), these scales were the three identified as the most heavily loaded for each of Evaluation, Potency and Activity in a factor analysis carried out by Osgood and Suci (1955, p336) and were typical of reliable scales. Each slider was initialised to the middle of the scale and produced values between 0 and 100.

Table 6.1: The nine scales used for rating the images

Evaluation	Potency	Activity
Awful - Nice	Light - Heavy	Slow - Fast
Ugly - Beautiful	Weak - Strong	Passive - Active
Dirty - Clean	Small - Large	Dull - Sharp

A button was included at the bottom of the screen in order to allow participants to move to the next image. Examples of the rating page are shown in Appendix H, Figure H.6.

6.2.3 Design

This experiment used a repeated measures design. The independent variables were the images in the test set that were to be rated. The dependent variables were the rating scores provided by the participants following presentation of an experimental image, using the sliders. The images were displayed in a random order.

6.2.4 Procedure

Participants were asked to rate every image by using every slider bar. Each slider was initialised to the middle of the scale and the participants moved the slider towards one extreme or the other, as indicated by the noun, in order to record their rating. Once the participant had completed the ratings, the 'Next Image' button at the bottom of the screen could be pressed to display the next trial.

The experiment was self paced with no time pressure, however, participants were asked to make their ratings quickly and not to spend too long thinking about them, in other words, participants were asked to go more on 'gut' feel. Once all of the ratings were completed, participants were presented with a thank you page.

6.3 Results

6.3.1 Factor analysis

The data collected from participants for each of the rating scales were first checked for 'factorability' using the Kaiser-Meyer-Olkin measure of sampling adequacy (.73) and with Bartletts test of sphericity ($\chi^2(36)$=2821.63, p <.001). On the basis of these checks, factor analysis was carried out on the data.

Three principal components with eigenvalues greater than 1 were revealed, accounting for 74.23% of the variance in the ratings. Table 6.2 shows loadings of the component matrix following a varimax rotation.

Though the factors were in a slightly different order from previous semantic differentials, three factors, consistent with previous findings, were revealed by the analysis. Factor one related to rating scales indicating Evaluation, Factor 2 to rating scales indicating Activity and Factor 3 to rating scales indicating Potency. The factor score coefficients calculated for each scale are shown in Table 6.3.

The purpose of the analysis was to relate the probability (or proportion) of a colour appearing in a set of images to the dimensions of a semantic differential produced by rating a set of images. Therefore, having calculated the semantic differential, a way to calculate the colour proportions in the images was required

Table 6.2: Rotated Component Matrix showing factor loadings greater than .5. Extracted using principal component analysis, with varimax rotation which converged in five iterations.

Scale	Factor 1	Factor 2	Factor 3
Awful - Nice	.889		
Ugly - Beautiful	.866		
Dirty - Clean	.861		
Light - Heavy	-.777		
Weak - Strong			.796
Small - Large			.849
Slow - Fast		.888	
Passive - Active		.905	
Dull - Sharp		.579	

Table 6.3: Factor score coefficients.

Scale	Factor 1	Factor 2	Factor 3
Awful - Nice	.293	.045	-.002
Ugly - Beautiful	.293	.006	.129
Dirty - Clean	.284	-.042	.045
Light - Heavy	-.245	-.016	.130
Weak - Strong	-.071	-.051	.494
Small - Large	.066	-.130	.572
Slow - Fast	-.031	.490	-.169
Passive - Active	-.001	.480	-.097
Dull - Sharp	.090	.243	.189

and for this an approach based on a Gaussian mixture model was chosen.

6.3.2 Gaussian mixture model

A Gaussian mixture model is a probabilistic model that assumes that data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. Fitting the best mixture of Gaussians for a given dataset (as measured by the log likelihood) results in a probability distribution of classes that can be used to predict the probability (posterior) of new data points belonging to those classes. Fitting Gaussian mixture models is an example of an unsupervised learning method, however, the computing required for fitting a mixture of Gaussians is exponential for the number of latent Gaussian distributions, so approximate inference techniques are often used. While this does not guarantee *the* optimal solution, models do converge quickly to a 'local' optimum. To improve the quality it is usual to fit many of these models and choose the model that best fits the data, often on the basis of log likelihood or similar approach. Here the Gaussian mixture model functions from the Netlab toolbox (Nabney & Bishop, 2004) are used, these functions initialise the model using a clustering process known as *k-means* and then use the expectation maximisation (EM) algorithm. EM is an iterative method for obtaining maximum likelihood estimates of parameters for models that depend on unobserved variables, in the present case a finite number of Gaussian distributions.

For comparison, Gaussian mixture models were built from two sources: *a*) colour data that is freely available at the World Color Survey website (Cook, Kay, & Regier, 2011); and *b*) a reference set of 233 randomly chosen digital images of everyday scenes (with mean width and height of 1512 and 1127 pixels respectively; see Appendix H, Figure H.7 for examples). The mixture models were built using CIEL*a*b* colour space values. A description of CIEL*a*b* colour space and how the values were obtained is provided in the next section.

6.3.3 Mixture model (World Color Survey)

The World Color Survey began towards the end of the 1970's in order to test the hypotheses that were proposed by Berlin and Kay (1969) and was specifically involved with research into the existence of colour naming across languages and cultures, together with the evolution of colour terms in languages over time. From the World Color Survey data three files were used in the following way to obtain values for the eleven basic colour terms that Berlin and Kay (1969) proposed and which would provide the centres of the Gaussian distributions in the mixture model.

Using the World Color Survey data meant that the centres of the Gaussian distributions that formed the model could be fixed at the outset (although the mean values were calculated for those colours consisting of multiple points, see Figure 6.1) and therefore did not require repeated processing in order to arrive at the best model, rather, a procedure was needed to fix the priors and covariance matrix of the model.

The definitive names, abbreviations and order of emergence for the colour terms were obtained from BK-dict.txt. The abbreviations were then used to obtain the ranges of cells in the World Color Survey colour chart (shown in Figure 6.1) that represent the basic colours.

These cell ranges were then used to obtain the CIEL*a*b* colour values from the cnum-vhcm-lab-new.txt table. Where multiple values represented a basic colour term, it was a simple matter to calculate the mean values for that colour. The colour terms, their numbers and CIEL*a*b* values are shown in Table 6.4 and a representation of the colours in Table 6.5. CIEL*a*b* is a perceptually uniform

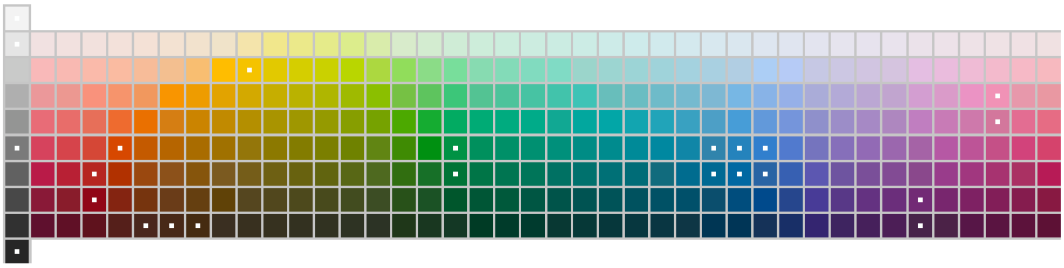


Figure 6.1: The World Color Survey colour chart with the centres representing the 'best' colours marked with white dots.

colour space, where perceptually uniform means that a change of a certain amount in a colour value should produce a similar change in visual importance. Unlike the RGB (used in most display devices) and CMYK colour models, the CIEL*a*b* colour space includes all perceivable colours and its gamut exceeds that of other colour models. It is clearly important to use a perceptually plausible colour model, however, the benefit may, to some extent, be obviated here because the images that were used were only available in RGB.

Table 6.4: The Berlin and Kay (1969) basic colour terms, their numbers and CIEL*a*b* values taken from the World Color Survey.

Term	Number	L*	a*	b*
Black	1	15.60	-0.02	0.02
White	2	93.54	-0.06	0.06
Red	3	36.00	54.64	37.53
Green	4	46.40	-58.43	26.12
Yellow	5	81.35	7.28	109.12
Blue	6	51.57	-10.34	-38.57
Brown	7	20.54	13.14	18.05
Purple	8	25.66	31.26	-22.32
Pink	9	66.65	40.06	-3.07
Orange	10	51.57	55.20	68.32
Grey	11	51.57	-0.03	0.04

Table 6.5: Representation of the Berlin and Kay (1969) basic colour terms using CIEL*a*b* colour values from the World Colour Survey, together with the colour terms and their numbers.

Term	Number	Colour(s)
Dark	1	
Light	2	
Red	3	
Green	4	
Yellow	5	
Blue	6	
Brown	7	
Purple	8	
Pink	9	
Orange	10	
Grey	11	

Once the necessary data (the CIEL*a*b* values) for the mixture model centres had been obtained, the mixture model was created based on the structure used in Netlab (Nabney & Bishop, 2004). In order to complete the model, a 1% sample of the pixels from every image in the reference set was taken. The samples were taken from each image as RGB values and converted to the CIEL*a*b* colour space, resulting in a sample of 4.04 million colour values; these values were used to calculate the prior probabilities for each of the centres together with their covariance matrix.

Using all of the colour values from every image in the rated test set of images, posterior probabilities for each colour (the probability that a colour appears in the image) were then calculated using the model that had been constructed. Mean posterior values for each image were calculated, which, along with the semantic differential discussed above, provided a data set that could be analysed using multiple regression.

6.3.4 Regression (World Color Survey)

To investigate the relationship between the World Color survey colours (i.e. the Berlin and Kay (1969) colours) and semantic differential obtained from the ratings of the set of test images, multiple regression was carried out using the mean values for each image, for each dimension of the semantic differential. The three multiple regressions used Evaluation/reward, Potency/risk and Activity/uncertainty in turn as the dependent variable, together with the mean values of the posterior probability of the colours for each of the images in the test set, as the independent variables.

Significant relationships were found for Evaluation/reward ($R = .49, R^2 = .24, aR^2 = .21, F(10, 224) = 7.232, p < .001$) and Potency/risk ($R = .32, R^2 = .10, aR^2 = .06, F(10, 224) = 2.570, p = .006$), while Activity/uncertainty was short of significance ($R = .25, R^2 = .06, aR^2 = .02, F(10, 224) = 1.520, p = n/s$). Colinearity diagnostics indicated that the tolerance for black was < 0.0001, so it was excluded from the analysis. Inspecting the beta coefficients for the regressions showed that the contribution of the colours was in the order shown in Table 6.6.

While this model provides a reference point for further comparisons, it is unrealistic, simply because the model centres were artificially set at the values given by Berlin and Kay (1969) and the World Color Survey. The model produced would be considerably better if the centres were learned, unaided, simply from the data.

Table 6.6: The standardised beta coefficients from multiple regressions using the semantic differential and the Gaussian mixture model posterior probabilities built with the World Color Survey colour centre values.

Evaluation (Reward)		Potency (Risk)		Activity (Uncertainty)	
Term	Beta	Term	Beta	Term	Beta
Grey	0.398	Brown	-0.226	Grey	0.174
White	0.201	Pink	0.173	Purple	0.121
Green	0.131	Yellow	0.117	Yellow	-0.080
Orange	0.120	White	0.087	Orange	0.080
Blue	0.079	Red	0.059	Green	-0.068
Red	0.074	Green	-0.044	Blue	0.052
Pink	0.058	Blue	0.041	White	0.038
Purple	0.048	Purple	0.024	Red	0.033
Brown	0.034	Grey	-0.024	Pink	0.009
Yellow	-0.007	Orange	-0.007	Brown	0.009

6.3.5 Establishing the number of centres

To establish the best estimate of the number of Gaussians (centres) that describe the colour data from the reference images, a large number of models were produced. For each of 100 relatively modest random samples of 100,000 pixels from the 4.04 million sampled from the reference set, mixture models were produced for 2 to 20 centres, each repeated 10 times. From the 10 repetitions for each of the models the best log likelihood was obtained and from these, the minimum description length calculated, as shown in Equation 6.1 (Nannen, 2003, p14).

$$L(D) = \min \left[-\log P(D|M_k) + k \log \sqrt{N} \right] \tag{6.1}$$

where D is the data set, M_k is the model, k is the number of Gaussians or centres and N is the number of points in the data set.

While obtaining the minimum description lengths, the number of centres responsible for that minimum description length was also obtained. From the 100 centre values obtained, the mean number of centres that were responsible for the minimum description lengths was calculated. The mean number of centres was 12.8 and the standard deviation was 4.96, this is illustrated in Figure 6.2. In view of these estimates 18 centres was set as the upper limit for the Gaussian mixture model used for the image data.

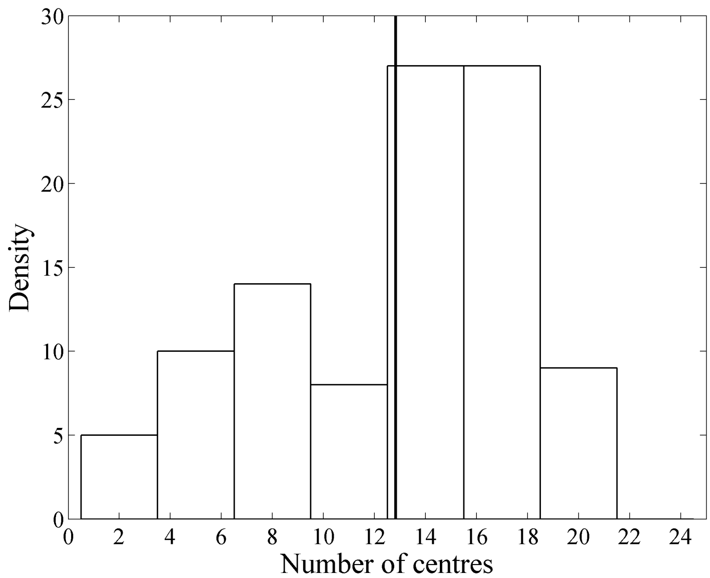


Figure 6.2: Histogram of the numbers of centres for each of the 100 minimum description lengths calculated for the mixture models and showing the mean (thick line).

6.3.6 Mixture model (image data)

Ten mixture models, each with 18 centres, were created using the 1% sample consisting of the 4.04 million colour values described above and the model best fitting the data was chosen using log likelihood. A graphical representation of the colours that were found by the mixture model is shown in Table 6.7.

It is interesting to note that, of the 18 centres that were found, four of the centres are clearly shades of blue and a further four centres were clearly shades of purple, also, two centres were clearly shades of brown. Not only does this bear striking similarity to the centres provided by the World Color Survey, but, if the multiple shades of blue, purple and brown are collapsed into a single colour value for blue, a single colour value for purple and a single colour value for brown, there are a total of eleven centres, exactly the number argued by Berlin and Kay (1969), the CIEL*a*b* values for these colours are shown in Table 6.8.

Table 6.7: Representation of the colour centres found by the Gaussian mixture model. The best matching Berlin and Kay (1969) basic colour terms and their numbers are included for reference.

Term	Number	Colour(s)
Dark	1	
Light	2	
Red	3	
Green	4	
Yellow	5	
Blue	6	
Brown	7	
Purple	8	
Pink	9	
Orange	10	
Grey	11	

Table 6.8: Colour centre CIEL*a*b* values generated by the Gaussian mixture model based on samples from the test image set, together with the matching Berlin and Kay (1969) colour terms and numbers for reference.

Term	Number	L*	a*	b*
Dark	1	15.60	-0.02	0.02
Light	2	96.00	-0.06	0.06
Red	3	59.17	59.36	19.67
Green	4	41.23	-33.18	41.15
Yellow	5	88.65	-33.80	12.41
Blue	6	52.10	-2.17	-35.23
Brown	7	36.06	12.34	36.94
Purple	8	48.98	29.49	-34.06
Pink	9	78.91	38.97	-10.98
Orange	10	76.48	38.22	29.22
Grey	11	56.64	-0.03	0.04

6.3.7 Regression (image data)

To investigate the relationship between the colours found by the mixture model and the semantic differential obtained from the ratings of the set of test images, multiple regression was carried out using the mean values for each image, for each dimension of the semantic differential. As with the regressions described above for the initial mixture model, the three multiple regressions used Evaluation/reward, Potency/risk and Activity/uncertainty in turn as the dependent variable, together with the mean values of the posterior probability of the colours for each of the images in the test set, as the independent variables.

Significant relationships were found for Evaluation/reward ($R = .47, R^2 = .22, aR^2 = .19, F(10, 224) = 6.403, p < .001$), Potency/risk ($R = .32, R^2 = .10, aR^2 = .06, F(10, 224) = 2.520, p = .007$), but, as with the previous model, Activity/uncertainty was found to be short of significance ($R = .23, R^2 = .05, aR^2 =$

$.01, F(10, 224) = 1.268, p = n/s)$. Colinearity diagnostics indicated that the tolerance for grey was < 0.0001, so it was excluded from the analysis. Inspecting the beta coefficients for the regressions showed that the contribution of the colours was in the order shown in Table 6.9.

Table 6.9: The standardised beta coefficients from multiple regressions using the semantic differential and the Gaussian mixture model posterior probabilities built with the image colour data centre values.

Evaluation (Reward)		Potency (Risk)		Activity (Uncertainty)	
Term	Beta	Term	Beta	Term	Beta
Black	-0.380	Yellow	0.159	Black	-0.137
Brown	-0.141	White	0.153	Yellow	-0.132
Orange	0.120	Orange	0.142	White	0.125
White	0.082	Red	0.126	Brown	0.037
Yellow	0.076	Brown	-0.107	Green	-0.028
Purple	-0.076	Black	0.106	Orange	0.017
Blue	0.055	Blue	0.104	Red	0.016
Pink	-0.033	Purple	-0.088	Blue	0.014
Red	0.031	Pink	-0.064	Purple	0.011
Green	0.027	Green	-0.043	Pink	-0.010

6.4 Discussion

This chapter addresses the idea that our reward structure, as represented by the semantic differential and as proposed in Chapter 3, will be available from perceptual information. Perceptual information in the form of colour, gained from the gist of a set of visual scenes, was expected to be related to how those scenes were rated, since it is known that the mere perception of colour triggers evaluative processes, and this is what was found. Initially, the rating experiment, based on ratings of exemplars of action, classical art, surreal art, nature and romance genres showed that a semantic differential was formed with the expected dimensions of Evaluation, Potency and Activity. However, what was of particular interest in the present experiment was how the semantic differential or reward structure related to the colour composition of the images.

In order to establish the colour composition, a Gaussian mixture model was created from mean colour values from the World Color Survey that represented the Berlin and Kay (1969) basic colour terms and samples from a set of images that provided a reference set of everyday scenes (which were different from the rating test set); multiple regression was then used to investigate the relationship between the model and the dimensions of the reward structure. The World Color Survey data has previously been analysed using clustering techniques (e.g. Lindsey & Brown, 2006) where the relationships between colour and language are argued for, however, an approach to these colour categories in terms of choice or reward has not been made. A significant relationship was found for Evaluation/reward and Potency/risk dimensions and the colour composition of the set of scenes used for testing.

However, in the present case, a mixture model built from the focal colour values contained in the World Color Survey, while successful, is unsatisfactory just because the values for the Gaussian centres (the focal colours) come from an external source and can not be said to have been learned by the model (although the priors and covariances clearly were). This approach can be criticised in the same way that Saunders and Brakel (1997, 2002) criticised Berlin and Kay (1969) as merely confirming what is already known by assuming that the basic colours are correct. More satisfactory is to use focal colour values that were learned by the model and this was what was carried out with a second Gaussian mixture model. First though, an upper limit for the number of centres for the model was needed.

Through running a Gaussian mixture model many times with a modest set of test data and with a number of centres from two to twenty, the log likelihood was used to calculate a minimum description length for each model and the number of centres best fitting the data established. The mean number of best fitting centres was calculated and an upper limit on the number of centres to use in the model set at 18. Running the model with a much larger sample from the reference set of scenes resulted in the mean colour centres shown in Table 6.11. Several of the centres were clearly shades of basic colours i.e. Blue, Purple and Brown, a similar outcome to the World Color Survey for several colours e.g. Blue, Purple, Brown, Red etc. Calculating the means for these multiple colour centres resulted in eleven colours that are plausible facsimiles of the basic colour terms defined by the World Color Survey and Berlin and Kay (1969). As with the initial mixture model, multiple regression was used in order to investigate the relationship between the model and the dimensions of the reward structure obtained from image ratings. Again, the regression analysis showed significant relationships with Evaluation/reward and Potency/risk dimensions of the reward structure, similar to the initial model.

Inspecting the beta coefficients for the regressions showed a different order in the contributions of each colour between the models. The reasons for this are unclear and are perhaps to do with the effect of chroma on preferences, which was highlighted by McManus (I.C. McManus, personal communication, October 15, 2009) and discussed above, nonetheless, the present analysis does not extend far enough to be able to draw any clear conclusions in this area. However, the model colours can be considered in a different way.

Using the World Color Survey colour chart and setting the central colour value in every cell that was named as that colour, provides a colour chart that illustrates the composition of each colour in terms of the chart; this is illustrated in Figure 6.3. Presumably, if the present model is a good representation of the basic colour terms, obtaining posterior probabilities for the World Color Survey colour chart colours using the model should provide a reasonable approximation of the centre version of the World Color Survey chart (Figure 6.3).

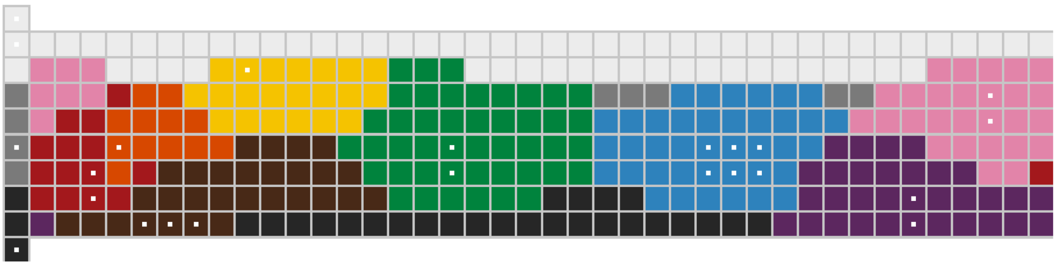


Figure 6.3: The World Color Survey colour chart with the central colour value set in every cell that was named as that colour, together with the cells representing the 'best' colours marked with white dots.

Rather than the 330 colours that form the World Color Survey colour chart as shown in Figure 6.1, the full range of colours that the World Color Survey used was 998 and these were used to produce the colour chart shown in Figure 6.4. While this colour chart appears to contain a representation of the basic colours, it could not be said to show ideal or even especially definite shades of those colours. However, neither the World Color Survey nor Berlin and Kay (1969) were interested in what people thought the mean or central value for a colour was, but rather what people considered the typical or best representation of a colour from a constrained set of colours that was presented to them. The present model should therefore be able to achieve a similar outcome, using the same data. It is argued that this is what happens in a phenomenon known as peak shift.

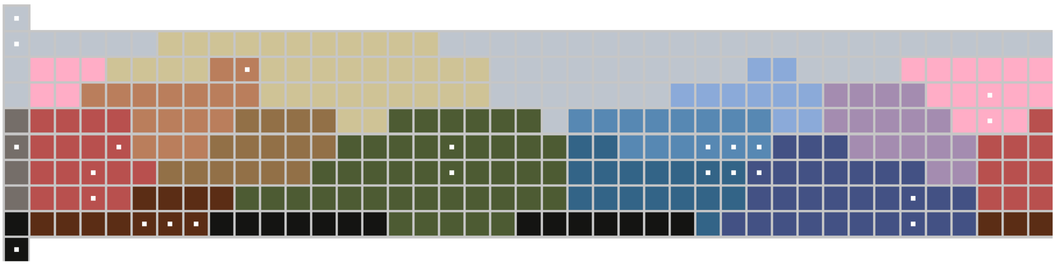


Figure 6.4: Colour chart based on mixture model centre values with the central colour value set in every cell that was closest to that colour centre, together with the cells representing the 'best' colours marked with white dots.

Peak shift is a well known psychological phenomenon that is found in animal discrimination learning and in humans (e.g. Thomas, Mood, Morrison, & Wiertelak, 1991; Wills & Mackintosh, 1998; Ramachandran & Hirstein, 1999; Baddeley, Osorio, & Jones, 2007). In the peak shift effect, animals respond more strongly to exaggerated versions of the training stimuli, as in the often quoted example of rats discriminating between squares and rectangles (e.g. Ramachandran & Hirstein, 1999). When rats are rewarded for discriminating a square from a rectangle they rapidly learn to choose the rectangle, interestingly though, response to an extreme, longer and thinner, rectangle provokes a significantly stronger response than the original training rectangle. Ramachandran and Hirstein (1999) argues that rats are learning a rule (of 'rectanglularity') rather than a prototype and that similar, peak shift, phenomena also occur in people where they apply to many visual experiences, especially to art. Instances of peak shift in art seem to be too common to be ignored, for example, in representations of the human figure by Michelangelo, Rubens, and Renoir, and the distortions of El Greco, Giacometti and Modigliani, although, perhaps the extremes of colour used by the Impressionists are more relevant to the present chapter. How then does the peak shift effect apply to, say, our choice of the reddest of reds?

It seems clear that what most people would choose as the 'ideal' or most representative red does not appear often in the natural world and even if it did it is hard to understand why a particular shade out of all of the shades would be chosen as constitutive as ideal redness. To paraphrase an example in Ramachandran and Hirstein (1999), a skilled caricaturist will subconsciously subtract average features, while exaggerating the differences, and thereby create a cartoon that is more like the original than the original (for example, imagine a Gerald Scarfe cartoon of Margret Thatcher or Tony Blair), it is argued that the rat responds to 'rectangularness', and the perfect red is identified, in the same way.

In order to establish a peak shift value for each colour centre (or in the present case the colours in the World Color Survey colour chart), first, the mixture model is used to calculate the posterior probability that each sampled colour point, used to create the model, belongs to each centre. The Netlab toolbox (Nabney & Bishop, 2004) contains a routine to allow this to be achieved easily (gmmpost). The Gaussian centre that each colour belongs to is then established, that is, for every sampled colour the nearest centre is established by calculating the Euclidian distance from the sampled colour to each centre and choosing the centre that is the minimum distance from it. Having established the points belonging to each centre the posterior probabilities are then used and the point with the maximum posterior value identified, this point is the peak shifted value that is used for the centre in question. Carrying out this procedure for the learned (image data) model results in the peak shifted colours shown in Table 6.10 and having the CIEL*a*b* values shown in Table 6.11.

Table 6.10: Peak shifted colour representation.

Term	Number	Colour(s)
Dark	1	
Light	2	
Red	3	
Green	4	
Yellow	5	
Blue	6	
Brown	7	
Purple	8	
Pink	9	
Orange	10	
Grey	11	

The peak shift effect for the centre version of the colour chart is shown in Figure 6.5. As can be seen, the representations for each of the colours is much more as expected for representative colours, with, for example, redder reds, and greener greens than the mean values shown in Figure 6.4. Indeed, given the limited 'world' that the mixture model has learned the colours from, the peak shifted colours are a remarkably good representation of the 11 basic colours hypothesised by Berlin and Kay (1969). The peak shifted version of the focal colours found by the Gaussian mixture model is indicating that what might be considered a

Table 6.11: Colour class CIEL*a*b* values generated by the Gaussian mixture model based on samples from the test image set that have been peak shifted. The best matching Berlin and Kay (1969) basic colour terms and their numbers are included for reference.

Term	Number	L*	a*	b*
Dark	1	15.600	-0.020	0.020
Light	2	96.000	-0.060	0.060
Red	3	51.570	58.010	30.520
Green	4	41.220	-19.480	36.560
Yellow	5	91.080	-5.250	45.240
Blue	6	53.702	-2.497	-29.760
Brown	7	39.876	16.290	38.240
Purple	8	61.700	20.070	-26.920
Pink	9	76.475	40.040	-2.700
Orange	10	71.600	34.960	37.540
Grey	11	56.635	-0.030	0.040

perfect, or ideal, representation of a colour is not the typical or average value that is typically seen or is most common; rather it is an extreme value. Does this provide support for Berlin and Kay (1969)?

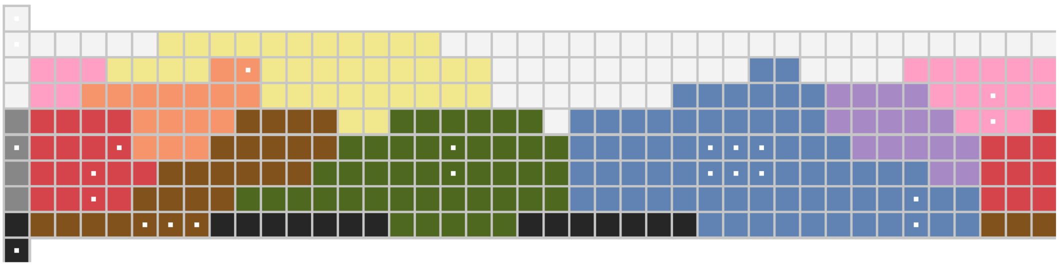


Figure 6.5: Colour chart based on peak shifted mixture model centre values with the central colour value set in every cell that was closest to that colour centre, together with the cells representing the 'best' colours marked with white dots.

The present experiment does not offer any insight into colour terms with respect to language. However, what it does provide is an objective method to assess the most effective number of colours that can be profitably used to describe the world. It can also be claimed that the criticisms of Saunders and Brakel (1997, 2002) do not apply to the present approach, at least insofar as particular ranges of colour are not assumed (except that all colours came from digital images). The number of colours was found to be around 11 and to the extent that this was the number of colour terms found by previous research, adds support to the Berlin and Kay (1969) hypothesis.

Despite this, the modelling approach used here could be used to extend the investigation of basic colour terms, particularly in specific environments or carrying out specific tasks. Taken together, the semantic differential, the Gaussian mixture model and the relationships that were revealed between them, suggest that perceptual colour information from gist, gained during short duration viewing of images representing a range of genres does evoke evaluative processes. These processes are the basic mechanisms that establish whether a stimulus is hostile or hospitable (Elliot & Covington, 2001), which in turn produce motivated behaviour such as choices.

The experiments that have been reported in Chapters 3, 4, 5 and 6 have argued that a single dimension of reward or utility is insufficient to represent the prior information that must be considered when making inductive inferences that lead to choices and that a richer three dimensional representation is needed. A large data source was used to investigate the plausibility of this idea and then, in subsequent experiments whether the hypothesised representation can be directly affected, or learned. This was investigated on the basis of a choice with rewards in the AlphaBet experiment and a choice based on a more subjective inference in the triangles experiment, with the present experiment investigating if this reward structure is apparent based on perceptual information. The chapters that follow provide a mathematical model, based on the ideas that have been investigated, consisting of a value function and probability weighting function, beginning with a brief discussion that reiterates the importance of uncertainty, introduces Bayesian modelling and discusses a feature of the decision making literature that has received attention recently; decisions from description versus decisions from experience.

Chapter 7

Choices

7.1 Introduction

Problems involving choices amongst such things as simple lotteries and gambles have been intensively studied over the last forty years. Research involving these sorts of choices is important because it represents in an abstract way how people deal with risky and uncertain decisions. As Neumann and Morgenstern (1944, §2.1.1, p. 8) observe, understanding these choices is also important because it is the decision making of individuals that constitutes the community. It is obvious that choice behaviour has implications in many economic contexts, from why we gamble and why we buy insurance, through to international relations. This sort of research is also of further importance because peoples behaviour is not well described by expected utility theory, which is taken as the normative theory of choice, particularly for economics.

Expected utility theory tells us that, for each of a set of alternatives that we want to choose between, we ought to assign a real number or value to an outcome, then multiply that value by its probability, add them up and then use the best result to choose an alternative. This is the expectation principle (Kahneman, 2011) and has been taken to be the essence of rational behaviour, however, it depends on having complete information; recall from Chapter 2 that the classical model of rationality requires knowledge of all the relevant alternatives, their consequences and probabilities, and a predictable world without surprises (Simon, 1979, p500). Indeed, Neumann and Morgenstern (1944) and Marschak (1950) explicitly consider only the situation where there is complete information, neither considers what rational behaviour might be under incomplete information, with all of the vagueness of uncertainty and risk (in the sense of danger or difficulty) that it might entail for the individual in a social economy, attempting to make a choice. Recall also from Chapter 2 the often quoted view associated with Knight (1921), that an uncertain alternative is one where some of the outcomes may occur with a probability that is known, others may occur with a probability that is unknown and others may be unknown.

This chapter briefly introduces the following three chapters, which consider how the empirical findings discussed in the preceding chapters might be pulled together into a coherent mathematical model. The research that led to the questioning of expected utility theory and production of prospect theory are briefly recapped and an outline of the approach taken here provided. One of the most significant issues is one of uncertainty, which is alluded to above.

7.2 Uncertainty

The main condition for expected utility theory to be rational is that the stated probability for an alternative has no uncertainty associated with it, that is, there is complete information for that alternative. If a probability of 50% has the same information content as the statement *we will observe* 5×10^{10} *successes out of* 1×10^{11} *trials*, then no amount of previous knowledge should affect it. This is at least plausible in artificial situations, for example, those involving randomly drawn coloured balls from urns, however, for the kinds of situations we are evolved to deal with there are four reasons that such extreme levels of certainty never occur.

The first reason is that a probability statement made by a signaller (or otherwise available to a person) is based on finite levels of experience. Not only does such knowledge have uncertainty associated with it, but the 'precision' of an estimate only improves very slowly with additional measurements (as the square root of the number of measurements for independent trials). (Jacob) Bernoulli (1713), after inventing much of what we understand of probability theory and promising many practical applications, is said to have given up when he calculated that the number of draws, from an urn containing twenty black and thirty white pebbles, that is required to reach 'moral certainty' that the ratio of black and white pebbles in the urn is 2:3, was larger than the population of Basel, Switzerland (Mlodinow, 2008; Polasek, 2000). Most probability signaller's experience of situations, indeed the experience of most people, is not so extensive. Even given extensive experience, the second problem of non-stationarity fundamentally limits certainty.

The fact that the world changes over time limits how confident we will or can be about future events. Statements about probability made in a changing world must be uncertain since, unsurprisingly, the world may have changed. Although finite, limited, experience and non-stationarity are the most fundamental sources of uncertainty they must also potentially be combined with problems associated with, the third problem, (inaccurate) memory and even, the fourth problem, the possible dishonesty of the person making the probability statement (or, to state matters more positively, the receiver has perfect trust in the communicator). Together these reasons mean that, except in rather artificial situations, real life probability statements rarely contain complete information and are never certain.

Given probabilistic statements with potentially incomplete or vague information, associated uncertainty and risk, the rational way of dealing with them is to combine them, using Bayes' rule, with prior experience. Indeed, rational models, often based on Bayesian inference (Oaksford & Chater, 2007; Griffiths, Kemp, & Tenenbaum, 2008), provide very good accounts of many other aspects of cognition.

7.3 Bayesian modelling

A Bayesian approach to a problem starts by defining three aspects of a given problem, beginning with the formulation of a model, that is, the representation of the problem needs to be defined together with the relevant quantities of interest, the hypothesis space (Griffiths et al., 2008). Then, the relationship between observations and the model (likelihood) can be established; and then all relevant knowledge of the quantities of interest (priors), that capture our beliefs, can also be established.

From these three definitions all other quantities are calculated using standard mathematical identities. Given the complexity of many Bayesian calculations some approximations often need to be specified, however for the problems described in the following chapters all of the relevant distributions can be calculated analytically. The result of many Bayesian calculations is a probability distribution and to ease interpretation of this distribution, some summary statistics are often required.

In the following chapters a Bayesian approach is adopted for modelling a value function and a probability weighting function. Accepting Kahneman and Tversky's empirical findings and that the two functions described by prospect theory are required in order to describe choice behaviour, several other factors must also be considered as the basis of an expected utility theory that matches human performance.

7.4 An Expected Utility Theory That Matches Human Performance.

As has been previously described, maximising expected utility has long been accepted as a valid, and very often the preferred, model of rational behaviour, however, it has limited predictive and descriptive accuracy simply because, in practice, people do not always behave in the prescribed way. Accordingly, this is interpreted as evidence that either humans are not rational; expected utility is not an appropriate characterisation of rationality; or, some combination of both of these. The following chapters argue that this observed behaviour can be considered rational and expected utility an appropriate characterisation of rationality, provided it is accepted that:

1. most utility/reward has no meaning unless it is in the presence of potential competitors;
2. there is uncertainty in the nature of the competitors;
3. all statements of probability are also associated with uncertainty;
4. utility is marginalised over uncertainty, with framing effects providing constraints; and
5. utility is also sensitive to risk (that is, the potential for punishment or danger), which when taken together with reward and uncertainty suggests a three dimensional representation of utility.

The model produced is tested against three groups of decision making paradoxes (preference reversals, dominance effects and context effects (Busemeyer & Johnson, 2008)) and found to be predictively accurate. While these so called decision making paradoxes are confined to what have become known as 'description based decisions' there is a distinction in the literature between these and 'experience based decisions' (e.g. Barron & Erev, 2003; Hertwig, Barron, Weber, & Erev, 2004; Newell & Rakow, 2007).

7.5 Decisions from experience

Description based decisions, such as the following prospect:

1. Would you prefer option A or option B?

   A £3 for certain.

   B £4 with probability 80% otherwise nothing.

2. Would you prefer option C or option D?

   C £300 with probability 25%.

   D £400 with probability 20%.

contrast the answers to two questions posed to the same participants. These prospects are based on explicit information about alternatives, particularly gains or losses and their probability of occurring. Decisions from experience, on the other hand, are based on information gained from the outcome of previous choices, be they good, bad or indifferent. Decision makers presumably learn about the outcomes associated with each alternative through repeated sampling. Interestingly however, the finding, for description based decisions, that low probability events are overweighted has been shown to be reversed for experience based decisions. For example, Hertwig et al. (2004) found that people make choices as if they underweight the probability of rare events and explored reliance on relatively small samples of information and overweighting of recently sampled information.

These findings though have been attributed to sampling errors (Fox & Hadar, 2006), however, according to other researchers this explanation may not be quite so straight forward. Ungemach, Chater, and Stewart (2009) carried out further research with experience based decisions that sought to eliminate the under sampling that was thought to potentially cause the difference in weightings between description based decisions and experience based decisions. The results showed that the extent of underweighting of small probabilities and overweighting of higher probabilities was reduced for experience based decisions but apparently not eliminated or reversed, suggesting that under sampling is not responsible for the entirety of the effect.

Alternatively, L. Hadar and Fox (2009) proposes that the cause of, what has become known as, the 'experience gap' is a combination of sampling error and the amount of information available in each of the paradigms (i.e. description based decisions compared with experience based decisions) and testing what they refer to as 'information asymmetry. They find that accounting for under sampling and information asymmetry drastically reduces the extent of overestimation for small probabilities with experience based decisions compared with description based decisions, but does not reverse it. Hadar and Fox (2009) conclude that there is no need for an alternative theory for decisions from experience. Decisions from experience, as well as the other areas discussed above are considered in the following chapters, which propose a mathematical model beginning with the value function.

Chapter 8

The value function

8.1 Introduction

Historically, choosing on a rational basis meant selecting an alternative with the greater or greatest expected value, E_V (Eq. 8.1). It is commonly accepted that Pascal's solution to a gambling problem, known as the problem of points or division of the stakes (Pascal, 1665), shown in Appendix G, is the first example of using expected values, which in turn was the basis of Huygens (1657) *De ratiociniis in ludo aleae*, itself the first systematic work on probability.

$$E_V = \sum_{i=1}^{n} p_i x_i \tag{8.1}$$

While basing rational decisions on expected value is an appealing and elegant solution it was clear that, particularly highlighted by another gambling problem first stated by Nicolas Bernoulli in a letter to Pierre Montmort (1713), it was not adequate. The St Petersberg paradox, or prospect, as it became known (shown in Appendix B), is a gambling problem that is intended to illustrate that a rational person would be willing to part with only modest amounts in order to play in a game that is considered to have infinite expected value.

The St Petersberg problem was eventually resolved by Nicholas Bernoulli's cousin Daniel (1738) by suggesting that the utility of money increases non-linearly at a decreasing rate as the absolute value increases:

> ...the determination of the *value* of an item must not be based on its *price*, but rather on the *utility* it yields. The price of the item is dependent only on the thing itself and is equal for everyone; the utility, however, is dependent on the particular circumstances of the person making the estimate. (D. Bernoulli, 1954, §3 p. 24)

In other words, the objective value of the monetary gain of the gamble was replaced by the subjective value, or the utility, of the gain. Expected utility is defined as follows:

$$E_U = \sum_{i=1}^{n} p_i U(x_i) \tag{8.2}$$

where $U(x_i)$ is a positive but reducing function of the monetary amount x_i:

$$dU = c\frac{dx}{x} \tag{8.3}$$

where c is a constant, dU is change in utility, x is a monetary amount and dx is the change in the amount, forming a logarithmic utility function.

Fast forwarding to the twentieth century, game theory attempts to extend the study of rational decision making to situations that involve other agents. In the Theory of Games and Economic Behavior (Neumann & Morgenstern, 1944) the authors, as well as axiomatising expected utility theory (see Appendix A), effectively returned to using the objective value of money, for simplicity (Neumann

& Morgenstern, 1944, p. 8): For vonNeumann and Morgenstern, those involved in games, decisions and the economy more generally, want to maximise monetary income rather than a Bernoulli type utility. Nevertheless the quantity was labelled as utility and expected utility theory in this guise rapidly became the most influential theory of choice behaviour.

However, experiments quickly showed systematic violations of expected utility theory's axioms and its standing, particularly as a descriptive theory of choice, was questioned and alternatives proposed. Although there are a number of alternatives to expected utility theory they retain the central idea of D. Bernoulli (1738) in multiplying the probability of an outcome by its value or utility. Probably the most influential of the modified theories is Kahneman and Tversky's Prospect Theory (Kahneman & Tversky, 1979), which proposes two functions, where the value (or utility) of an outcome is multiplied by a subjective decision weight.

The modification to the value function in Prospect Theory is to replace the absolute utility function with a non-linear value function that is expressed relative to an individual's current position, convex (risk averse) for gains, concave (risk seeking) for losses and the curve for losses is steeper than for gains; in other words, things that make your world better are good, and things that make your world worse are bad. The individual's current position or reference point, which divides the area between losses and gains will be non-zero, dynamically changing over time and between decisions. Prospect theory suggests that because they are risk averse above the reference point and risk seeking below it people are loss averse; Figure 8.1 shows a hypothetical value function.

The reference point and shape of the value function are clearly important, indeed, Kahneman (2011, Part IV, Chapter 25 & 26) cites inclusion of a reference point, utility (value) being based on changes in wealth (though, of course, this is reminiscent of (D. Bernoulli, 1738)), and losses being treated differently from gains as the significant contributions that prospect theory made; Kahneman (2011) also expresses considerable surprise that what, on reflection, appear to be rather obvious anomalies in expected utility theory, had been overlooked for so long.

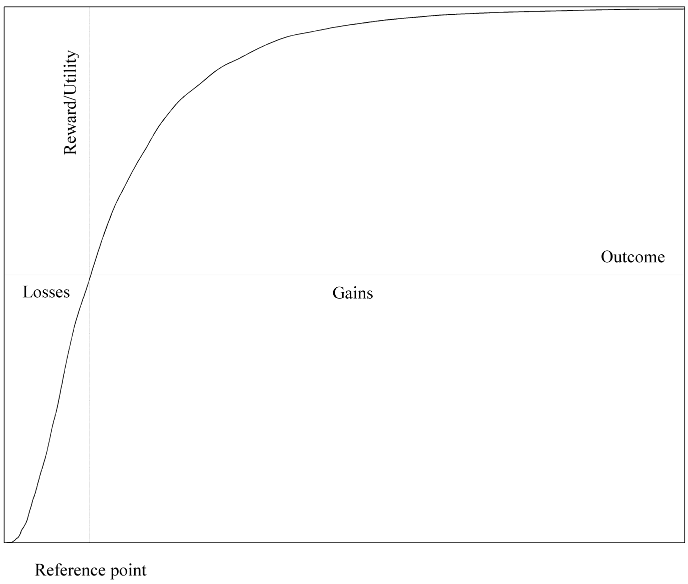


Figure 8.1: A hypothetical value function.

Accepting Kahneman and Tversky's empirical findings and that the two functions of Prospect Theory are required in order to describe choice behaviour, the following sections present a series of models of the value function, each offering a solution to a problem (or problems) raised by the previous one. The value function is cast in terms of a competition where an independent observer would judge the decision maker to have won a competition against one or more competitors:

This would be straight forward if the nature of the perceptual system of the judge and the competitors were known, however, in a real situation this can only be uncertain at best.

Proposing an independent judge may seem strange, but it harks all the way back to a time before Pascal (1665) and the expected value hypothesis, arguably to the heart of what expected value and, more importantly here, expected utility means. Teira (2006) argues that Nicholas and Daniel Bernoulli did not consider the St Petersburg problem as just a probabilistic puzzle but rather as a contradiction between expected value and 'common sense' and that the key to Daniel Bernoulli's solution (D. Bernoulli, 1738) was transferring the idea of expectation from a legal frame to an economic frame. Teira's (2006) argument is that ideas about expectation can be traced back to the work of Spanish Dominican, Domingo de Soto (1495–1560) and analogies that were drawn between a gambling game, which is a kind of contract voluntarily arranged by the gamblers and an agreement to insure seaborne commodities. The details are not of concern except to the extent that gambling games as contracts include a juridical standard of fairness, or just prices, which can arguably be seen in Bernoulli's solution (e.g. D. Bernoulli, 1954, p. 24).

8.2 Perfect Judge

Considered as a competition, a representation of both 'my' current position and winning potential together with the inferred position and winning potential of the opposition is required (and then combined) in order to produce the value function. For the simplest model, the value function consists of a single known competitor and a judge with a perfect memory and perfect perceptual system. In this case making a judgement is simple and resembles a step function as illustrated in Figure 8.2.

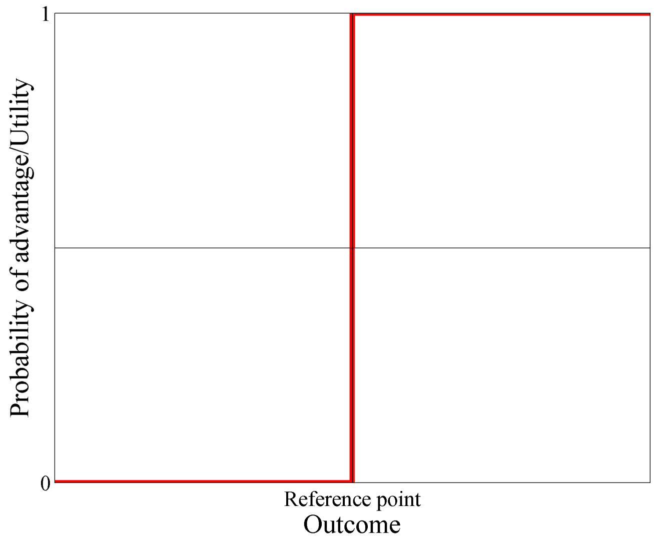


Figure 8.2: A simple step function representing a perfect judge where the slightest change in gain from below to above a reference point is represented by a snap from no probability of advantage to certainty of advantage.

To the left of the reference point there is a loss in a competition between 'me' and the competitor and to the right of the reference point there is a win against the competitor. The significance of this is that the perfect judge will be sensitive to the smallest change and snap from a certain probability of losing advantage when the outcome is below the reference point, to a certain probability of gaining advantage, when the outcome is above the reference point. However, this situation, as well as being the simplest, is the most unrealistic and changes are required in order to account for uncertainty.

8.3 Imperfect Judge

It is unrealistic to suppose that the judge will be perfectly sensitive, or error free, when making choices based on infinitesimal changes to gains. In other words, the perceptual system will be noisy and this must be represented in the value function. In addition the reference point must be considered. The the reference point, or current position, for 'me' is straight forward; it can be arbitrarily set at some value and the winning potential calculated. However, the opposition is more difficult because their current position is not 'known', that is, their reference point is uncertain and must be inferred. Fortunately, these values can be estimated on the basis of a psychometric function.

Psychometric functions describe the relationship between a physical stimulus and the responses of a person who has to make a judgement about the stimulus, for example, whether the difference in weight between two stimuli can be perceived; this resembles a sigmoid with the probability of a correct judgement on the y-axis and the stimulus magnitude on the x-axis; in general it can be calculated as follows:

$$p = \frac{1}{1 + \exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)} \tag{8.4}$$

where p is the probability that a change will be judged to have been made, $x \in \mathbb{R}^+$, μ is the mean and σ is the standard deviation (providing the slope for the function or its sensitivity). The smallest detectable difference (sensitivity) between two stimuli that a person can detect is referred to as a just noticeable difference (or jnd). In many sensory modalities the jnd is a constant proportion of the initial stimulus with the ratio of jnd to initial stimulus being approximately constant:

$$\frac{\Delta I}{I} = k \tag{8.5}$$

where I is the initial stimulus intensity, ΔI is the change in the initial stimulus intensity in order for it to be detected i.e. the jnd, and k is a constant. An alternative representation is:

$$\Delta S = k\frac{\Delta I}{I} \tag{8.6}$$

where ΔS is change in sensation; Note here the similarity with Bernoulli's (1738) relation between wealth and utility, above. This is known as Weber's Law or the Fechner-Weber Law and k as the Weber constant or fraction (Ekman, 1959).

The Weber constant provides the slope for the sigmoidal psychometric function. This logarithmic relationship (which can also be represented as $p = k \log\left(\frac{I_1}{I_0}\right)$ where I_0 is the initial stimulus intensity and I_1 a further stimulus intensity) approximately holds in many modalities and it is assumed that it also approximately holds for the value function. In order to introduce perceptual noise into the model, realistic values for the Weber fraction were sought and those from Teghtsoonian (1971) were used. Table 8.1 shows these typical values for the Weber fraction for various modalities and, since a Weber fraction would be difficult to establish across a wide range of choices, the overall mean of these values was used as an approximation for the model.

It is straight forward to introduce the perceptual noise into the model and it can be achieved as follows:

$$p = \frac{1}{1 + \exp\left(-\left(\left(\frac{x-o}{k}\right) + C\right)\right)} \tag{8.7}$$

where o represents the reference point or offset and k the Weber constant or slope

Table 8.1: Typical Weber fraction (or constant) values for different modalities.

Continuum		$\frac{\Delta I}{I}$
Brightness		0.08
Loudness		0.05
Finger span		0.02
Heaviness		0.02
Length		0.03
Taste, NaCl		0.08
Saturation, red		0.02
Electric shock		0.01
Vibration		
60 Hz	0.04	
125 Hz	0.05	
250 Hz	0.05	
Mean Vibration		0.05
Overall Mean		0.04

of the function (which replaces the σ shown in Equation 8.4). C represents the competitor and is introduced at this stage simply to indicate that a competitor is present (the parameters and assumptions for competitors are discussed in sections 8.4 and 8.5). A representation of the effect of incorporating a 'noisy judge' on the model is shown in Figure 8.3, where it can be seen that greater gains (or losses) are now needed in order for higher probabilities of winning or gaining advantage to be judged.

As alluded to above, Bernoulli's (1738) relation between wealth and utility and Weber's Law are logarithmic functions, indeed, it is commonly held that perception is approximately logarithmic. Accordingly the value function models will be calculated in 'log space'. This provides the characteristic shape of the value function of prospect theory (Kahneman & Tversky, 1979; Tversky & Kah-

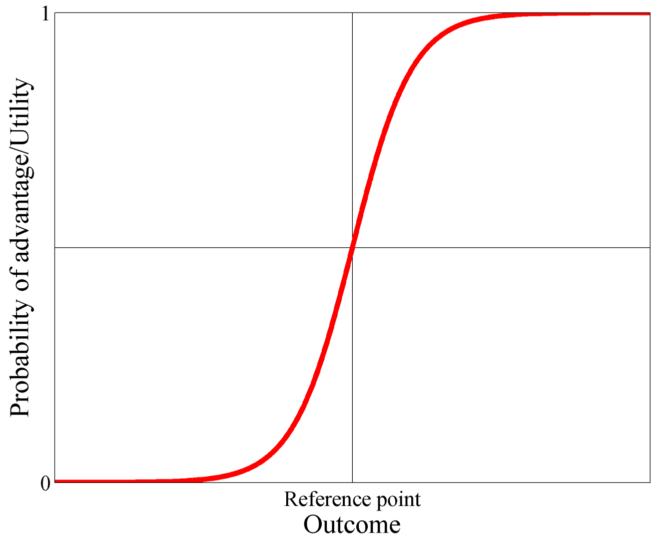


Figure 8.3: The effect of including perceptual noise based on a Weber constant into the value function model.

neman, 1992) where the part of the curve representing losses is steeper than that representing gains; this is illustrated in Figure 8.4.

The model now incorporates an imperfect judge, introducing uncertainty through use of the Weber constant; as 'my' personal reference point gets larger, greater gains are needed in order for the higher probabilities of gaining advantage to be judged. The value function as it is now calculated exhibits loss aversion, that is to say, risk aversion for gains greater than the reference point and risk seeking for gains less than the reference point, the curve for gains less than the reference point is also steeper than for gains greater than the reference point. However, until this point only one competitor has been considered and even then their reference point is not known. As implied above and discussed in Chapter 2 §2.3, this is inadequate since multiple competitors must be considered in a social economy.

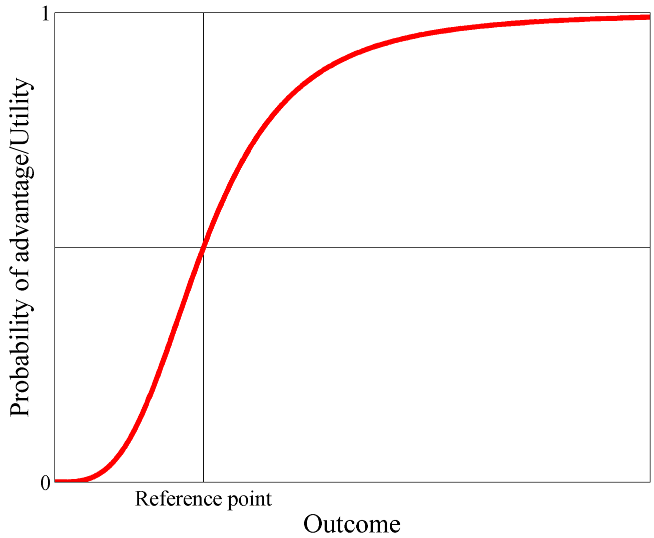


Figure 8.4: The effect of carrying out the calculations on 'log space' and including perceptual noise based on a Weber constant into the value function model.

8.4 Multiple Competitors

The idea of competition, however, immediately raises the problem of what our competition is and how it can be assessed. First though, a way to combine the psychometric curves for potentially many competitors with the psychometric curve for 'me' is required.

On the basis of a competition, given a choice between prospects (and depending on how the prospects are framed, together with the magnitudes involved), the rational way of dealing with them is to combine what I know of my current position with what can be inferred, based on what I know of one or more competitors, using Bayes' rule, choosing the prospect with the higher probability of gaining an advantage. For the value function this requires combining what I know of my own current position with a prior that represents what I can infer for the current position of potential competitors.

This is achieved using a generalisation of the sigmoid activation function, which implements Bayes' rule, as follows (adapted from Elkan, 1997): Of interest is the probability of winning or gaining advantage, W, given a number of competitors, n, $P(W|C_1, \cdots C_n)$; by bayes' rule

$$P(W|C_1, \cdots C_n) = \frac{P(C_1, \cdots C_n|W)P(W)}{P(C_1 \cdots C_n)} \tag{8.8}$$

then the probability of winning would be

$$p = P(W = win|C_1, \cdots C_n) = \left(\prod_{i=1}^{n} P(C_i|W = win)\right) P(W = win) \tag{8.9}$$

and the probability of losing would be

$$q = P(W = lose|C_1, \cdots C_n) = \left(\prod_{i=1}^{n} P(C_i|W = lose)\right) P(W = lose) \tag{8.10}$$

let $W_i = P(C_i|W = win)$ and $L_i = P(C_i|W = lose)$ and changing to logarithms gives

$$log(p) - log(q) = \left(\sum_{i=1}^{n} log(W_i) - log(L_i)\right) + log(P(W = win)) - log(P(W = lose)) \tag{8.11}$$

For the present model, the prior probability of winning or losing is equal so $log(P(W = win)) - log(P(W = lose))$ will equal 1. Making $C_i = log(W_i) - log(L_i$ gives

$$log\frac{1-p}{p} = -\sum_{i=1}^{n} C_i \tag{8.12}$$

Exponentiating and rearranging this equation gives

$$p = \frac{1}{1 + exp\left(-\sum_{i=1}^{n} C_i\right)} \tag{8.13}$$

This type of function is typically found in neural network applications (Elkan, 1997) and its purpose is to make the sum of an output neuron responses equal to one. The outputs are interpretable as posterior probabilities and are thought to be a biologically plausible approximation (Cadieu et al., 2007). To this the values representing 'me' can be added to the sum and providing the basis of the posterior probability, p, of winning in a competition between 'me' and a number, n, of unknown competitors:

$$p = \frac{1}{1 + \exp\left(-\left(\left(\frac{x-o}{k}\right) + \sum_{i=1}^{n} C_i\right)\right)} \tag{8.14}$$

How, though, can competitors be assessed, what can be inferred about them? Using random competitors i.e. using random reference points, would seem to be intuitively inadequate since the world is not random and presumably our choices are not random either. What can be plausibly assumed about competitors that are unknown?

8.5 Shared Environment

It seems plausible to assume that 'me' and my competitors exist in an environment that is essentially common or shared and in the context of the present thesis competitors are assumed to be 'something like me'. This is based on the common sense idea that the competitors in a particular scenario are likely to have the same concerns and that these will be the influences on the choices that they make.

Imagine, for example, the thirty players on the field in a game of rugby union; they are all playing in the same limited space, attempting to score points in a limited number of ways, based on a limited number of set pieces - individual skill and flair is certainly evident from time to time, but for the most part the general run of play is predictable and experience of playing the game is of great benefit. It might be said that those competing exist in a 'small world'. In addition to the idea of a shared environment, support is found in psychology from areas such as Meltzoffs "like me" hypothesis (e.g. A. Meltzoff & Moore, 1977), which attempts, from an infants intrinsic ability to imitate others, to identify how an understanding of other minds is developed. In this development of the value function, it is shown that, rather than being random, the prior for the competitors can be estimated based on the assumption that they are 'something like me' and then combined with 'my position'.

The foregoing does not imply a real competition, but rather, one that is metaphorical and private, where the values for competitors, C, must be estimated. The values for each competitor are calculated in the same way as they are for 'me' i.e. $P(x|o, k)$, where o is the reference point and k is the Weber constant or slope of the function. The value for k is assumed to be the same as the value for me, however, the reference point for a competitor must be estimated. In order to achieve this, the reference point for competitors will be represented by a normal distribution having the same mean as 'me', but the distribution of competitors will have some variance. In other words, this distribution will represent the uncertainty in competitors. From experience (and considering the rugby union analogy), it is clear that other people may be like me, but they vary widely and certainly are not all the same as me.

A competitor therefore is calculated from the weighted marginal probability based on a number of samples from this normal distribution:

$$C_i = \frac{P(x|o_i, k)}{w_i} \tag{8.15}$$

where o_i comes from $P(x|o, \sigma) = \mathcal{N}(o, \sigma^2)$ a normal distribution representing competitors based on the same mean or offset, o, as 'me' and with standard deviation, σ, representing the uncertainty in the competitors. o_i is then $\sum_{s=1}^{t} P(x_s|o, \sigma)$, the marginal probability of t samples from this distribution, where t is proposed to represent working memory capacity. It would, of course, be possible to vary the weight, w, for competitors, in other words making some competitor more likely to win by perhaps giving that competitor a higher weighting.

To reiterate, for the competition that is proposed as the basis of a choice, the competitors must be inferred. Inferring the psychometric function for competitors is based on the plausible idea that they are 'something like me' and that values representing me can be used as a basis for this inference: o is the inferred reference point or mean of the distribution, the same as the reference point, or offset, for 'me'; σ is a new parameter, the standard deviation for the competitor distribution and accordingly higher values represent greater uncertainty; w is also a new parameter, the weight for a competitor (although all competitors are taken as being of equal weight in the testing that follows); and t is the number of samples taken from the competitor distribution, perhaps representing working memory capacity.

8.6 The Value Function

Having obtained these values we can calculate the value function for a particular option in a prospect, which will be the mean (expected value) of a number of sigmoids produced for 'me' and the competitors:

$$U = \frac{1}{d}\sum_{i=1}^{d} p_i \tag{8.16}$$

where U is simply the expected value or the mean probability of winning or gaining an advantage, d is the number of iterations and perhaps represents the amount of risk; and p is the probability of winning or gaining an advantage for a particular iteration calculated using Equation 8.14.

In order to establish how each of the parameters (d, the number of times to repeat the process to establish the mean; t, the number of samples from the inner distribution representing a competitor; σ the standard deviation of the distribution representing a competitor; and n, the number of competitors) affects the value function, each parameter was systematically varied while keeping the others unchanged (see table 8.2).

Table 8.2: Values used for the value function parameters

Parameter	Values
d	8 16 32 64
t	2 3 5 10
σ	1 5 10 50
n	1 6 50 100

8.7 Results

The number of times that the estimation process is repeated is represented by d and is taken to indicate the amount of difficulty, deliberation or risk that is inherent in the options that have to be decided on. Figure 8.5 shows the effect of changing the number of times that the process is repeated, d, which has been tentatively labelled 'difficulty'; larger values of d produce a value function that is less risk averse (i.e. the value function gets more step-like as the number of repetitions increases).

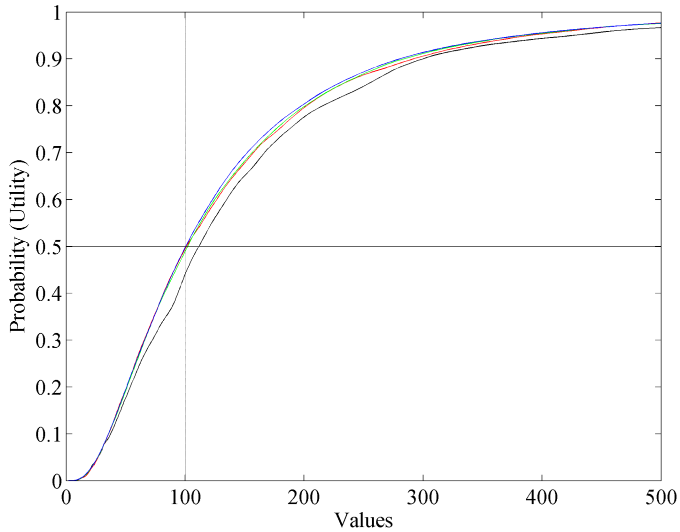


Figure 8.5: The value function showing increasing values (black, red, green, blue) for difficulty, d. It can be seen that, as the values increase, the shape of the curve gets more concave.

In the present context t represents the extent of 'my' experience of competitors in some domain. Figure 8.6 shows the effect of changing the number of samples, t, that are taken from the normal distribution in order to represent a competitor, which has been tentatively labelled 'experience'; larger values of t produce a value function that is less risk averse (i.e. the value function gets more step-like as the number of samples increases). This might suggest that higher values of this parameter imply more experience with an uncertain problem or perhaps that a problem has been experienced before.

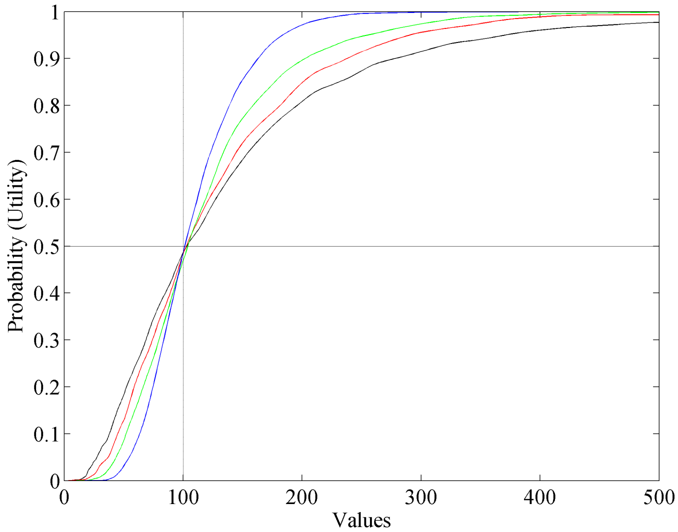


Figure 8.6: The value function showing increasing values (black, red, green, blue) for experience, t. It can be seen that, as the values increase the shape of the curve gets more concave.

The variability in 'my' experience of competitors, in other words 'my' uncertainty, in some domain is represented by σ. Figure 8.7 shows the effect of changing the variance of the normal distribution representing competitors, σ, which has been tentatively labelled 'uncertainty'; larger values for variance produces a value function that is more risk averse (i.e. the value function gets less step-like as the variance increases), suggesting that larger values of variance relate to greater uncertainty in the potential competitors.

The number of competitors that is assumed is given by n. Figure 8.8 shows the effect of changing the number of competitors, n, for a particular problem.

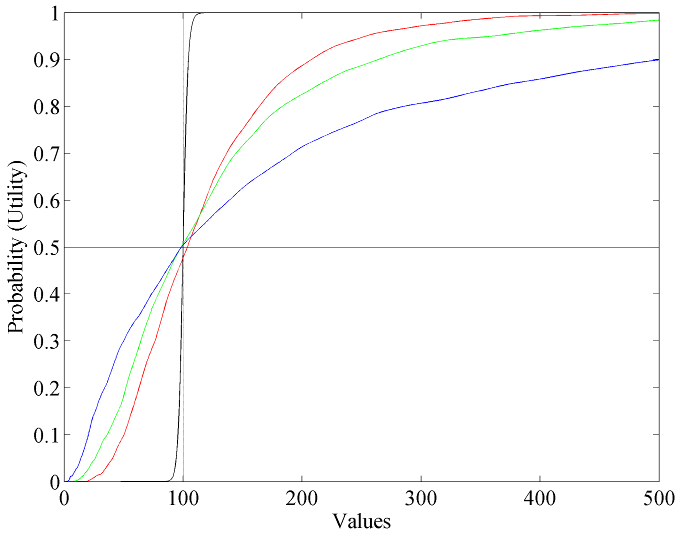


Figure 8.7: The value function showing increasing values (black, red, green, blue) for uncertainty, σ. It can be seen that, as values increase, the shape of the curve gets less concave i.e. more risk averse.

Larger numbers of competitors produce a value function that is less risk averse (i.e. the value function gets more step-like as the number of samples increases). This surprising result suggests that, as the number of competitors grows, the more risk someone might be prepared to accept.

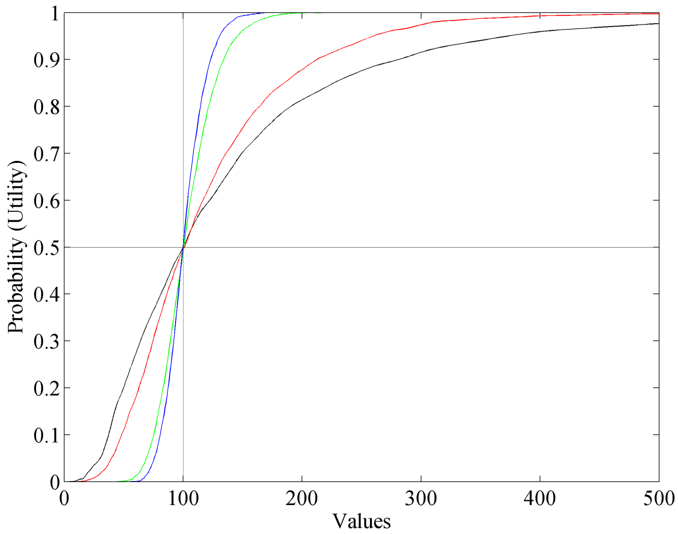


Figure 8.8: The value function showing increasing values (black, red, green, blue) for the number of competitors, n. It can be seen that, as the values increase, the shape of the curve gets more concave i.e. less risk averse.

8.8 Discussion

The present chapter has shown that a value function based on the idea of a competition, where one or more competitors are assumed to be 'like me', shows all of the features of the value function from Prospect Theory; that is, a function that is risk averse above the reference point, risk seeking below the reference point and loss averse overall. As well as the 'like me' or shared environment assumption, the value function also incorporates the idea of winning or gaining advantage based on a fair but not necessarily unbiased observer judging that 'I' would have won the competition.

A juridical element to expected utility arguably goes back to a time before Bernoulli's solution to the St Petersburg paradox and involves an analogy between insuring goods on sea voyages and games of chance (Teira, 2006) and it can also be found in Savages personalistic view of probabilities (e.g. Savage, 1972, §2, p57). What is judged fair is based on the experience of the observer and this is incorporated into the present model using a form of Bayes' rule (illustrated in §8.4). In addition, according to D. Bernoulli (1738) the relationship between gains and utility is logarithmic as is (approximately) the relationship between stimulus and perception in many modalities according to the Weber-Fechner law, which has been the subject of much empirical research. Accordingly, the present model is calculated in 'log space', which helps in defining the characteristic pattern of the value function produced from the competition.

The competition being referred to was stated as being metaphorical and private (§8.5), this is because, even in a real game of say, chess, for example, the choice of next move is made on an entirely endogenous basis, inferring how your opponent might react to any given move of yours and how you might gain best advantage. It was suggested above (§8.5) that this shared environment for making choices might be referred to as a 'small world'; referring to a shared environment in this way inevitably draws comparison with small worlds as talked about by Savage in The Foundations of Statistics (Savage, 1972).

In The Foundations of Statistics (Savage, 1972), the phrase 'look before you leap' is used to describe small worlds and 'you can cross that bridge when you come to it' for expanded worlds. However, as pointed out and quoted by Binmore (2007), Savage explicitly states (1972, p. 16) that "...the look before you leap principle is preposterous if carried to extremes...", which is problematic for the present value function. Nevertheless, in the remainder of paragraph that this quote comes from (Savage, 1972, p. 16), rather than rejecting small worlds as suggested by Binmore (2007), Savage goes on to say that to cross bridges when they are arrived at actually means to tackle relatively simple choice problems by constraining attention to a small world where the 'look before you leap' does apply. Savage goes on to say that he can not formulate specific criteria for selecting small worlds in this way but believes it to be based on judgement and experience. Shared environments as envisaged here are therefore very similar to those of Savage. Despite the plausibility of this argument no compelling evidence has been offered that potential competitors are 'something like me'; fortunately however, much better empirical evidence comes from psychology by way of Andrew Meltzoff's 'like me' hypothesis.

The first step in this process is the infants intrinsic ability to connect observed acts and similar executed acts i.e. to imitate others (A. Meltzoff & Moore, 1977, 1983). Through this imitation and everyday experience, infants take the second step, associating their actions with their own mental states. Thereafter, the third step, infants project their own experiences onto similar acts that are performed by others. Meltzoff argues that, in this way, infants develop an understanding of others minds and their mental states. While there is continued development, especially in sophistication, through childhood and into adulthood people associate the acts of others with connotative meaning because others are processed "like me" (A. Meltzoff, 2005; A. N. Meltzoff, 2007).

Based on these ideas, in addition to the similarities with Prospect Theory (Kahneman & Tversky, 1979) there are some novel predictions that come from a value function construed this way, particularly in terms of risk and competitors. It is interesting to note that larger values of risk, d, appear to produce a more concave, or more step like, value function.

Less risk aversion would seem to be counter intuitive since apparently larger amounts of risk ought, presumably, to cause greater risk aversion. Note though that from the earlier experiments with the semantic differential, particularly in Chapter 3, that as concept values on the potency dimension became more potent, the potency value got more negative or smaller. The potency dimension, which was relabelled as risk, is therefore consistent with the proposed three dimensional reward structure based on the semantic differential. It is also worthy of note that as the value function gets steeper, the more stark or black and white, the choice will become. A further prediction of the model is that as the number of competitors increases, the more risk someone might bear.

Again, bearing more risk might seem to be odd if more competitors equates to more risk and more risk simply makes us more risk averse. However, it is argued here that, on the contrary, this might be intuitively correct if it is considered that, as the number of competitors increases, the more risk someone would be prepared to bear in order to win or to gain an advantage and as with risk as the value function gets steeper, the more stark or black and white, the choice will become. This argument would seem to be all the more plausible if what was at stake was a reward such as food.

The value function, however, is not the only function found in prospect theory, because people behave as if they use a transformed version of a given probability, a further function, the probability weighting function, weights the given probability by a subjective version of it. The probability weighting function is the subject of the next chapter.

Chapter 9

The probability weighting function

9.1 Introduction

Empirical research has shown that, when making choices based on probabilistic options, people behave as if they overestimate small probabilities, underestimate large probabilities, and treat positive and negative outcomes differently. These distortions have been modelled using a non-linear probability weighting function, which is found in several non-expected utility theories, including rank-dependent models and prospect theory (see Figure 9.1); here a Bayesian approach to the probability weighting function is proposed and with it, a psychological rationale.

As already stated in Chapter 7, this function is an inverted 'S' shaped function and has four robust characteristics; it increases the effective probability of improbable events; decreases the probability of probable events; crosses the line of equality at a probability less than 0.5; and is systematically different for the assessment of positive and negative prospects. Here it is proposed that uncertainty plus prior experience gives rise to this function and that it is based on Bayes' rule.

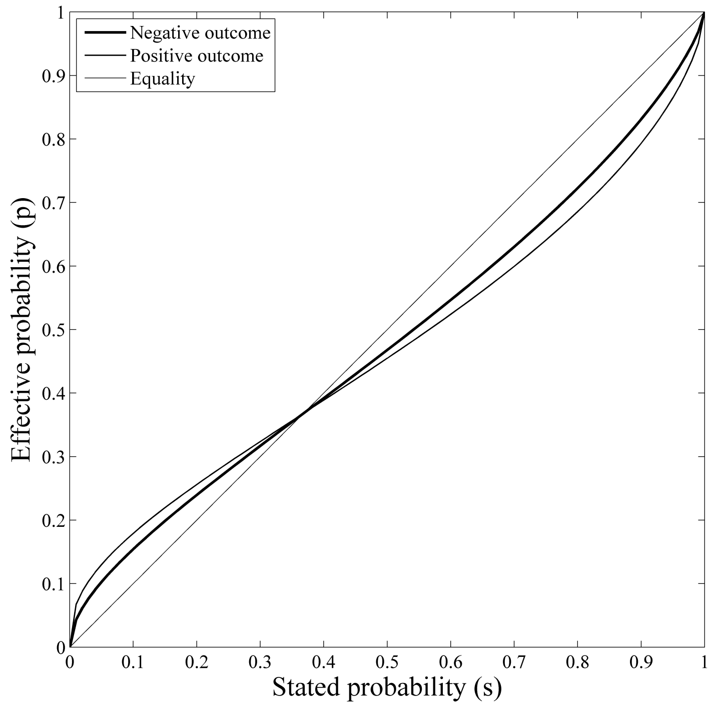


Figure 9.1: The probability weighting function of cumulative prospect theory for positive (thin line) and negative outcomes (thick line) (Tversky & Kahneman, 1992). The line of equality, as predicted by a expected utility theory is shown for comparison (diagonal line). Given a stated probability s, subjects behave as if the probability were p. This function has four main and robust deviations from expected utility theory; small probabilities are over weighted, large probabilities are underweighted, the function crosses the line of equality at a point below 0.5 and the distortions for positive outcomes are more extreme than for negative.

Logically there are two classes of prior; either a situation is like those that have been encountered before and an empirical prior of inference should be used, or a situation is novel and unlike anything that has been previously encountered and a prior representing ignorance should be used. Both classes of prior are routinely used in Bayesian analysis, but unfortunately, on their own, each has problems. Empirical priors rely strongly on the fact that the situation is similar to those encountered before and can be very biased if this is not true, while ignorance priors that ignore previous experience are, potentially, disastrously inefficient.

For the present probability weighting function it is argued that the efficiency of empirical priors and the robustness of ignorance priors should both be used by combining them using standard Bayesian model comparison techniques. This is illustrated using the approach to internet blog searching from Chapter 3 to estimate the prior of inference for generic contexts of 'good' and 'bad' events. The weighting function that is produced accounts for all the major characteristics of the probability weighting function of Prospect theory.

Given any uncertainty in a probabilistic statement (i.e. some probability distribution over the probability), the optimal response for a receiver of such a statement is to combine it, using Bayes rule, with any previous knowledge they have. The model that is described below allows for the possibility that previous experience is not relevant by utilising a combined prior based on a mixture of a form of uniform prior, expressing the possibility of novelty, with a prior based on what has been learned from previous experience.

It is proposed that it is this combination that results in the systematic distortions shown by the probability weighting function. Related methods have been proposed in order to perform robust statistics within the Bayesian framework (e.g. by Jaynes, 1995); a similar approach is used in computer vision systems to model background objects and also in models of animal categorisation to allow for the possibility that an object is novel (Baddeley et al., 2007). It is also proposed that this model can account for the distinction in the decision making literature between description based decisions and experience based decisions, which was discussed in Chapter 7.

In the following the basics of Bayesian modelling that were described in Chapter 7 are expanded to show how the relevant prior distributions were estimated and how these can be combined, using a Bayesian model comparison framework, to form a posterior probability distribution over probabilities given a particular statement. The median of the function very closely resembles the probability weighting function of Prospect theory.

9.2 Method and results

Prospects are statements consisting of the probability of good or bad events happening, so in this case the representation simply consists of s, the stated probability; the inferred distribution of probabilities, p; and the context, c, whether the stated event is good or bad.

It is proposed that when a signaller states that there is an "80% *chance*" of an event happening, what they mean is closer to a statement like "*I have experienced the situation* 10 *times and* 8 *were successful*" rather than something like "*I have experienced the situation* 2×10^{20} *times and* 1.6×10^{20} *were successful*" as implicitly assumed using objective (frequentist) probabilities. More specifically, it is proposed that when someone states a given probability it is interpreted in terms of two components; an explicit probability, s, and an implicit number of events that the statement is based on, N. Statements made by trusted and experienced people about unchanging characteristics of the world are associated with high implicit N (and with modifiers such as "precisely" or "exactly"), while statements made by less reliable people in unusual situations about changeable characteristics (and associated with modifiers such as "roughly", "approximately" or "about") are associated with lower implicit N. Note that in this way N is akin to a measure of confidence or certainty in the statement (N does not refer to the amount of experience or numbers of samples that may have been used to form the prior distribution of inference).

This means that a given statement of probability is not an exact statement, but equivalent to a statement of the form $s \times N$ out of N. If N approaches infinity the prospect reduces to the uncertainty free definition usually used, but for humans it is proposed that $N \approx 5 - 100$ where it is lower in situations where there is less confidence in a prospect e.g. the signaller has modest experience or is not wholly trusted or the characteristics of the world are believed to be changeable (or some combination of these). For example, a stated 80% would be represented as 8 out of 10 (or $s \times N = .8 \times 10$ out of $N = 10$) for modest uncertainty or say 164 out of 205 (or $s \times N = .8 \times 205$ out of $N = 205$) for greater certainty. For present purposes, unless explicitly stated, it is assumed that probability statements are made in terms of the number of successes out of 10 ($N = 10$), but the sensitivity of the model to this assumption is discussed. Given such a definition, the likelihood is binomial and may be calculated using a Bernoulli probability statement (Equation 9.1).

$$P(p|s, N) = \binom{N}{k} s^k (1 - s)^{N-k} \tag{9.1}$$

This defines the representation and the likelihood, but leaves the prior; before specification of the prior, one general property of most plausible priors should be acknowledged. Given a particular probability statement associated with uncertainty and a prior with a given mean, after applying Bayes rule the posterior distribution of probabilities will have a mean that is closer to the prior than the original statement. This implies that as long as the mean of the prior for the probability is roughly 50%, small probabilities will be 'overestimated', and large probabilities 'underestimated', and if we assume, as argued in Chapter 2, that Bayes rule is rational, for rational reasons as well. Figure 9.2 shows this for a $B(5,5)$ prior, a prior with a beta distribution and a mean of 0.5. As can be seen, though this probability weighting function does not match the details that are measured empirically, it shows that the over and underweighting can be simply due to applying a prior. As long as probability statements are associated with uncertainty, applying a prior results in systematic distortions of the effective probability weighting function, but leaves open what kind of prior people might be expected to use.

There are two forms of prior that are relevant here. Either the situation is unlike any encountered before, in which case the relevant prior captures this ignorance (a prior of ignorance), or in contrast, the situation could be like good or bad situations that have been encountered before (an empirical prior of inference). Both types of prior have virtues for the statistical analysis of data. Indeed empirical priors are almost always assumed in models of human performance. Despite this, individually, they have strong drawbacks as the basis of inference in humans. These drawbacks are illustrated below using the simple case of estimating the probability of heads or tails occurring in a coin toss.

Most people have reasonably extensive experience of coin tosses, for example, experiencing 1000 would not be unreasonable in an average student. Such empirical experience would be naturally summarised by the empirical prior $B(501, 501)$, a prior very peaked at a probability of 0.5 and representing previous experience of 500 heads and 500 tails. If a person encounters a new situation involving a coin toss and the coin is 'normal', all proceeds well with such an empirical prior. If

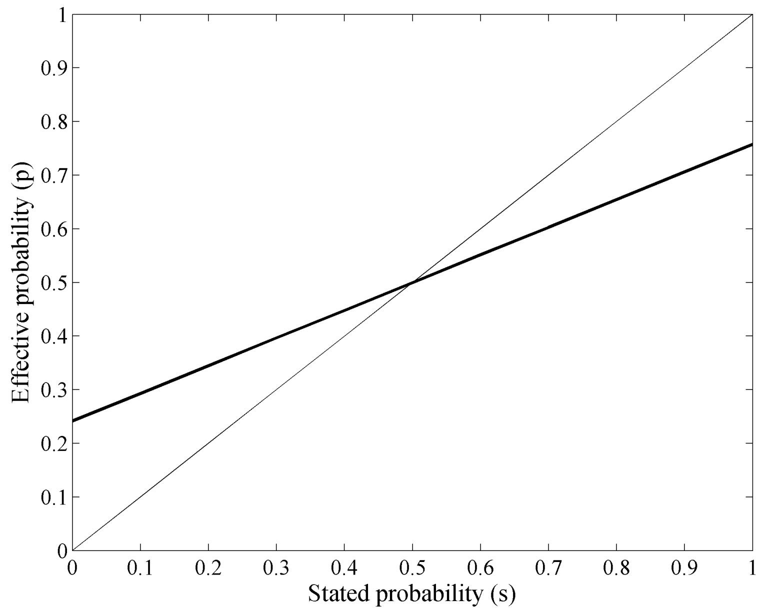


Figure 9.2: A probability weighting function using a single prior based on a $B(5,5)$, a prior with a mean of 0.5 (thick line). This probability weighting function demonstrates that over and underweighting can be attributed simply to applying a prior, though it does not match the details that have been shown empirically.

someone starts with a prior with a mean around 0.5 then observes a few more fair coin tosses, given the highly peaked prior, the result of these further coin tosses will be essentially irrelevant and the posterior will be dominated by previous experience; the person will still have a posterior with a mean of approximately 0.5.

This situation is not maintained, however, if the person now encounters a biased coin (say, one with two heads). After viewing twenty coin tosses, each coming up heads, a reasonable person would become suspicious that their original model of the coin might somehow be flawed. However, even in the face of such strong evidence, a standard empirical prior will still dominate. Compared to previous experience, 20 coin tosses is very few, and based on an empirical prior, will continue to predict that the probability of a tail was roughly 0.5. In fact, even after observing 500 heads in a row, the guess will still be that the probability of a tail occurring is $\frac{1}{3}$. The empirical prior is therefore very efficient for estimating events that are known to have the same characteristics as when encountered before, but disastrously biased if the world has changed, or the event is misclassified.

Priors of ignorance do not share this problem though; if we know that the current situation is unlike any we have encountered before, then it makes sense that our prior should express our initial ignorance of the situation. Much discussion has taken place on what prior best represents our ignorance about a probability. Laplace's original suggestion was that ignorance is a uniform distribution, that is, all probabilities are equally likely. In terms of a beta distribution, this corresponds to a $B(1,1)$ prior. Reasonable arguments have been made for other choices e.g. $B(\frac{1}{2},\frac{1}{2})$ and $B(0,0)$ (Tuyl, Gerlach, & Mengersen, 2008), however, for the present situation, all of these choices generate essentially equivalent predictions and Laplace's uniform distribution is used as the prior of ignorance. With such a prior of ignorance, encountering a two headed coin presents no problem. With such a weak prior, experience will rapidly allow us to conclude that the coin is biased. After observing 20 heads, our mean probability of a tail will be $\frac{1}{21}(or\ 0.05)$, but this robustness to novel situations comes at a high cost. By effectively treating all situations as novel, we are throwing away all of our previous experience. This experience is expensive and potentially dangerous to collect, and while throwing it away may be robust, it is ridiculously inefficient.

Clearly both commonly used priors have problems in certain situations, but it seems clear that what is required is the benefit of both; employ the efficiency of an empirical prior when this is consistent with the evidence we have encountered, but fall back on our prior of ignorance when the evidence is incompatible with our prior experience. Fortunately, this is straight forward to do using the standard Bayesian machinery. The details are given below, but essentially we can effectively use both priors, weighting each by the (posterior) probability that it is appropriate. The probability that one or other prior is appropriate can be calculated using the evidence based model comparison method (based on MacKay, 2003) and by doing this we can efficiently utilise our previous experience for situations we have come across before, but be robust when the world is not the same as we have previously encountered.

In order to address this and because much of the data used in research on choices does not have a well established context, it is important to base the prior of inference on 'real' data from the environment. Accordingly, the data obtained from internet blogs and described in Chapter 3 was used as a representative data source. The details of how the blog data were obtained and used is given in Chapter 3, however, in essence, 1500 concepts were searched for, each associated with good or bad events occurring (by combining them with with a set of modifier words that are relatively unambiguous in their goodness and badness e.g. happy, evil). Marginal probability was then obtained by summing (or integrating, more generally) the conditional probabilities over all outcomes and obtaining the probability distribution of these probabilities.

To illustrate this, consider that birthdays, weddings and Christmas, for example, are associated with good things happening, while house fires, and earthquakes are associated with bad things happening. Nevertheless, these concepts are not exclusively associated with good or bad things, we all know of fights at weddings and arguments at Christmas, while disasters such as earthquakes can be the scenes of good acts and heroism. There are also those occasions when an outcome is neither particularly good or bad, but indifferent. It is clear that, given certain contexts or concepts, there is a higher or lower probability of good or bad things happening; in other words, there is a distribution of good and bad events and these distributions can be different for different concepts. On the basis of the millions blogs indexed it is straightforward to estimate the probability of good or bad events happening. Across all concepts, a distribution of probabilities is obtained and to characterise each of these two distributions (one for good events and one for bad) the best beta distribution was fitted using maximum likelihood. Figure 9.3 shows the distribution of good and bad events together with the best fitting beta distribution. As can be seen, there is significant uncertainty associated with the distribution of probabilities, the average probability is less than half, good things are more common than bad, and the data is well summarised by a beta distribution.

Given these priors and a probability statement, posterior distributions can be calculated, each associated with good inference and bad inference, then combined with the ignorance prior. In order to combine ignorance with good inference and ignorance with bad inference to form the distributions required for the weighting function, an evidence based Bayesian model averaging framework (based on MacKay, 2003) was used. This last step is important as it provides the means to identify the extent to which a situation is like something we have encountered before and therefore the weight to apply to it.

It should be noted that a single data set was not relied on in order to establish these priors, rather, three separate sets of data were explored. Figure 9.4, in addition to the technorati data set that is duplicated from Figure 9.3 for comparison, shows a data set from an alternative blog search engine, Blogscope, together with

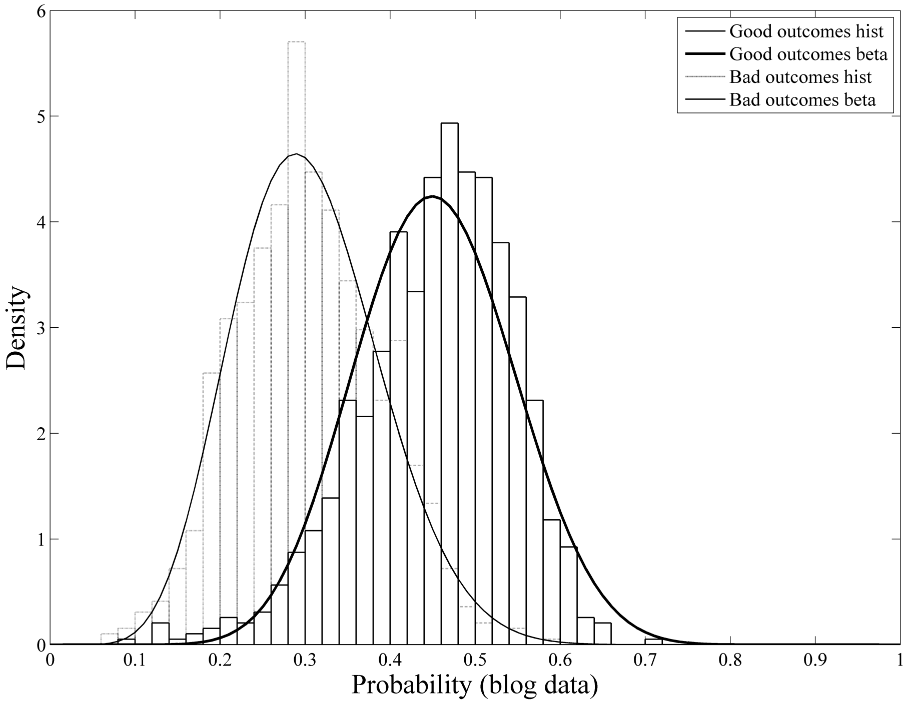


Figure 9.3: The probability distribution of the probability of good (thick line) and bad (thin line) outcomes. The figure shows the data from analysing 1500 different "word" contexts across millions blogs using technorati.com. The lines superimposed show the best fitting (maximum likelihood) beta distributions. There are three robust properties: the mean is less than 50%, there is a fair amount of spread, and they show a Pollyanna effect: "good" events are more probable than bad.

data from analysing 19,043 articles from Reuters-21578 data set (Lewis, 1997). While the exact details across the data sets are different three robust properties can be seen; the mean is less than 50%, there is a fair amount of spread and they show "good" events are more probable than bad, even in the somewhat cynical and pessimistic world described in newspapers.

Essentially the two priors (ignorance and inference) are simply models of the world and their relative probability is determined by the compatibility of the

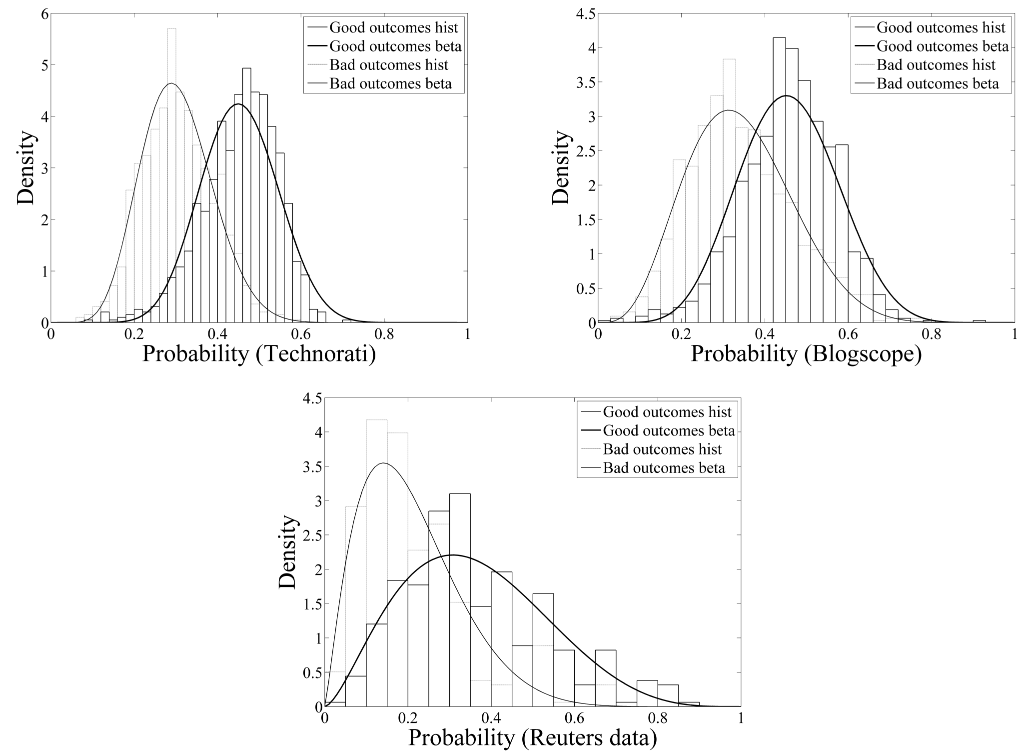


Figure 9.4: The probability distribution of the probability of good (thick line) and bad (thin line) outcomes. The top left panel shows the data from analysing 133 million blogs using the technorati search engine, the top right panel the results from the blogscope search engine and the bottom middle panel shows the results from analysing 19,043 articles from Reuters-21578. The lines superimposed show the best fitting (maximum likelihood) beta distributions.

probability statement with the probability distribution associated with each prior. In statements compatible with previous experience, the effective distribution is dominated by the prior of inference and for statements incompatible with previous experience, the effective distribution is dominated by the prior of ignorance. This is achieved automatically by the application of the rules of probability theory.

In order to calculate the distribution of the inferred probability p and hence its mean and median, the following approach and calculations were used. The posterior distribution is a function of the stated probability, s, the uncertainty in the statement given by an implicit N and is based on two models, one representing ignorance, ig, and the other inference, in, calculated from the internet blog search data. Each of these models consists of a beta distribution with parameters α and β that depend on the model (as stated the ignorance model is simply a $B(1,1)$ distribution). For completeness, a prior probability for the models is also provided, which unless otherwise specified, is assumed to be 0.5. The effective posterior distribution that is of interest, given these two models, is simply the sum of the posteriors for each of the models (ignorance and inference), weighted by the probabilities that they are correct (Equation 9.2).

$$P(p|s,N) = P(p|s,N,M=ig)P(M=ig|s,N) + P(p|s,N,M=in)P(M=in|s,N) \tag{9.2}$$

To be able to calculate this value two quantities for each of the models need to be calculated: First the likelihood of a probability, given a stated probability, inferred N and model (ignorance or inference), $P(p|s,N,M)$; and second, the probability that a model is correct given a stated probability and an inferred N, $P(M|s,N)$. Calculation of the first of these quantities,$P(p|s,N,M)$, was achieved using the beta-binomial Equation 9.3, because the priors for the two models (ignorance and inference) are expressed as beta distributions with parameters α and β, an explicit s and implicit N:

$$P(p|s,N,M) = \binom{N}{Ns}\frac{B(Ns+\alpha, N(1-s)+\beta)}{B(\alpha,\beta)} \tag{9.3}$$

Next, calculating $P(M|s,N)$, relies on Bayes' rule and because there are only two possible models, it can be calculated for the ignorance model as shown in

Equation 9.4. To calculate this value for the other model, the inference model, appropriate values would be substituted into the numerator.

$$P(M = ig|s, N) = \frac{P(s|N, M = ig)P(M = ig)}{P(s|N, M = ig)P(M = ig) + P(s|N, M = in)P(M = in)} \tag{9.4}$$

There are two things to note about this calculation: *a*) this stage is where the prior probabilities of the two models enter the calculations. In general these are assumed to be both 0.5, but in practice, framing effects are expected and the general level of trust in the source of the probability to have an effect on this value; and *b*) that this calculation is different from the more common "Bayes' factor", which is the log ratio of the two probabilities. However, in order to carry out this calculation a further value is needed for each model, $P(s|N, M)$.

This calculation (shown in Equation 9.5) is arguably the most technical stage; in essence the problem here is that the two models have different flexibility, the ignorance model will provide a good account of almost any data, while the inference model will only provide a good model for data that is compatible with previous experience. The problem of comparing highly flexible models (such as high order polynomials), with simple ones (say a linear model), is common in statistics. The Bayesian "evidence" solution to appropriately penalise complicated models is to integrate the likelihoods over the prior. In general this step is intractable and requires either approximations or sampling based methods, however, for the models proposed here, it is possible to analytically derive the relevant quantity:

$$P(s|N, M = ig) = \frac{B(Ns + \alpha, N(1 - s) + \beta)}{B(\alpha, \beta)} \tag{9.5}$$

where α and β are the shape parameters from the beta distribution prior for the ignorance model. The equation for the inference model is the same but with the appropriate values of α and β.

The equations above provide all of the mathematical machinery required here, but a more detailed treatment of related problems is given in MacKay (2003). Figure 9.5 shows the posterior distributions for two probability statements (top panels), together with the combined probability (bottom panels) of the statement being due to the prior of inference (thick line), or ignorance (thin line). Extreme probability statements that are incompatible with previous experience are dominated by P(ignorance). The top left panel shows the posterior probability distribution for the probability statement of *30%*. This is compatible with previous experience and in this case the distribution is dominated by the prior for good events (*P(good)=63.85%*), and is relatively tightly peaked (*mean = 38.43%, median = 39.06%, mode = 36.92%, SD = 9.72%*). The right top panel shows the posterior for a probability statement of *70%* that is incompatible with previous experience of "good" events (*P(good) = 17.54%*), and is therefore dominated by the prior of ignorance. This distribution is based on less "experience" and is therefore more vague (*mean = 72.28%, median = 75.42%, mode = 77.99%, SD = 11.78%*).

The results of all the calculations are (for any given probability statement, s, and implicit uncertainty, N) a probability distribution over the posterior probabilities. To simplify this the median was calculated, which is the solution to $\int_0^1 P(p|s,N)dp = 0.5$. Figure 9.6 shows the mean posterior probability for good (thick line) and bad (thin line) statements, based on priors of inference from the analysis of internet blogs (left panel), the Reuters data set (middle panel) and on the right the probability weighting function based on Prelec (1998) for comparison. As can be seen the present Bayesian function shows all the main features

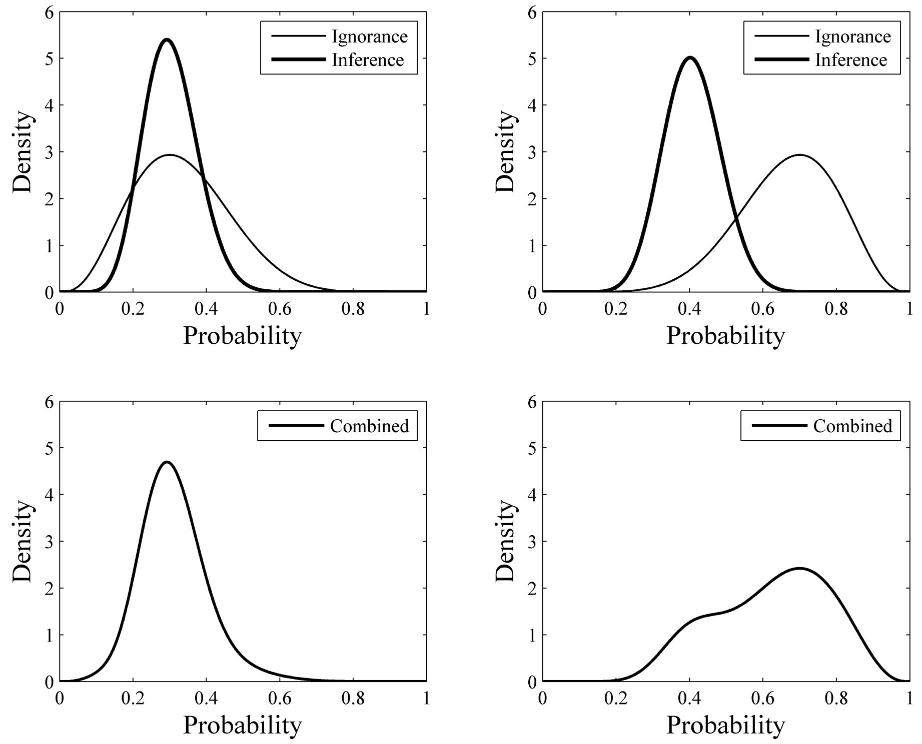


Figure 9.5: Posterior probability distributions for two probability statements. The top figures show the posterior distributions for priors of inference and ignorance separately and the bottom figures show them combined. The figures on the left show the posterior probability distribution for the statement *30% (implicit N=10)* that is reasonably compatible with previous experience of good events. The combined distribution is therefore dominated by the prior for good events. On the right is shown the posterior distributions for a probability statement of *70% (implicit N = 10)* that is incompatible with previous experience of "good" events and is therefore dominated by the prior of ignorance.

of the probability weighting function of Prospect Theory; in both cases it over weights small probabilities, under weights large probabilities, crosses equality below $s = 0.5$, with the differences between good and bad statements mirroring empirical observations.

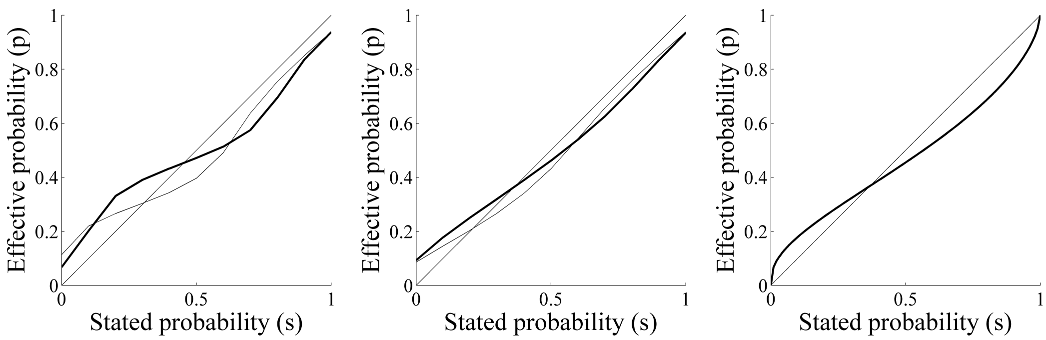


Figure 9.6: The median posterior inferred probability for good (thick line) and bad (thin line) statements (*implicit N=10*) based on priors of inference from the analysis of blogs (left panel) or the Reuters-21578 data set (middle panel). Also shown is the probability weighting function based on Prelec (1998). As can seen the Bayesian functions shows all the main features of the probability weighting function of prospect theory.

9.3 Discussion

It was proposed that because of the inevitable uncertainty that exists in all but exceptional cases when making a choice, that a weighting function is constructed from experience based on a mixture of two priors, one that represents ignorance and one that represents inference. In order to test this hypothesis real world data are needed and two internet blog search engines together with the Reuters-21578 data set were chosen. By proposing that probability statements are uncertain and that people use Bayes rule to incorporate previous knowledge to calculate the most probable probability, it was found that the present Bayesian weighting function accounted for all of the main features of the probability weighting function of cumulative Prospect Theory. This is based on three assumptions.

The first assumption is that a 50% probability of winning is interpreted as equivalent not to an infinite number of wins out of twice this number of bets, but more like five wins out of ten. The important difference is that whereas the former has no associated uncertainty, the latter does, in other words, it is consistent with a range of probabilities, with 0.5 being only the most probable. Given the finite and non-stationary nature of the world, treating probability statements as having uncertainty is rational. It also makes sense of such statements as *exactly 50%*, indeed it is not uncommon to hear children saying "I'll have the bigger half" when they are trying to share, say, a bar of chocolate equally, which under the traditional interpretation is tautological. When given probability estimation problems, such as *what is the chance of you being home by 6pm?*, people naturally prefer to use verbal labels like *probably*, with their answers appropriately associated with uncertainty, rather than to reply *52.2%* with some unrealistic level of certainty (Windschitl & Wells, 1996).

Uncertainty implies that the optimal way to use this information is to update a probability distribution over probabilities. To some, this concept of the probability of a probability is very unnatural, but it is routinely used in diverse areas from machine learning to the analysis of neuronal data. It also implies that any explicit probabilistic statement is (often) associated with an implicit statement conveying the uncertainty associated with it; normally conveyed by context, perceived knowledge, trust of the conveyor of the probability and by adjectives such as *about*, *roughly* or *exactly*. This is quantified in terms of the equivalent number of measurements where the more trusted, knowledgeable and unchanging the probability is believed to be, the closer it is to the line of equality. Significantly, even with relatively high equivalent N statements, the effects of the prior are still evident (Figure 9.7 shows the effect of this quantity).

The second proposal is that evaluating probabilities is, at least partly, based on prior experience. The prospects were interpreted as being unfamiliar and the probability distribution of quantities of rather generic good and bad events was

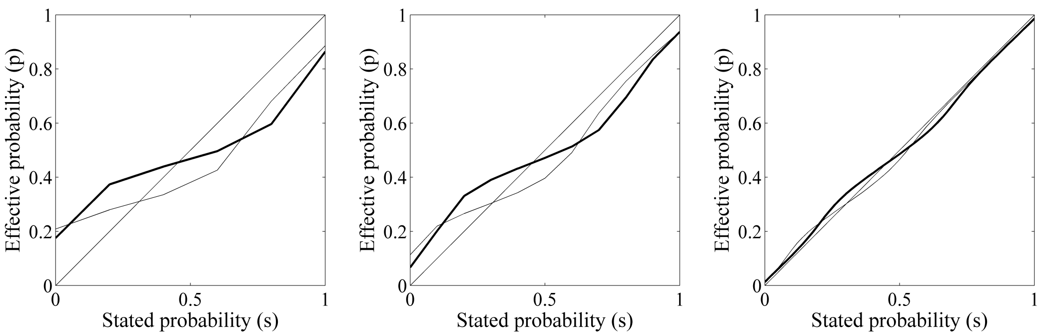


Figure 9.7: The effect of inferred confidence in a "good" probability statement. The left hand panel shows the median posterior probabilities for good outcomes (thick line) and bad outcomes (thin line) for a very uncertain statement (*implicit N=5*); the middle panel a statement with medium implicit uncertainty (*implicit N=10*), and the right panel, a statement of very low uncertainty (*implicit N=50*). As can be seen, the greater the uncertainty associated with the statement, the larger the effect of any prior experience and the greater the deviation from expected utility theory.

calculated. A very large collection of blogs was taken as a representative data set and searched in order to calculate these probabilities, but there is a worry that this may be a rather biased data source, however, two arguments militate against this. Firstly, essentially equivalent results were obtained using different data sources; an alternative blog search engine 'blogscope' and the Reuters-21578 data set (a collection of 19,043 articles that appeared on the Reuters newswire in 1987). Secondly, the characteristics of the distribution required to obtain the results are general i.e. there should be a range of probabilities, that the average probability should be less than 0.5, and that good things should be more probable than bad.

That good things are more probable than bad should not be surprising, the Pollyanna hypothesis, which is a tendency to use positive evaluations more frequently than negative evaluations in communicating, is widely recognised and has been extensively researched (e.g. Boucher & Osgood, 1969; Matlin & Stang, 1978). The effect has also been found across cultures and languages, and argued to be a universal tendency (Boucher & Osgood, 1969). Interestingly for the present proposal, research on autobiographical memory has also shown that positive events come to mind more readily than negative ones (Holland & Kensinger, 2010). However, it is possible that there is a bias to reporting unusual events in blogs; data sources such as Twitter may provide a better characterisation of everyday probabilities, but at the time of writing no machine readable database was known to be available to search.

The last and important assumption is that the prior allows for the possibility that previous experience is not relevant. The idea of utilizing a combined prior based on a mixture of some form of uniform prior, expressing the possibility of novelty, with a prior based on previous experience, is not a new idea and has been applied to a number of different areas of analysis. The main work of appropriately applying the prior is achieved by the model comparison calculations but these, in turn, require a prior for each model (which for the example calculations here was assumed to be 0.5). This though should be susceptible to framing effects where situations described as strange or unusual will be associated with larger values of *P(ignorant)* (Figure 9.8 shows the effect of varying this parameter). On this basis the present model provides for a prior for a new situation to be built from a position of extreme ignorance, based on the experience gained.

Accepting the conclusions of L. Hadar and Fox (2009) to the extent that under sampling and information asymmetry would be the result of building a new prior of inference in a situation that has been recognised as novel (because the likelihood for the inference prior is insignificant): When a situation does not conform to our previous experience the present model is sensitive to it and rather

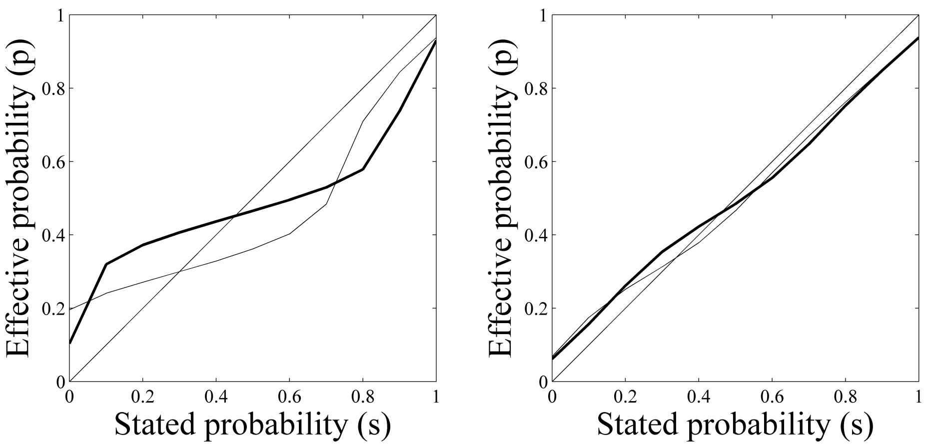


Figure 9.8: The effect of varying the prior probability that this is a novel context: left panel P(ignorant) = 0.15; P(inference) =0.85; right panel P(ignorant) = 0.85; P(inference) =0.15. As can be seen, the more clues that this is a novel context; the less the probability weighting function is distorted.

than slavishly sticking to something that is inappropriate, as might be the case if there were only a single prior, a new prior for that situation can be built from the position of ignorance, based on the experience gained. In building up this new prior it will be susceptible to bias and all the problems of sampling error and information asymmetry that are familiar from the experience based decisions literature and it is argued, accounts for many of the findings in that literature. The top left panel of Figure 9.9 illustrates the results of the present model using an inference prior representing an extreme event, *beta(1,30)*, and the posterior distributions for a stated probability of *10%*.

So far the similarities to the predictions of Prospect Theory have been emphasised, but there is a difference. In Prospect Theory, if you are told a probability is zero, you believe it. In the present account, particularly for bad outcomes, as

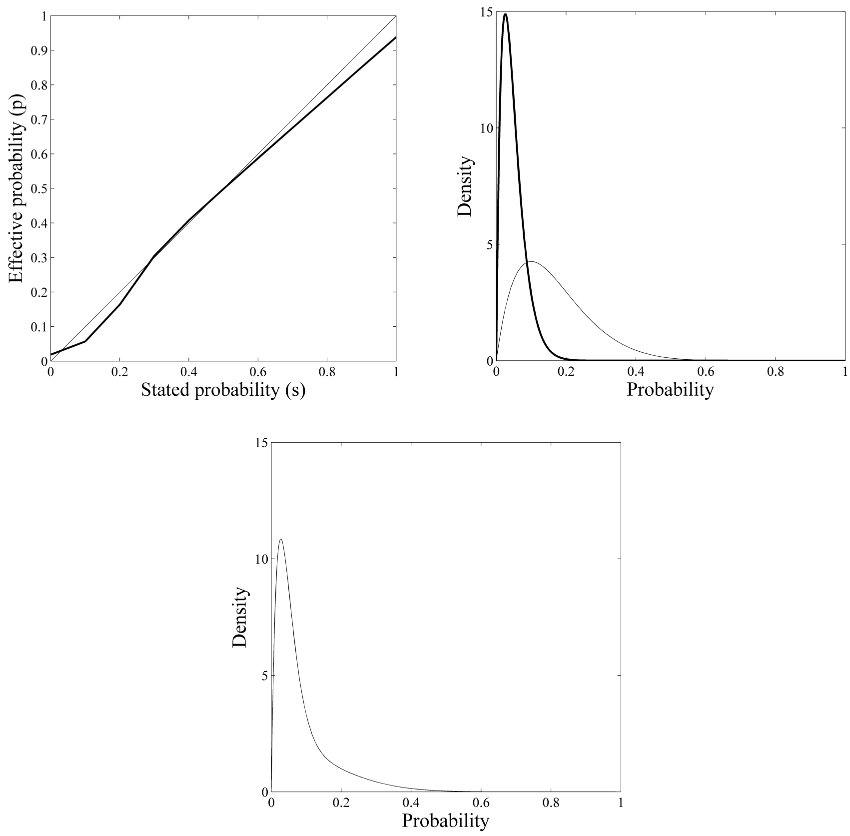


Figure 9.9: Underweighting of small probabilities. The top left panel shows the results of the model based on an inference prior representing an extreme event, *beta(1,30)*, the top right panel the posterior distributions for the priors of inference (thick line) and ignorance (thin line) for a stated probability, s, of 10% and implicit $N = 10$. The bottom middle panel shows the combined effective distribution for these posteriors.

can be seen in Figure 9.7, the probability never goes to zero; you never believe the impossibility of a negative event. Picking up on one of Andersons (1990) arguments, this function embodies an element of regression towards the mean; discounting statements that state or imply certainty, for example, *I'm sure that will happen* or *there hasn't been a run on a bank for 150 years*, is the rational thing to do. This is not, though, in any way to question Allais' paradox, for example, that offers no uncertainty and a guaranteed gain or loss.

Consider the following choice:

> You have £250,000 to buy a house. There are two houses on offer, which are identical in design, location etc. The difference between them is that one house costs £250,000 and the other £249,000.
>
> However, the cheaper one has an unexploded bomb under it, but it's okay, the bomb has been checked and estate agent has assured you that there is a 0% chance of it exploding.
>
> Which house will you buy?

This has not been investigated except informally, but without fail people, unsurprisingly, opt for the more expensive house: they do not accept the stated impossibility.

Taken together the value and probability weighting functions described in the present and previous chapters provide all that is required to be able to model a choice. Each of the functions has been tested in isolation and found to exhibit the expected results, but now it remains for both of the functions to be tested together as a complete model. The testing uses known problem cases from the decision making literature and is reported in the next chapter.

Chapter 10

Putting it all together - the model and testing

10.1 Introduction

This chapter identifies the challenges (often called paradoxes) that research has made to models of decision making, and expected utility theory in particular, which credible models need to be able to meet (Busemeyer & Johnson, 2008). The present model is an expected utility model, but, should produce the same patterns of results that have been found in empirical research, yet do so on the basis of simple maximisation of utility. The present model is tested against each of the types of challenge identified and the results presented.

10.2 Classes of challenge

In order for the full model, consisting of both the value and the probability weighting functions identified in Chapters 8 and 9 respectively, to be credible it must be able to predict the findings from empirical research that other models to a greater or lesser extent find problematic. These findings are known as paradoxes and broadly consist of three classes: preference reversals, dominance effects and context effects (Busemeyer & Johnson, 2008). In this section these broad classes of paradox are briefly described and the violations of Expected Utility theory identified.

10.2.1 Common consequence and common ratio effects

More commonly known as Allais' paradox (Allais, 1953), the common consequence and common ratio effects are probably the best known of all the decision making paradoxes. Allais' original common consequence problem (Allais, 1953, p527) is given below and was provided, by Allais (1953), as a direct counter example to expected utility theory:

1. Préférez-vous la situation A à la situation B?

   SITUATION A   Certitude de recevoir 100 millions

   $$\text{SITUATION B} \left\{ \begin{array}{l} \text{10 chances sur 100 de gagner 500 millions.} \\ \text{89 chances sur 100 de gagner 100 millions.} \\ \text{1 chance sur 100 de ne rien gagner.} \end{array} \right.$$

2. Préférez-vous la situation C à la situation D?

   $$\text{SITUATION C} \left\{ \begin{array}{l} \text{11 chances sur 100 de gagner 100 millions.} \\ \text{89 chances sur 100 de ne rien gagner.} \end{array} \right.$$

   $$\text{SITUATION D} \left\{ \begin{array}{l} \text{10 chances sur 100 de gagner 500 millions.} \\ \text{90 chances sur 100 de ne rien gagner} \end{array} \right.$$

Participants were asked to consider the first pair of prospects, indicating which they would prefer, then to consider the second pair of prospects, also indicating which they would prefer. In order to be consistent with expected utility Theory prospect A and then C or prospect B and then D ought to be chosen. However, this is found not to be the case in empirical research where consistently the preference is found to be A and then D. A quick calculation shows that the preferences are in fact for the lower, less risky, amount for the first prospect and the higher, more risky, amount for the second prospect.

However, this is called the common consequence effect because, according to expected utility theory, altering an outcome by a fixed amount in each of the two gambles should not change the preference for one gamble over the other. Consider a simple example: Suppose you are thinking about buying a new house and there two possible alternatives; does it matter whether both houses are equally furnished or that neither house is furnished? The independence axiom says it should not; whether both houses are furnished or unfurnished is irrelevant to the choice being made, nonetheless, human behaviour appears to be contrary. In Allais' example above, it can be seen that subtracting 89% of 100 millions for options A and B above and subtracting 89% of nothing for options C and D, results in the same gamble. 89% of 100 millions and 89% of nothing are the common consequence in each of the gambles. This is true in Allais' original experiment and in other research, for example Kahneman and Tversky (1979) where the effect is shown with monetary and non monetary outcomes; all in violation of expected utility's independence axiom (the Neumann and Morgenstern (1944) axioms for expected utility theory are given in Appendix A).

A further effect that also violates the independence axiom is known as the common ratio effect. An example of the common ratio effect is shown below:

1. Would you prefer option A or option B?

   A £3000 for certain.

   B £4000 with probability 80% otherwise nothing.

2. Would you prefer option C or option D?

   C £3000 with probability 25%.

   D £4000 with probability 20%.

This effect is also attributable to Allais (1953); participants have a similar preference for A and D rather than B and D, which would be predicted by expected utility theory. Again, a quick calculation shows that not only are B and D the higher value options, but the amounts for the options in both of the prospects are in the same ratio. Theories that can account for common consequence and common ratio effects have been developed but these must also be able to cope with violations of monotonicity or stochastic dominance effects.

10.2.2 Violations of monotonicity

In the context of decision making, stochastic dominance is often referred to in terms of first and second orders, where second order dominance is a weaker form than first order dominance. If a prospect, B, offers at least as good a chance as another prospect, A, of obtaining each possible outcome or better, then prospect B (first order) stochastically dominates prospect A (J. Hadar & Russell, 1969; Bawa, 1975) (strictly speaking B must also offer a better chance for at least one outcome otherwise the prospects are identical). In order to have second order stochastic dominance, prospect B must be more predictable (less risky) and have at least as high a mean or expected value than the alternative, prospect A (J. Hadar & Russell, 1969; Bawa, 1975). In order to illustrate stochastic dominance and its associated problems consider the example below from Tversky and Kahneman (1986):

The following pair of lotteries, described by the percentage of marbles of different colors in each box and the amount of money you win or lose depending on the color of a randomly drawn marble. Which lottery do you prefer?

Option A	90% white	6% red	1% green	1% blue	2% yellow
	$0	win $45	win $30	lose $15	lose $15
Option B	90% white	6% red	1% green	1% blue	2% yellow
	$0	win $45	win $45	lose $10	lose $15

With this first example it is quite easy to see that Option B dominates Option A, indeed, Tversky and Kahneman (1986) found that all of the participants who were presented with this choice chose Option B. However, things were not so obvious in a similar choice:

Option C	90% white	6% red	1% green	3% yellow
	$0	win $45	win $30	lose $15
Option D	90% white	7% red	1% green	2% yellow
	$0	win $45	lose $10	lose $15

In this choice, the options are equivalent from a stochastic dominance perspective, with Option D dominating Option C, but Tversky and Kahneman (1986) found this time that 58% of participants presented with the choice went for the dominated Option C. In order to better understand what is happening with stochastic dominance, consider a further example from Birnbaum and Zimmermann (1998):

Option F	85%	5%	10%
	$98	$90	$12
Option G	90%	5%	5%
	$98	$14	$12

Presented with these prospects most participants preferred F, but F is stochastically dominated by G. This can be seen by rearranging the prospect in the following way:

$$\text{Option F'} \begin{cases} 85\% & 5\% & 5\% & 5\% \\ \$98 & \$90 & \$12 & \$12 \end{cases}$$

$$\text{Option G'} \begin{cases} 85\% & 5\% & 5\% & 5\% \\ \$98 & \$98 & \$14 & \$12 \end{cases}$$

Not only is it easy to see that G' is the dominant option, but, when presented with the prospect in this form, participants chose G'. Examples of experimental violations of stochastic dominance and independence are perhaps the most common in the decision making literature, but, nevertheless maybe not the most difficult to resolve: A further challenge that is also taken seriously in the literature is preference reversals.

10.2.3 Preference reversals

Preference reversals are considered to be problematic for most utility theories because they conflict with a fundamental prediction of those theories (Busemeyer & Johnson, 2008). If one prospect is preferred to another, it follows that it has higher utility; accordingly this implies that the price equivalent of the former prospect is greater than the latter. However, consider the following example from Grether and Plott (1979):

P: win $4 with 35/36 probability;
D: win $16 with 11/36 probability.

When offered this prospect and asked for a direct preference, significantly more participants chose option P than option D. However, when asked to provide a price equivalent option D was preferred. Preferences also reverse when participants are required to provide different types of prices: minimum selling prices or willingness to accept (WTA) and maximum purchase prices or willingness to pay (WTP) (Birnbaum & Zimmermann, 1998).

Nonetheless, potentially at the risk of being controversial, the following alternative interpretation is offered, which is consistent with the present thesis: If I accept one of the two prospects offered above (P or D), then I will presumably opt for the one with the greater chance of winning and since P gives me a 97.25% chance of winning (against D of 30.56%), I will choose P, consistent with empirical data. Alternatively, if I offer you (i.e. I stand to lose if you win) the same choice then I will price them according to what I stand to gain, so for P I will stand a 2.75% chance to gain against D where I will stand a 69.44% chance and presumably I will choose D, pricing it accordingly. If this is right then it appears that participants are being consistent in their choices, choosing the prospect that offers the greatest advantage in each case. In addition to preference reversals, there is a further class of challenges that are concerned with the problem of context.

10.2.4 Context dependent preferences

Context dependent preference effects violate a principle called independence from irrelevant alternatives (Tversky & Simonson, 1993) and have been found with both humans and animals (e.g. Hurly & Oseen, 1999). Two context dependent effects, similarity and attraction, are identified here. Both of these effects involve adding a third prospect to an original two prospects offered to participants.

10.2.4.1 Similarity

The first context dependent effect is known as the similarity effect and occurs when there is initially a choice from two alternatives, say A and B, and one of them is chosen more frequently than the other, say B. Adding a third alternative, C, similar, but not dominating B, causes the preference to change to A more frequently. Consider the following example from Tversky and Sattath (1979):

Context similarity Problem 1

Candidate A: Intelligence = 60, Motivation = 90

Candidate B: Intelligence = 78, Motivation = 25

When participants were presented with context similarity Problem 1, preference was for candidate B. However, when participants were offered context similarity Problem 2, which is shown below with an additional option, C, preference switched from candidate B to candidate A

Context similarity Problem 2

Candidate A Intelligence = 60, Enthusiasm = 90

Candidate B Intelligence = 78, Enthusiasm = 25

Candidate C Intelligence = 75, Enthusiasm = 35

10.2.4.2 Attraction

This effect occurs when there is initially a choice from two alternatives, say A and B, as in the similarity effect and one of them is chosen more frequently than the other, say B. When a new option, C, is added that is similar to A but is dominated by A, preference changes to alternative A.

In one experiment (Simonson, 1989), participants were required to choose between cars that differed in miles per gallon and ride quality:

Attraction problem 1

Brand A: 73 rating on ride quality, 33 miles per gallon (mpg)

Brand B: 83 rating on ride quality, 24 mpg

In this initial attraction problem, participants preference was for car brand B. A similar problem with a third brand added, which was similar to but dominated by brand A, is shown below:

Attraction problem 2

Brand A: 73 rating on ride quality, 33 mpg

Brand B: 83 rating on ride quality, 24 mpg

Brand D: 70 rating on ride quality, 33 mpg

In a reversal of the preference found for attraction problem 1, participants presented with this choice, preferred brand A.

The five challenges (or paradoxes) identified above form the basis for testing the present full model. Although the model is an expected utility model, it is expected that, with utility construed as it is, the same pattern of preferences will be seen as in experimental research based on human performance.

10.3 Method

The full model consists of the value function identified in Chapter 8 and the probability weighting function identified in Chapter 9. Because it is an expected utility theory, the preferred alternative in a choice problem will be the one that attracts the maximum utility, as calculated by the model, for each alternative. The implications of choosing between alternatives, A, are explored below.

For an expected utility theory the utility of a particular option is the sum of the utilities for all of the sub options, m, multiplied by the probability of those sub options (Equation 10.1):

$$U^{preferred} = \operatorname*{argmax}_{a=1}^{A} \left[\sum_{i=1}^{m} P(U_i^a) U_i^a \right] \tag{10.1}$$

For each of the paradoxes identified in the sections above, Equation 10.1 was implemented in a separate Matlab function in the following way:

1. obtain average sigmoids over competitors for each part of the prospect. For the Allais paradox prospects options 1 and 2, this meant obtaining values for 100 and 500 based on a reference point that was arbitrarily set at 100 (note that the values tested were reduced from Allais' original by dividing each by 1 million);

2. obtain weighted probabilities for each of the stated probabilities. Again, for Allais' paradox options 1 and 2 these were 89%, 10% and 11%;

3. utility is then a simple matter of calculating the product of these quantities, which can then be compared and the counts of the preferences recorded.

Assuming the three dimensional reward structure discussed in Chapter 3 suggests that three of the parameters that are required for the model consist of the dimensions from that reward structure, where r is the reward dimension; d is the risk dimension; and σ the uncertainty dimension. This leaves t which, it is proposed, is governed by short term memory capacity, for example, the magic number 7 ± 2 (Miller, 1956). Each of the Matlab functions was tested using a fixed set of parameter values, each for a hundred repetitions. The parameter values were as follows and were chosen arbitrarily since there is no *a priori* principled reason that the parameters should be different between paradoxes (though the parameters may be different if the paradoxes were tested across different (external) conditions):

1. personal reference point $r = 100$ + reward determined by the paradox;
2. risk (or danger or difficulty) $d = 10$;
3. uncertainty $\sigma = 4$ a modest variation given r;
4. experience of a context $t = 5$ perhaps governed by short term memory capacity;
5. the number of competitors $n = 1$.

10.4 Results

Each of the paradoxes was tested in the same way, simply by setting the parameters shown above and recording the results for 100 repetitions of the model. For each paradox these results are given in Table 10.1 and shown graphically in the subsections below.

Table 10.1: Preferred options chosen by the model for each prospect of the paradoxes that were tested. It can be seen that the inconsistent choices seen in empirical research are also seen with a semi arbitrary parameterisation of the model. It should be noted that, although the numbering of prospects and options differs from the descriptions of the original research given above, in the table below the first prospect is always referred to as Situation 1 and the second as Situation 2 with the respective options sequentially numbered 1, 2 and 3 if required, this has been done for ease of presentation.

Paradox				Situation 1		Situation 2		
				Option	Option	Option	Option	Option
	σ	t	j	1	2	1	2	3
Allais	4	5	10	100	0	27	73	n/a
Stochastic dominance	4	5	10	84	16	30	70	n/a
Preference reversals	4	5	10	99	1	0	100	n/a
Similarity	4	5	10	44	56	39	29	32
Attraction	4	5	10	26	74	37	29	34

10.4.1 Allais paradox

Allais' paradox offers two situations, each of which has two options, one more certain than the other. When presented with the first situation, participants typically prefer the more certain option (option 1 in the present example). Conversely, when presented with the second situation, the less certain option is typically chosen (option 2 for in the second situation in the present example). Figure

10.1 shows the preferences found by the model for the example of the Allais paradox described in §10.2.1. The differences between the preferred and non preferred options for each of the prospects were tested using the binomial test and found to be significant; all $p < .0001$.

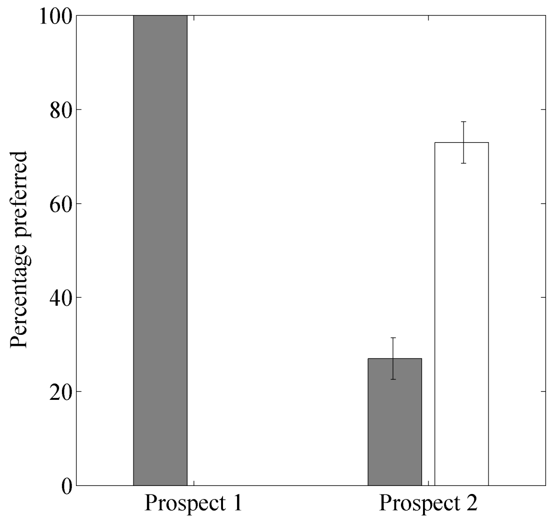


Figure 10.1: The preferences for each of options for each of the prospects of Allais' paradox are shown, illustrating the change in preference from option 1 in the first prospect to option 2 in the second prospect. Error bars are standard errors.

10.4.2 Stochastic dominance

Recall that if an option offers at least as good a chance as another of obtaining each possible outcome and a better chance for at least one outcome, then it stochastically dominates the other option. Using the example from Birnbaum and Zimmermann (1998) described in §10.2.2, where preference changed, from the dominated option in the first situation (option 1 for the present test), to the non dominated option in the second situation (option 2), Figure 10.2 shows the preferences found by the model. The differences between the preferred and non preferred options for each of the prospects were tested using the binomial test and found to be significant; all $p < .0001$.

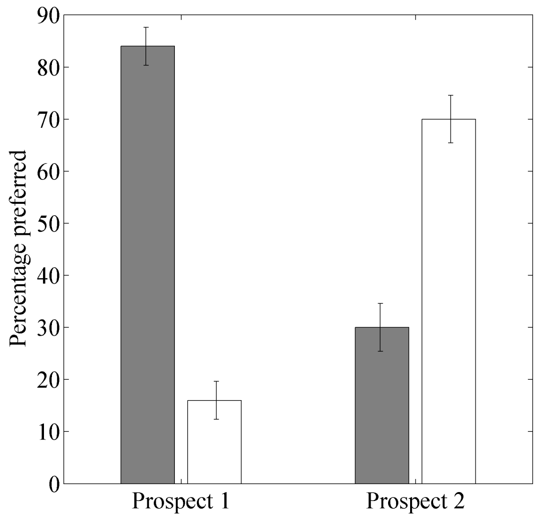


Figure 10.2: Preference results for the example of stochastic dominance. The preferences for each of options for each of the prospects are shown and illustrate the change in preference from option 1 in the first prospect to option 2 in the second prospect. Error bars are standard errors.

10.4.3 Preference reversals

Although preference reversals, as described in the literature (e.g. Grether & Plott, 1979), are not of the form described in Section 10.2.3 above, the reversal in preference can be clearly seen in Figure 10.3. Using the example from Grether and Plott (1979), preference reverses, from option 1, in a buying situation, to option 2, in a selling situation. The differences between the preferred and non preferred options for each of the prospects were tested using the binomial test and found to be significant; all $p < .0001$.

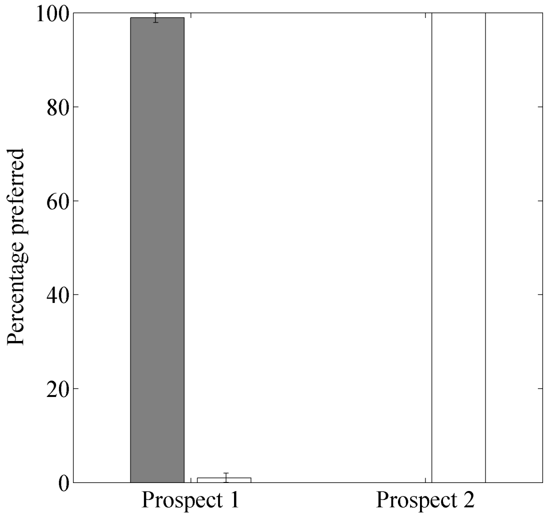


Figure 10.3: Preference reversal results. The reversal in preference is due to the perceived difference between values when buying and selling. The preferences for each of options for each of the prospects are shown and illustrate the change in preference from option 1 in the first prospect to option 2 in the second prospect. Error bars are standard errors.

10.4.4 Similarity

The similarity effect is tested using the example from Tversky and Sattath (1979) and described in described in §10.2.4.1. Figure 10.4 shows the preferences found by the model for a first situation where option 2 is preferred out of the two options presented and a second situation where preference changed to the original option 1 when a third option, similar to the originally preferred option, is added. The differences between the preferred and non preferred options for each of the prospects were tested using the binomial test and found to be significant; all $p < .0001$.

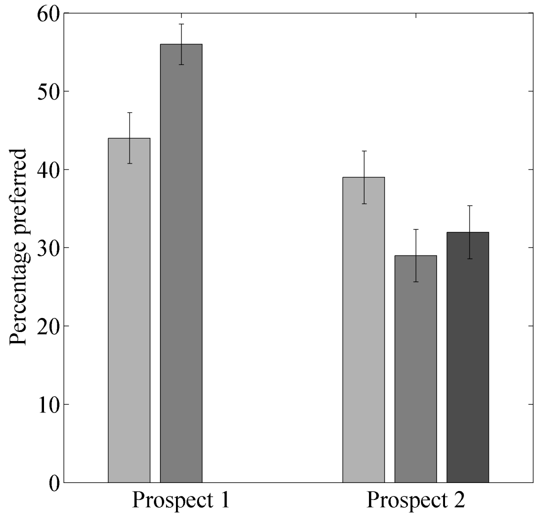


Figure 10.4: Preference results for the example of the similarity effect. The preferences for each of options 1 and 2 for each of the prospects are shown and illustrate the change in preference from option 2 in the first prospect to option 1 in the second prospect when a third option, similar to the originally preferred option, is added. Error bars are standard errors.

10.4.5 Attraction

This test uses the example of the attraction effect described in §10.2.4.2 from (Simonson, 1989). Figure 10.5 shows the preferences found by the model for an initial situation where option 2 is preferred out of the two options presented and a second situation where preference changed to the original option 1 when a third option, similar to the originally non preferred option, is added. The differences between the preferred and non preferred options for each of the prospects were tested using the binomial test and found to be significant; all $p < .0001$.

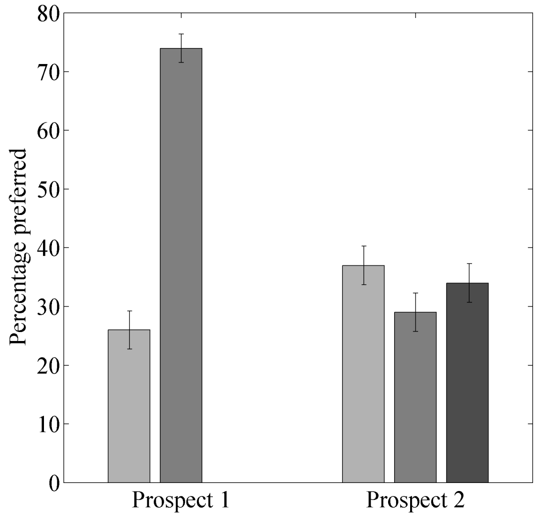


Figure 10.5: Preference results for the example of the attraction effect. The preferences for each of options 1 and 2 for each of the prospects are shown and illustrate the change in preference from option 2 in the first prospect to option 1 in the second prospect when a third option, similar to the originally non preferred option, is added. Error bars are standard errors.

10.5 Discussion

The main classes of challenge to decision making models were identified as common consequence and common ratio effects, violations of stochastic dominance, preference reversals and two forms of context dependent effect; similarity and attraction. The full model consisting of the value function and probability weighting function, identified in Chapters 8 and 9 respectively, was tested against examples from each of these classes. Testing preference reversals used a novel and potentially controversial interpretation of what participants are doing when this effect is seen. If this interpretation is right then not only are the results from the model showing the same pattern as experiments, but the results are rather unsurprising because the comparison appears simply to be between a gain and a loss.

In all cases it was found that the pattern of results that were produced by the model were similar to those found in the cited experiments but does so based on the maximisation of an expected utility. Although implemented in separate Matlab functions, all of the challenges were implemented the same way, as identified in Section 10.3, and used the same set of parameter values.

Clearly, it is difficult to identify a particular parameterisation as correct, and meaningful statistics, beyond the binomial tests that were carried out, are difficult, however, the results can be considered in terms of parsimony, suggesting that the parameters do not need to be altered in an arbitrary manner per paradox. It may be, of course, that alternative parameters are required in order to model a situation of particular framing or particular context, however, this was not tested.

Further testing should be carried out, identifying scenarios that are capable of being implemented with other sets of parameters and levels of competition. However, despite these caveats, it has been shown that an expected utility theory that takes account of a social element can deal qualitatively and in many cases quantitatively, can meet the challenges that are posed in a range of, so called, decision making paradoxes.

Chapter 11

General discussion

11.1 Introduction

This thesis has fallen naturally into two parts, an experimental part, consisting broadly of four experiments, and a modelling part, consisting of a mathematical model of a decision process and its testing, also falling broadly into three parts, which were motivated by some simple observations. This chapter draws the thesis together, summarising the findings of each chapter, identifying its implications and weaknesses, and suggesting areas of further research.

11.2 Motivation

The present thesis was initially motivated by the challenge of Antonio Damasio's Somatic Marker hypothesis (1994, 1996, 2000) to the traditional or classical view that rationality is a cold calculation requiring a world without surprises and knowledge of all the relevant alternatives and their consequences (e.g. Simon, 1979). Indeed, Descartes' error was mind/body dualism and, for Damasio (1994), the separation of rationality and emotion.

Chapter 2 began by reviewing the Somatic Marker hypothesis and discusses the evidence that supports it, venturing into other areas of research including moral psychology and automaticity. As a result, it seems that our choice making behaviour may be largely automatic and based on affective processes but that these can be overridden by (or operate in conjunction with) a conscious deliberative process. A significant issue is also identified: Maximisation of a simple reward is unsatisfactory because it ignores the risk and uncertainty that is, or might be, associated with gaining the reward and consistent with Rushworth and Behrens (2008) it is proposed that models of decision making can be improved by considering a richer conception of reward consisting of reward, risk (or cost) and uncertainty; this is one of the central parts of the present thesis which can also be motivated by considering the environment in which choices are made.

In their classic treatment of expected utility Neumann and Morgenstern (1944) consider rational behaviour and contrast a very simple economy, populated by a single person referred to as Robinson Crusoe, with a more complex social economy. While Robinson Crusoe's problem is one of simple maximisation, to the extent that external conditions are given and all he has to do is make choices that make his situation as good as it can be, people in a social economy must enter into competition with others as well as having their own preferences. This is also illustrated and supported with examples from cross cultural research on the Ultimatum game and subtle cues. Due to the changing nature of the decision making environment and the inherent uncertainties of competitive relationships Neumann and Morgenstern (1944) argues that, what for Robinson Crusoe is a simple maximising problem, turns into a complex and disconcerting mixture of conflicting maximising problems for people in a social economy. It seems that the only way that this can be managed is through applying previous experiences to new situations and this can only be achieved through a process of induction.

Nevertheless, the following section concerning new knowledge identifies that induction is a potentially thorny issue, particularly philosophically, and, despite its very long history, cannot be justified from a reasoned logical point of view. Frank Ramsey (Ramsey, 1926) summed up the position concerning induction in the following quote “We are all convinced by inductive arguments and our conviction is reasonable because the world is so constituted that inductive arguments lead on the whole to true opinions. We are not, therefore, able to help trusting induction, nor, if we could help it, do we see any reason why we should”. Assuming the choices that we make are inductive, learning from experience and applying our prior experiences to new, but similar, situations, allows a different question from induction itself being justified to be posed: Whether beliefs and opinions can be modified in a logically justified way based on additional experience or evidence. There is. Using Bayes’ rule.

Using Bayes rule accepts the descriptive explanation of induction, whether it is reasonable or not, and identifies a model of reasonable change in belief that is sufficient for being rational in a changing world. Taking induction, a multi dimensional reward structure and the somatic marker hypothesis together implies that there is a marker that can be maintained for every concept that we know. The semantic differential (Osgood et al., 1957) is a well researched psychological instrument that is thought to measure the connotative meaning of a concept in three dimensions, however, in spite of all of the research on and with the semantic differential, the question ‘what is the semantic differential really measuring’ may still be posed, indeed, although its usefulness and reliability remains unquestioned, this has been a recurring question for most of the semantic differentials history (e.g. Miron, 1969; Osgood, 1969). It is proposed in the present thesis that the semantic differential is a representation of our reward structure. Early support for this approach came from David Heise, a long standing practitioner with the semantic differential and pioneer of Affect Control Theory, who said, with respect to the semantic differential, that had he been entering the field today he would be interested in choices and decision making because to him it is "obvious that those kinds of decisions involve emotion and affect" (personal communication, December 25, 2010).

The remainder of Chapter 2 identifies problems for expected utility theory, expecially in terms of axiom violations and the changes that were made to it by prospect theory, which has become the most popular alternative to the expected utility theory. Chapter 2 also includes a brief discussion of how learning and reinforcement sufficient for making choices might exist at the level of neurophysiology, in particular surrounding the actions of the neurotransmitters dopamine, serotonin, acetylcholine and norepinephrine.

So to sum up the background to and motivation for the present thesis, it was identified that: *a*) making choices is to a great extent affective and that the Somatic Marker hypothesis offers the basis of a Bayesian type prior; *b*) making choices is competitive, whether the competitors are actual or perceived; *c*) there is inherent uncertainty in the world and in the competitors; *d*) in order to maximise advantage, cost and uncertainty should be considered alongside reward, rather than just reward; and *e*) the semantic differential may offer a reliable representation of that reward structure.

11.3 Experimental findings

Chapter 3 began the main experimental work by investigating whether the semantic differential can be plausibly considered to measure a three dimensional structure representing reward, risk/cost and uncertainty. In order to be able to establish whether a relationship existed between peoples experiences and the semantic differential, a set of rated concepts and an assessment of the experiences that related to them was required. The rated concepts were easy to acquire from a publicly available source (i.e. Francis & Heise, 2006) and required little further work to make them usable, however, a novel approach was needed to gathering data that could be considered to represent peoples' good and bad experiences.

These 'experiences' data were collected by carrying out automated searches of internet weblogs or blogs through the use of two blog search engines, technorati and blogscope. Each search consisted of a concept word that had been pre-rated using the semantic differential and a disjunctive list of modifier words. From the data that were collected, six measures of reward were derived and used with regression techniques to investigate the relationships with the semantic differential.

The key findings in this experiment were similar for both search engines, one of which was unexpected. There was essentially no correlation with the absolute proportion of positively rewarded events: Evaluation is to first approximation simply relative reward representing the ratio of good events to bad, it does not represent the probability of good events happening as might be intuitively thought. The second dimension, Potency, essentially measures the risk (of bad things happening), making it clear why Potency needs to be represented for every object we can make decisions about; it is important to know not only the average reward associated with an option, but what the cost (risk/danger) might be in obtaining it. Associating the semantic differential with a multi dimensional representation of reward is also consistent with the action of the neurotransmitters discussed in Chapter 2 §2.7 and all the more so if Evaluation represents relative reward as found in this experiment.

It might be argued, however, that reducing prior experience to the apparent experience of two outcomes is too limited, but it is argued that this is not the case, what is potentially limiting though is the data sources that were used. Although two blog search engines were used, it might be argued that they are in fact just two access points to the same source and in any case might suffer the same biases. It would be desirable to have other data sources, of the same scale, that could be searched in a structured manner in order to provide an alternative to, and a cross check for, the present findings. Facebook statuses have the potential to provide the necessary information, however, Twitter, with the storage of 'tweets' by the National Library of Congress has the greater potential, though neither potential data source offered an accessible interface at the time of writing.

The blog search experiment was correlational only and it can not be inferred that the good and bad experiences accessed through the blog search are causing the reward structure, and thereby the semantic differential, to be a particular way or otherwise altered. In the next experiment, labelled AlphaBet and reminiscent of the Iowa gambling task that has been used extensively to test the prediction of the somatic marker hypothesis, Chapter 4 successfully tried to influence a semantic differential through differing distributions of rewards.

The AlphaBet experiment demonstrated that our reward structure and as a result, the semantic differential, could be manipulated. Participants were asked to make a series of bets on the outcome for arbitrarily coloured shapes. Each shape was associated with a different probability of gaining a reward, nothing happening and loss, for which the participant was given explicit feedback. This was thought to represent an economic type choice that was experienced as good, bad or indifferent.

Ratings that were given for each shape after participants had completed a series of bets formed the expected factors of the semantic differential, as with the blog search experiment, these, together with the experience data (given by the trial outcomes) were analysed using multiple regression techniques. The same pattern of results was found in the AlphaBet experiment as was found in the blog search experiment, particularly that the Evaluation dimension relates to relative reward and that Potency relates to risk. The main claim of the AlphaBet experiment therefore is that, on the basis of experiences from betting on arbitrary shapes (representing an economic decision), that the semantic differential can be manipulated and offers confirmation of the findings of the blog search experiment that, to a first approximation, the semantic differential is a summary of the reward history. It is interesting to note that, despite the massive difference in the size of the data sets that were used, the pattern of the relationship between Evaluation and Potency (or reward and cost), illustrated in Figure 4.5, is surprisingly similar.

While the winning shape distributions in the AlphaBet experiment are the only ones to have a reduced level of activity, which is consistent with the idea that this dimension represents uncertainty, it is not significant. Indeed, neither the AlphaBet nor blog search experiments directly or explicitly address the Activity dimension. Based on the idea that the Activity dimension of the semantic differential is more accurately described as control or certainty, Chapter 5 attempts to focus on this dimension with an experiment labelled as the Triangles experiment.

By keeping other variables, such as the stimulus colour and shape, static and just manipulating the way that the size of the shape changed, the Triangles experiment focussed on how good the predictions that participants made were. At one of the extremes were random changes to the shape size, which were expected to be the most unpredictable and consequently the most uncertain (or most active) and at the other extreme an unchanging or uniform size, which was expected to be most certain (or least active). In addition, because, unlike the AlphaBet shapes experiment, no explicit rewards were provided, only the subjective feeling that a good choice has been made is available, this experiment was considered to be more like many everyday choices.

The uniform distribution, which by any standard must be the least active, presumably be the easiest to see and hence be certain of, was, unsurprisingly, found to be the least active and most certain. This was supported by the numbers of correct predictions and squared prediction errors for the uniform distribution. Although it was found to be significantly different from the uniform shape, the randomly changing shape was not found to be significantly different from the other two change distributions. It was, however, rated significantly higher for the Potency/risk factor, that is, it was considered more risky, than the other distributions and was again supported by the objective measure of the lower number of correct predictions and the significantly higher squared prediction error for the random distribution.

The Triangles experiment suggests that people are sensitive to more subtle changes in experience, as predicted by the Somatic Marker hypothesis and that explicit rewards are probably not required in order for our reward structure to be maintained. While in some ways it seems obvious from the Somatic Marker hypothesis (Damasio, 1994) and from other research on affect, such as Zajonc (1980), that subtle changes at a perceptual level must also affect our reward structure it was unclear whether it could be captured based on the present approach.

Chapter 6 sought to test this based on the premise that the somatic marker hypothesis, and hence our reward structure, is created through physiological states and learned associations. Since it is known that the mere perception of colour triggers evaluative processes (Elliot & Maier, 2007), Chapter 6 investigated whether our reward structure will be evident from 'lower level' perceptual information in the form of colour gleaned from limited exposure to scenes, often referred to as gist.

One of the drawbacks with attempting to investigate perceptual information on the basis of colour is that colour is complicated, whether considering colour preferences or colour spaces, nonetheless the ratings of a test image set that were provided by the small number of participants in this experiment produced the expected semantic differential. More problematic, however, was considering what the independent colour variables should be.

This consideration led to the basic colour terms and colours of Berlin and Kay (1969) and also the World Color Survey. Though in some ways controversial, the Berlin and Kay (1969) hypothesis has been tested across cultures and languages through the WorlColor Survey. It should be stated clearly that what is of interest is not whether we can perceive many hues at many levels of luminance, but a range of colour categories or terms that are descriptive of the colours that they represent; the Berlin and Kay (1969) basic colours serve this requirement and an approach based on a finite mixture of Gaussians was selected in order to model the colours.

An initial model was created using the Berlin and Kay (1969)/World Color Survey colours and multiple regression used to investigate the relationship with the semantic differential that was found from the ratings of the images. The analysis showed strong and significant relationships for Evaluation/reward and Potency/cost, but not Activity/uncertainty, suggesting that our reward structure is evident from low level perceptual information.

However, it was felt that, because the Berlin and Kay (1969) colours were externally generated from a set of 'ideal' Munsell colour chips, the initial model did not adequately represent learning from the environment and that a further model should be created that learned its colours rather than them being given. A further model was therefore created where the environment, the 'world' as it were, was limited to a set of everyday scenes. Again, the relationship between the learned colours and the Evaluation/reward and Potency/cost dimensions showed strong and significant relationships. Visual inspection of the colour centres though, showed that while they could be described in terms of the basic colours, they were somewhat dull and did not appear to be what might intuitively be an 'ideal' colour.

Arguably though, because, for example, 'ideal red' would not be the same as 'average red', these centres do not represent what we mean when we point to an ideal representation of the colour. Something else is happening and it is proposed that this is a phenomenon known as peak shift. Applying a peak shift to the model colours produces a representation that is much more like the ideal colours expected.

To summarise, the experimental chapters of the present thesis have addressed three of the areas that were highlighted at the outset these are: *a*) making choices is to a great extent affective and the extent of prior experience is evident for all concepts we know; *b*) in order to maximise advantage, cost and uncertainty should be considered alongside reward; and *c*) the semantic differential may be a reliable representation of that reward structure. It was found from an investigation of a very large data source and from two controlled experiments that three dimensions provided a good representation of experiences, more generally reward, and that this representation is evident from the semantic differential. Further, that this reward structure is also apparent when considering low level perceptual information such as the perception of colour. It is argued that the semantic differential represents or describes the three dimensions of our reward structure and that these dimensions might be better labelled as Reward, Risk/Cost and Uncertainty.

Based on the premise that the Somatic Marker hypothesis provides a Bayesian type prior and that the basis for choices must be inferred, the importance of uncertainty is reiterated and the approach to Bayesian modelling provided in Chapter 7 by way of a brief introduction to the second part of the thesis, which is concerned with mathematical modelling.

11.4 Mathematical model

Expected utility theory as proposed by D. Bernoulli (1738) and axiomatised by Neumann and Morgenstern (1944) has been questioned, particularly as a descriptive theory of choice. A number of alternatives to expected utility theory have been proposed, many of which retain the central idea of D. Bernoulli (1738) in multiplying the probability of an outcome by its value or utility. Probably the most widely known and most influential of the modified theories is Kahneman and Tverskys prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Prospect theory proposes two functions, where the value (or utility) of an outcome is multiplied by a subjective decision weight; both of these functions, with broadly the same characteristics, are retained for the present thesis.

Chapter 8 proposes a value (or utility) function that introduces the novel ideas of competition and fair judges, where utility is the probability (or how certain) that we are to gain an advantage over real or hypothetical competitors.

Based around Bayesian techniques, the value function retains the idea of a personal reference point introduced by Kahneman and Tversky (1979), but, harking back to the very origins of expected utility, that the winning or gaining advantage over the proposed competitors is based on the idea that the result would be as if it were decided by a fair judge. The value function that is constructed takes four parameters, each of which was systematically varied for the purposes of testing the function.

The first parameter, d, has been tentatively labelled difficulty/deliberation. Although the effect seems to be quite modest, larger values for difficulty/deliberation produce a value function that is less risk averse, this may seem strange, but it should be considered that the more step like the function becomes, the more stark, or black and white, it becomes also. The second parameter, t, has been considered to represent 'my' experience in some domain, as with difficulty/deliberation higher values for this parameter suggests an easier, more black and white choice. The third parameter, σ^2, represents the variance or uncertainty in competitors and larger values for this parameter, unsurprisingly, leads to greater risk aversion. The fourth parameter, n, represents the number of competitors with the novel and interesting prediction that with a greater number of competitors a less risk averse function is produced; this suggests that, as the number of competitors grows, the more risk someone might be prepared to bear, nevertheless as with the first two parameters this would have the effect of making the function more step like and consequently a more straightforward choice.

It can be clearly seen that whereas reward relates to utility, the other parameters relate to the risk and uncertainty dimensions of the reward structure that is hypothesised. In addition, the value or utility function that is produced has all of the key features of the empirically derived value function of prospect theory; that is, a function that is risk averse above the reference point, risk seeking below the reference point and loss averse overall. However, because people are observed to behave as if they use a transformed version of a given probability, in prospect theory a further function, the probability weighting function, is used to weight the given probability subjectively. The idea of a probability weighting function is retained by the present thesis, which is the subject of Chapter 9.

The probability weighting function proposed in the present thesis argues that, because uncertainty is ubiquitous, the optimal strategy is to combine probability statements with prior information using Bayes rule. Further, that the prior distribution that is used is the adaptive combination of two classes of prior, an informative empirical prior, that represents previous experience and an uninformative prior of ignorance.

Both classes of prior are commonly used in Bayesian analysis, but, unfortunately, each has potential problems when used alone. Empirical priors rely strongly on the fact that the situation is similar to those encountered before and can be biased if this is not true, while ignorance priors that ignore this previous experience are, potentially, inefficient. The present probability weighting function combines the efficiency of empirical priors with the robustness of ignorance priors using Bayesian model comparison techniques.

The benefits of this approach are illustrated for generic contexts of "good" and "bad" events, based on data gathered from the internet (as described in Chapter 3), and how they can be used to estimate the prior of inference. The resulting model accounts for all the major characteristics of the probability weighting function found in prospect theory.

The combination of priors used in the present probability weighting function also potentially addresses the differences highlighted in the experience based decisions literature that were discussed in Chapter 7. When a situation does not conform to previous experience the present probability weighting function is sensitive to it and rather than slavishly sticking to an empirical prior that is inappropriate, as might be the case if there were only a single prior, a new prior for that situation can be built from the position of ignorance, based on the experience gained. As illustrated in Chapter 9, under weighting of small probabilities that are due potentially to under sampling and information asymmetry, familiar from the experience based decisions literature (e.g. L. Hadar & Fox, 2009), that would be the result of building a new prior of inference in that situation.

The value function and probability weighting function, together forming a complete model, are tested against known problem cases for decision making in Chapter 10. The main classes of challenge to decision making models identified in the decision making literature (e.g. Busemeyer & Johnson, 2008) are common consequence and common ratio effects, violations of stochastic dominance, preference reversals and context dependent effects.

In testing the model, it is difficult to identify a particular parameterisation as correct and meaningful statistics on this basis are also difficult, however, the results can be considered in terms of parsimony i.e. the lowest parameter sets generating the required preferences and in terms of typical values i.e. the mean values of the parameters generating the required preferences. In all cases, however, it was found that the pattern of results that were produced by the model were similar to those found in cited experiments but does so based on the maximisation of an expected utility, perhaps better termed maximisation of potential advantage in the present case.

11.5 Implications and further research

As a result of the research contained in the present thesis a number of plausible predictions are made that can be followed up in further research, in addition, some of the approaches used for data collection were interesting. For example, using internet data sources, as discussed above, has been useful to the present thesis and bears further use and investigation.

Use of these sources is an important approach if it can collect data that are representative of behaviour "in the raw", as it were, rather than as it might occur in a psychology lab where, in certain types of experiments, people have the opportunity to modify their behaviour according to what they think might be required or have their behaviour artificially limited by dichotomous options such as in the moral dilemmas. Fox, Rogers, and Tversky (1996), for example, collected data from participants who were professional options traders in order to investigate the probability weighting function and Stewart, Chater, and Brown (2006) used real banking data for the Decisions By Sampling model. Furthermore, while not directly concerned with choices, the mixture model approach, that was used to investigate colour, warrants further attention as a route for exploring the Berlin and Kay (1969) basic colour term hypothesis.

Intuitively, considering the potential risk and uncertainty associated with obtaining a reward is important when making choices, a point that is also made by in Rushworth and Behrens (2008), however, these variables are rarely considered in the reinforcement learning and decision making literature. The present thesis considers them as central to making choices in terms of a multi dimensional reward structure and investigates this structure in a series of experiments. As a result it is argued that the three dimensions of reward, risk/cost and uncertainty should be considered in research involving choices.

In addition, a long standing and well researched instrument, the semantic differential, has been identified as a potential representation (or measure or description) of the richer reward structure. While this will need further investigation as an instrument for use with choices, a potential answer has been provided for a long standing theoretical question for the semantic differential, which is what is it really representing.

The present thesis treats utility differently to the way that it has been treated in other theories, that is, utility is considered as potential advantage, rather than potential gain, which is perhaps more consistent with how people actually operate in a social economy. When advantage is considered in terms of a social economy it gives rise to an important and novel idea; the introduction of idea of competitors into the value function.

However, because competitors are assumed to be randomly sampled and a particular choice will be determined by the competitors that are sampled, peoples choices will be changeable dependent on the samples that are actually taken. The potential for priming (e.g. Bargh & Chartrand, 1999; Duckworth et al., 2002) and other memory type effects (e.g. Kahneman, Slovic, & Tversky, 1982; Furnham, 1989) to bias choices would therefore seem to be clear; indeed, a similar view is expressed about the Decision By Sampling (DbS) model (Stewart et al., 2006). The probability weighting function on the other hand, built as it is on a combination of priors, offers a potential explanation for the differences that have been seen in the experience based decisions literature. This combination of priors approach bears further research, especially in terms of other data sources and specific choice situations.

In conclusion, the present thesis has two strong advantages. First, rather than simply being an arbitrary function fitted to data, all the parameters have psychological meaning. This means that the theory makes extensive and plausible predictions: The number of competitors, how certain a person is of those competitors; the level of certainty expressed in a probability (or the way it is signalled), the level of trust in a probability (or the way it was signalled); the amount and nature of previous experience a person has of similar situations; cues that this situation is different to those in the past and others, are predicted to have affects on peoples choices. These variables are also very likely to vary between people, leading to between subject differences in choices.

Second, the present thesis also sheds light on another issue with descriptive theories that maintain the expected utility hypothesis as the normative model. If these models are correct, on the assumption that expected utility theory would potentially be simple to implement and that evolution has had millennia to optimise our decision making machinery, it is unclear why we have not been out competed by animals that do employ expected utility maximisation. It has been shown that the observed biases are compatible with expected utility theory; it is simply that the expectation is carried out over the uncertainty. Taking previous knowledge into account (when appropriate) is not irrational, and not inconsistent with expected utility theory, simply the uncertainty free version that is normally explored.

This last section has attempted to pinpoint some of the findings that are novel and have potentially advanced the investigation of how people go about making choices in their everyday lives. However, it is a continuing challenge to attempt to account for all of the variations in behaviour and for the present thesis, choice behaviour. To paraphrase Oaksford, Chater, and Stewart (2012), human cognition is striking in its ability to handle, even to a modest extent, making choices in novel, hypothetical, verbally stated and real scenarios for which our past experience and evolutionary history may have provided only minimal preparation; there is a long way to go before we will be able to claim an acceptable understanding of it.

References

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica: Journal of the Econometric Society*, 503–546.

Baddeley, R. J., Osorio, D., & Jones, C. D. (2007). Generalization of Color by Chickens: Experimental Observations and a Bayesian Model. *The American Naturalist*, *169*(s1), S27–S41.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*(7), 462–479.

Bargh, J. A., Gollwitzer, P. M., Lee-chai, A., Barndollar, K., & Trotschel, R. (2001). The Automated Will : Nonconscious Activation and Pursuit of Behavioral Goals. *Journal of Personality and Social Psychology*, *81*(6), 1014–1027.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233.

Baskerville, K. A., Schweitzer, J. B., & Herron, P. (1997). Effects of cholinergic depletion on experience-dependent plasticity in the cortex of the rat. *Neuroscience*, *80*(4), 1159–1169.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*(3), 412–414.

Bawa, V. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, *2*(1), 95–121.

Bechara, A. (2000). Emotion, Decision Making and the Orbitofrontal Cortex.

Cerebral Cortex, *10*(3), 295–307.

Bechara, A., & Damasio, A. R. (2005). The somatic marker hypothesis: A neural theory of economic decision. *Games and Economic Behavior*, *52*(2), 336–372.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7–15.

Bechara, A., Tranel, D., Damasio, H., & Damasio, A. R. (1996). Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cerebral Cortex*, *6*(2), 215–225. Available from `http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8670652`

Berlin, B., & Kay, P. (1969). *Basic color terms.* Univ. of California Press.

Bernoulli, D. (1738). *Specimen theoriae novae de mensura sortis (Exposition of a New Theory on the Measurement of Risk).*

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, *22*(1), 23–36.

Bernoulli, J. (1713). *Ars conjectandi.* Basel, Switzerland.

Binmore, K. (2007). Making Decisions in Large Worlds. *Annales dEconomie Statistique*, *86*(4), 1–19.

Birnbaum, M., & Zimmermann, J. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes*, *74*(2), 145–187.

Blume, L. E., & Easley, D. (2008). Rationality. In S. N. Durlauf & L. E. Blume

(Eds.), *The new palgrave dictionary of economics* (2nd ed., pp. 884–893). Basingstoke: Nature Publishing Group.

Boucher, J., & Osgood, C. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *8*(1), 1–8.

Boureau, Y. L., & Dayan, P. (2010). Opponency Revisited: Competition and Cooperation Between Dopamine and Serotonin. *Neuropsychopharmacology*.

Boynton, R. M., & Olson, C. X. (1990). Salience of chromatic basic color terms confirmed by three measures. *Vision Research*, *30*(9), 1311–1317.

Busemeyer, J. R., & Johnson, J. G. (2008). Micro-Process Models of Decision-Making. In R. Sun (Ed.), *Cambridge handbook of computational cognitive modeling* (pp. 302–321). Cambridge University Press.

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., & Poggio, T. (2007). A model of V4 shape selectivity and invariance. *Journal of Neurophysiology*, *98*(3), 1733.

Cook, R., Kay, P., & Regier, T. (2011). *WCS Data Archives.* Available from `http://www.icsi.berkeley.edu/wcs/data.html`

Cools, R., & Nakamura, K. (2010). Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology*.

Cornelissen, F. W., Brenner, E., & Smeets, J. (2003). True color only exists in the eye of the observer. *Behavioral and Brain Sciences*, *26*(01), 26–27.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*(1), 1–73.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature reviews. Neuroscience*, *3*(8), 655–66.

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology*, *13*(4), 500–505.

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature neuroscience*, *7*(2), 189–95.

Crockett, M. J., Clark, L., & Robbins, T. W. (2009). Reconciling the Role of Serotonin in Behavioral Inhibition and Aversion: Acute Tryptophan Depletion Abolishes Punishment-Induced Inhibition in Humans. *Journal of Neuroscience*, *29*(38), 11993.

Dalton, P., Maute, C., Oshida, A., Hikichi, S., & Izumi, Y. (2008). The Use Of Semantic Differential Scaling To Define The Multidimensional Representation Of Odors. *Journal of Sensory Studies*, *23*, 485–497.

Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain.* Putnam New York.

Damasio, A. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex [and discussion]. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413–1420.

Damasio, A. (1998). Emotion in the perspective of an integrated nervous system. *Brain Research Reviews*, *26*(2-3), 83–86.

Damasio, A. (2000). *The Feeling Of What Happens: Body, Emotion and the Making of Consciousness.* Vintage.

Damasio, A., Grabowski, T., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., et al. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature neuroscience*, *3*(10), 1049–56.

Davidson, R. R. J., & Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, *3*(1), 11–21.

Davis, J. B., Hands, D. W., & Maki, U. (1998). *The Handbook of Economic Methodology.* Edward Elgar Publishing.

Dayan, P., & Yu, A. J. (2002). ACh, uncertainty, and cortical inference. *Advances in neural information processing systems*, *1*, 189–196.

Dayan, P., & Yu, A. J. (2006). Phasic norepinephrine: A neural interrupt signal for unexpected events. *Network: Computation in Neural Systems*, *17*(4), 335–350.

Detenber, B., & Reeves, B. (1996). A Bio-Informational Theory of Emotion: Motion and Image Size Effects on Viewers. *Journal of Communication*, *46*(3), 66–84.

DiVesta, F. J. (1966). A developmental model of the semantic structures of children. *Journal of Verbal Learning and Verbal Behavior*(5), 249–259.

Duckworth, K., Bargh, J., Garcia, M., & Chaiken, S. (2002). The automatic evaluation of novel stimuli. *Psychological Science*, *13*(6), 513.

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: a critical evaluation. *Neuroscience & Biobehavioral Reviews*, *30*(2), 239–271. Available from `http://www.ncbi.nlm.nih.gov/pubmed/16197997`

Elkan, C. (1997). Boosting and naive Bayesian learning. *proceeding of KDD97 New Port beach CA.*

Elliot, A. J., & Covington, M. V. (2001). Approach and avoidance motivation. *Educational Psychology Review*, *13*(2), 73–92.

Elliot, A. J., & Maier, M. A. (2007). Color and Psychological Functioning.

Current Directions in Psychological Science, *16*(5), 250–254.

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: the effect of red on performance attainment. *Journal of experimental psychology General*, *136*(1), 154–168.

Elliot, A. J., & Niesta, D. (2008). Romantic red: red enhances men's attraction to women. *Journal of Personality and Social Psychology*, *95*(5), 1150–1164.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision*, *7*(1), 10.

Fox, C. R., & Hadar, L. (2006). Decisions from experience= sampling error+ prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making*, *1*(2), 159–161.

Fox, C. R., Rogers, B. A., & Tversky, A. (1996). Options traders exhibit subadditive decision weights. *Journal of Risk and Uncertainty*, *13*(1), 5–17.

Francis, C., & Heise, D. (2006). *Mean Affective Ratings of 1,500 Concepts by Indiana University Undergraduates in 2002-3 [Computer file].* Available from `http://www.indiana.edu/~socpsy/ACT/interact/JavaInteract.html%3E`

Furnham, A. (1989). The robustness of the recency effect: Studies using legal evidence. *The Journal of General Psychology*, *113*(4), 351–357.

Gilboa, I. (2010, September). Questions in Decision Theory. *Annual Review of Economics*, *2*(1), 1–19.

Goldie, P., & Spicer, F. (2002). Introduction. In P. Goldie (Ed.), *Understanding emotions: Mind and morals.* Ashgate Pub Ltd.

Greene, J. (2002). *The Terrible, Horrible, No Good, Very Bad Truth about Morality and What to Do About it.* Unpublished doctoral dissertation, Princeton

University.

Greene, J. (2007). *The secret joke of Kants soul.* MIT Press, Cambridge, MA.

Greene, J., Nystrom, L., Engell, A., Darley, J., & Cohen, J. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389–400.

Greene, J., Sommerville, R., Nystrom, L., & Darley, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*(5537), 2105.

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

Griffiths, T., Kemp, C., & Tenenbaum, J. (2008). Bayesian models of cognition. *Cambridge handbook of computational cognitive modeling*.

Gu, Q. (2002). Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience*, *111*(4), 815–835.

Guadagnoli, E., & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265.

Hacking, I. (2001). *An Introduction to Probability and Inductive Logic* (Vol. 24) (No. 2). Cambridge University Press.

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *The American Economic Review*, *59*(1), 25–34.

Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*(4), 317–325.

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: a critique of some recent interpretations of experimental games. *Theoretical*

Population Biology, *69*(3), 339–48.

Hagtvedt, H., Hagtvedt, R., & Patrick, V. M. (2008). The Perception and Evaluation of Visual Art. *Empirical Studies of the Arts*, *26*(2), 197–218.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814.

Haidt, J., & Bjorklund, F. (2007). Social Intuitionists Answer Six Questions About Morality. In Sinnott-Armstrong W. (Ed.), *Moral psychology.* Oxford University Press.

Haley, K., & Fessler, D. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*(3), 245–256.

Hauser, M. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong.* Ecco Press.

Hays, D. G., Margolis, E., Naroll, R., & Perkins, D. R. (1972). Color Term Salience. *American Anthropologist*, *74*(5), 1107–1121.

Heise, D. (1969). Some methodological issues in semantic differential research. *Psychological Bulletin*, *72*(6), 406–422.

Heise, D. (2001). Project Magellan: Collecting Cross-cultural Affective Meanings Via The Internet. *Electronic Journal of Sociology*, *5*.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (2004). *Foundations of Human Sociality* (Vol. 1) (No. September). Oxford University Press.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Mcelreath, R., et al. (2005). Economic Man in Cross-cultural Perspective : Behavioral Experiments in 15 Small-scale Societies. *Social Dynamics*, *28*(6), 795–815.

Hering, E. (1964). *Outlines of a theory of the light sense* (No. 1905). Harvard University Press.

Herstein, I. N., & Milnor, J. (1953). An Axiomatic Approach to Measurable Utility. *Econometrica*, *21*(2), 291–297.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534.

Holland, A. C., & Kensinger, E. A. (2010). Emotion and autobiographical memory. *Physics of Life Reviews*, *7*(1), 88–131.

Hume, D. (1739). *Treatise of Human Nature.*

Hume, D. (1748). *An Enquiry Concerning Human Understanding.*

Humphrey, N. (1983). The colour currency of nature. In T. Porter & B. Mikellides (Eds.), *Consciousness regained chapters in the development of mind* (pp. 144–152). London: Studio-Vista.

Hurly, T., & Oseen, M. D. (1999). Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. *Animal Behaviour*, *58*, 59–66.

Hurvich, L. M., & Jameson, D. (1957). An opponent process theory of color vision. *Psychological Review*, *64*, 384–404.

Hurwicz, L. (1945). The theory of economic behavior. *The American Economic Review*, *35*(5), 909–925.

Huygens, C. (1657). *De ratiociniis in ludo aleae* (F. van Schooten, Trans.).

Huys, Q. J., & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328.

Jacobs, G. H. (1981). *Comparative Color Vision (Academic Press Series in Cognition & Perception).* Academic Pr.

James, W. (1894). The physical basis of emotion. *Psychological Review*, *1*(5), 516–529.

Jaynes, E. T. (1995). *Probability theory: the logic of science.*

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under Uncertainty: Heuristics and Biases.* Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kilgard, M. P., & Merzenich, M. M. (1998). Cortical map reorganization enabled by nucleus basalis activity. *Science*, *279*(5357), 1714–1718.

Kim, N. H., & Kang, H. H. (2009). The aesthetic evaluation of coastal landscape. *KSCE Journal of Civil Engineering*, *13*(2), 65–74.

Knight, F. (1921). *Risk, Uncertainty and Profit* (Vol. XXXI). Houghton Mifflin Company.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nature Reviews Neuroscience*, *6*(9), 691–702.

Kruschke, J. (2005). Category Learning. In K. Lamberts & R. L. Goldstone (Eds.), *The handbook of cognition* (pp. 183—-201). London: Sage.

Lange. (1885). The mechanism of the emotions. *The Emotions*, 672–685.

LeDoux, J. (1993). Emotional memory systems in the brain. *Behavioural Brain Research*, *58*(1-2), 69–79. Available from `http://www.ncbi.nlm.nih.gov/pubmed/8136051`

LeDoux, J. (1996). *The Emotional Brain* (Vol. 256; J. M. Jenkins, K. Oatley, & N. L. Stein, Eds.). Simon & Schuster. Available from `http://www.amazon.com/dp/0684836599`

Lewis, D. D. (1997). Reuters-21578 text categorization test collection. Distribution 1.0. 1997. *URL www. daviddlewis. com/resources/testcollections*.

Lindsey, D. T., & Brown, A. M. (2006). Universality of color names. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(44), 16608–13.

Machina, M. (2006). Choice under uncertainty. *Encyclopedia of Cognitive Science*.

MacKay, D. (2003). *Information theory, inference, and learning algorithms.* Cambridge Univ Pr.

Marschak, J. (1950). Rational Behavior, Uncertain Prospects, and Measurable Utility. *Econometrica*, *18*(2), 111–141.

MATLAB. (2008). *version 7.6.0.324 (R2008a).* Natick, Massachusetts: The MathWorks Inc.

Matlin, M., & Stang, D. (1978). *The Pollyanna principle: Selectivity in language, memory, and thought.* Schenkman Cambridge, MA.

McManus, I. (1983). Basic colour terms in literature. *Language and Speech*, *26*, 247—-252.

McManus, I. (1997). Note: half-a-million basic colour words: Berlin and Kay and the usage of colour words in literature and science. *Perception*, *26*(3), 367–370.

McManus, I., Jones, A. L., & Cottrell, J. (1982). The aesthetics of colour. *Perception*, *10*(6), 651–666.

Mehta, R., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, *323*(5918), 1226–1229.

Meltzoff, A. (2005). Imitation and other minds: the like mehypothesis. *Perspec-*

tives on imitation: Imitation, human development, and culture, 55e77.

Meltzoff, A., & Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*(4312), 75.

Meltzoff, A., & Moore, M. (1983). Newborn infants imitate adult facial gestures. *Child development*, 702–709.

Meltzoff, A. N. (2007). The like me framework for recognizing and becoming an intentional agent. *Acta psychologica*, *124*(1), 26–43.

Miller, G. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Miron, M. (1969). What is it that is being differentiated by the semantic differential?

Mlodinow, L. (2008). *The Drunkard's Walk: How Randomness Rules Our Lives.* Pantheon.

Mollon, J. D. (1989). Tho she kneeld in that place where they grew. The uses and origins of primate colour vision. *Journal of Experimental Biology*, *146*, 21–38.

Mongin, P. (1997). Expected Utility Theory. In J. Davis, W. Hands, & U. Maki (Eds.), *Handbook of economic methodology* (pp. 342–350). London: Edward Elgar Publishing.

Montmort, P. R. de. (1713). Essay d'analyse sur les jeux de hazard. Seconde edition. *Revue \& augment{\'e}e de plusieurs Lettres.*.

Morris, W. E. (2011). *David Hume.*

Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. *The Journal of Political Economy*, *59*, 371–404.

Nabney, I., & Bishop, C. (2004). *Netlab Neural Network Software.*

Nannen, V. (2003). A Short Introduction to Model Selection , Kolmogorov Complexity and Minimum Description Length (MDL). *Complexity*, *48*(Mdl), 1–23.

Naqvi, N., Shiv, B., & Bechara, A. (2006). The Role of Emotion in Decision Making: A Cognitive Neuroscience Perspective. *Current Directions in Psychological Science*, *15*(5), 260–264.

Neumann, J. von, & Morgenstern, O. (1944). *Theory of games and economic behavior.* Princeton university press Princeton, NJ.

Newell, B. R., & Rakow, T. (2007). The role of experience in decisions from description. *Psychonomic bulletin & review*, *14*(6), 1133.

Nichols, S. (2004). *Sentimental rules: on the natural foundations of moral judgment.* Oxford University Press.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, *84*(3), 231. Available from http://psycnet.apa.org/journals/rev/84/3/231/

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning.* Oxford University Press, USA.

Oaksford, M., Chater, N., & Stewart, N. (2009). Reasoning and decision making. In K. Frankish & W. Ramsey (Eds.), *The cambridge handbook of cognitive science.* Cambridge: Cambridge University Press.

Oaksford, M., Chater, N., & Stewart, N. (2012). Reasoning and decision making. In K. Frankish, W. Ramsey, & C. O'Callaghan (Eds.), *The cambridge handbook of cognitive science.* Cambridge University Press.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). *Abstract reward and punishment representations in the human orbitofrontal* *cortex.* (Vol. 4; Tech. Rep. No. 1). Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principal components analysis. *Practical Assessment, Research & Evaluation*, *9*(11), 8.

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, *49*(3), 197–237.

Osgood, C. E. (1969). On the whys and wherefores of E, P, and A. *Journal of Personality and Social Psychology*, *12*(3), 194–199.

Osgood, C. E., May, W. H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning.* University of Illinois Press.

Osgood, C. E., Suci, G., & Tannenbaum, P. (1957). The measurement of meaning. *Urbana: University of Illinois Press*.

Osgood, C. E., & Suci, G. J. (1955). Factor analysis of meaning. *Journal of experimental psychology*, *50*(5), 325–338.

Pascal, B. (1665). *Traité du triangle arithmétique (Treatise on the arithmetic triangle).*

Pizarro, D. (2000, December). Nothing More than Feelings? The Role of Emotions in Moral Judgment. *Journal for the Theory of Social Behaviour*, *30*(4), 355–375.

Polasek, W. (2000). The Bernoullis and the origin of probability theory: Looking back after 300 years. *Resonance*, *5*(8), 26–42.

Potter, M. C. (1975). Meaning in visual search. *Science*, *187*(4180), 965–966.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of experimental psychology Human learning and memory*, *2*(5), 509–522.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 497–527.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9*(1), 29–43.

Quine, W. V. (1970). Natural kinds. In R. Boyd, P. Gasper, & J. D. Trout (Eds.), *Metaphysics an anthology* (pp. 41–56). Open Court.

Ramachandran, V. S., & Hirstein, W. (1999). The Science of Art. *Journal of Consciousness Studies*, *6*(6), 15–35.

Ramsey, F. P. (1926). Truth and probability. In R. B. Braithwaite (Ed.), *Foundations of mathematics and other logical essays* (pp. 156–198). London: Kegan, Paul, Trench, Trubner & Co. Ltd.

Rescorla, R., & Wagner, A. (1972). Variations in the Effectiveness of Reinforcement and Nonreinforcement. *New York: Classical Conditioning II: Current Research and Theory, Appleton-Century-Crofts*.

Robson, A. J. (2001). The biological basis of economic behavior. *Journal of Economic Literature*, 11–33.

Rolls, E. T. (2006). Brain mechanisms of emotion and decision-making. *International Congress Series*, *1291*, 3–13.

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the gist of real-world natural scenes. *Visual Cognition*, *12*(6), 852–877.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature neuroscience*, *11*(4), 389–97.

Saunders, B. A. C., & Brakel, J. van. (1997). Are there non-trivial constraints on colour categorisation. *Behavioural and Brain Sciences*, *20*, 167–228.

Saunders, B. A. C., & Brakel, J. van. (2002). The Trajectory of Color. *Perspectives on Science*, *10*(3), 302–355.

Savage, L. J. (1972). *The Foundations of Statistics (2nd revised edition).* . New York: Dover Publications.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of neurophysiology*, *80*(1), 1.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions*, *6*(1), 24.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593.

Schwarz, N. (2010). Feelings-as-Information Theory. *chapter*, *In P. Van*(January), 1–32.

Schwarz, N., & Clore, G. (2007). Feelings and Phenomenal Experiences. *Social psychology Handbook of basic principles*, *433*, 385–407.

Shenhav, A., & Greene, J. D. (2010). Moral Judgments Recruit Domain-General Valuation Mechanisms to Integrate Representations of Probability and Magnitude. *Neuron*, *67*(4), 667–677.

Simon, H. (1979). Rational decision making in business organizations. *The American economic review*, *69*(4), 493–513.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of consumer research*, 158–174.

Sinnott-Armstrong, W. (2011). *Consequentialism.*

Snider, J. G., & Osgood, C. E. (1969). *Semantic differential technique; a sourcebook.* Aldine Pub. Co.

Soubrié, P. (1986). Reconciling the role of serotonin neurons in human and animal behavior. *Behavioral and Brain Sciences*, *9*, 319–364.

Springbett, B. (1960). The Semantic Differential and Meaning in Non-Objective

Art. *Perceptual and Motor Skills*, *10*, 231—-240.

Squire, L. R. (1992). Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243.

Stanovich, K. E., & West, R. F. (2001). Advancing the rationality debate. *Behavioral and Brain Sciences*, *23*(05), 701–717.

Stewart, N., Chater, N., & Brown, G. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1–26.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* The MIT press.

Taft, C., & Sivik, L. (1997). Salient color terms in four languages. *Scandinavian Journal of Psychology*, *38*, 29–34.

Takase, L. F., Nogueira, M. I., Baratta, M., Bland, S. T., Watkins, L. R., Maier, S. F., et al. (2004). Inescapable shock activates serotonergic neurons in all raphe nuclei of rat. *Behavioural Brain Research*, *153*(1), 233–239.

Teghtsoonian, R. (1971). On the exponents in Stevens' law and the constant in Ekman's law. *Psychological Review*, *78*(1), 71–80.

Teira, D. (2006). On the normative dimension of the St. Petersburg paradox. *Studies In History and Philosophy of Science Part A*, *37*(2), 210—-223.

Thomas, D. R., Mood, K., Morrison, S., & Wiertelak, E. (1991). Peak shift revisited: a test of alternative interpretations. *Journal of experimental psychology Animal behavior processes*, *17*(2), 130–140.

Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*(5715), 1642.

Tracer, D. P. (2003). Selfishness and Fairness in Economic and Evolutionary Perspective: An Experimental Economic Study in Papua New Guinea. *Current Anthropology*, *44*(3), 432–438.

Tuyl, F., Gerlach, R., & Mengersen, K. (2008). A Comparison of Bayes-Laplace, Jeffreys, and Other Priors. *The American Statistician*, *62*(1), 40–44.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293—-315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, 251–278.

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

Tversky, A., & Sattath, S. (1979). Preference trees. *Psychological Review*, *86*(6), 542–573.

Tversky, A., & Simonson, I. (1993). Context-Dependent Preferences. *Management Science*, *39*(10), 1179–1189.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are Probabilities Overweighted or Underweighted When Rare Outcomes Are Experienced (Rarely)? *Psychological Science*, *20*(4), 473–479.

Walton, M. E., Devlin, J. T., & Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*(11), 1259–1265.

Weinreich, U. (1958). Travels through semantic space. *Word*, *14*, 346–366.

Wills, S., & Mackintosh, N. J. (1998). Peak Shift on an Artificial Dimension. *The Quarterly Journal Of Experimental Psychology*, *51B*(1), 1–32.

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: verbal versus numeric methods. *Journal of Experimental Psychology Applied*, *2*, 343–364.

Yaqub, M. Z., Saz, G., & Hussain, D. (2009). A Meta Analysis of the Empirical Evidence on Expected Utility Theory. *European Journal of Economics Finance and Administrative Sciences*, *15*, 117–133.

Yu, A. J. (2003). Expected and Unexpected Uncertainty : ACh and NE in the Neocortex. *Analysis*, *15*, 173.

Yu, A. J., & Dayan, P. (2002). Acetylcholine in cortical inference. *Neural Networks*, *15*(4-6), 719–730.

Yu, A. J., & Dayan, P. (2003). Expected and Unexpected Uncertainty: ACh & NE in the Neocortex. *Advances in Neural Information Processing Systems 15*, *15*, 157–164.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inferences. *American psychologist*, *35*(2), 151.

Appendix A

Expected Utility Axioms

A von Neumann and Morgenstern type expected utility function that ranks lotteries or probability distributions according to an individuals preference ($\succeq$) is a consequence of four axioms: completeness, transitivity, continuity, and independence. These axioms are illustrated as follows (Machina, 2006):

Completeness

For all lotteries X and Y:

either $X \succeq Y$ or $Y \succeq X$ or both.

Transitivity

For all lotteries X, Y and Z

if $X \succeq Y$ and $Y \succeq Z$ then $X \succeq Z$.

Continuity

For all lotteries X, Y and Z

if $X \succeq Y$ and $Y \succeq Z$ then for some $\alpha \in (0, 1)$

Y is indifferent to $\alpha X + (1 - \alpha)Z$.

Independence

For all lotteries X, Y and Z and all $\alpha \in (0,1)$

if $X \succeq Y$ then $\alpha X + (1-\alpha)Z \succeq \alpha Y + (1-\alpha)Z$.

Appendix B

The St Petersberg prospect

Bernoulli's (1738) description of the St. Petersburg prospect:

> "Peter tosses a coin and continues to do so until it should land 'heads' when it comes to the ground. He agrees to give Paul one ducat if he gets 'heads' on the very first throw, two ducats if he gets it on the second, four if on the third, eight if on the fourth, and so on, so that with each additional throw the number of ducats he must pay is doubled. Suppose we seek to determine the value of Paul's expectation".

In other words, what is of interest is how much (or the most) a rational person would pay to enter the game. In this game, as defined above, the winnings are $x_i = 2^i$, where i is the number of times that the coin is tossed, with probability $p_i = \left(\frac{1}{2}\right)^i$; so based on expected value, E_V:

$$
\begin{aligned}
E_V &= \sum_{i=1}^{\infty} p_i x_i \\
&= \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i 2^i \\
&= 1+1+1+1+1+1+... \\
&= \infty
\end{aligned} \tag{B.1}
$$

The paradox, clearly, is that (based on the expected value) no rational person would pay such an amount to take part in the game.

Appendix C

Technorati concept list

Table C.1: Technorati concept words

abandon	bachelor	calm	compulsive	demean	entertain
abortionist	back	camp	conceited	demote	enthusiastic
abuse	badger	campground	concert	denigrate	entreat
accommodating	baldy	campus	condemn	denounce	entrepreneur
accuse	banquet	capitalist	confidant	dependable	envious
address	baptize	capture	confident	dependent	escape
admonish	bar	car	confine	deprecate	euphoric
adolescent	bash	careless	confront	deride	evangelist
adult	bathe	caress	congratulate	despondent	examine
adulterer	battlefield	carnival	conscientious	detective	exasperated
adulteress	beach	casino	conservative	devil	excited
adventurous	beckonto	catch	considerate	direct	excuse
advise	bed	cathedral	console	disappointed	execute
advisor	beg	Catholic	consultant	disapproving	execution
affectionate	belittle	caution	contemplate	disciplinarian	executioner
afraid	berate	cautious	contemptuous	discipline	executive
aggravated	beseech	celebration	contented	disco	exploit
aggressive	bill	celebrity	contradict	discontented	extrovert
agitated	bind	cemetery	contrite	discourage	extroverted
agnostic	bite	challenge	convalescent	discouraged	eye
aide	bitter	chapel	convict	disgusted	failure
airplane	Black	charmed	cooperative	disheartened	father
alarmed	blame	chase	correct	dismayed	fearful
alcoholic	bless	chastise	counsel	disobey	felon
aloof	blonde	chatterbox	courageous	disparage	female
alumnus	blue	cheer	courtroom	displeased	female
ambitious	bootlick	cheerful	cousin	disrobe	feminine
amuse	bootlicker	cheerless	cowardly	dissatisfied	feminist
analyze	boozer	chide	criminal	dissuade	festival
angry	bore	child	critic	distract	fight
anguished	boss	childish	criticize	distressed	fight
annoyed	bossy	choke	crook	divorce	fine
antiSemite	bouncer	Christmas	crowd	doctor	finicky
antisocial	boy	church	crushed	dogmatic	firstborn
anxious	brat	citizen	cue	domineering	flatter
apathetic	brave	classmate	curse	downhearted	flee
applaud	bribe	classroom	cuss	dress	flirt
applicant	bride	clergyman	customer	drinkto	flophouse
apprehend	bridesmaid	client	cynical	dropout	flunk
apprehensive	brief	clinic	damn	drunk	flunky
apprentice	bright	club	dance	dummy	flustered
approach	brothel	coach	dare	dyke	foe
arrest	brother	coach	daring	eager	follower
arrogant	browbeat	cocky	date	earnest	foolish
ashamed	brunette	coddle	daughter	easygoing	forget
assail	brutalize	coed	debate	ecstatic	forgive
assist	brute	coerce	debrief	egotistical	forgiving
assistant	buddy	cold	defeat	elated	freeloader
athlete	bully	colleague	defend	elbow	friend
attendto	bully	combat	defendant	elder	frightened
attorney	bum	comfort	defensive	embarrassed	frustrated
auction	bureaucrat	command	defiant	embrace	funeral
aunt	businessman	compassionate	deflated	employee	furious
authoritarian	businesswoman	compensate	defy	employer	fussover
authority	cafe	competent	degrade	encourage	gangster
awestruck	cafeteria	competitive	dejected	enemy	gay
baby	cajole	competitor	delinquent	enraged	genius

gentle	hero	insecure	lecture	mouthpiece	parent
gentleman	heroine	insensitive	lecturer	moved	parody
ghetto	heterosexual	insider	lesbian	mug	parolee
gigolo	hideout	insincere	liar	murder	partner
girl	hire	inspect	liberal	murderer	passerby
girlfriend	hit	instruct	librarian	murderess	passionate
glad	hold	instructor	library	museum	pastor
gleeful	home	insult	loafer	nag	patient
gloomy	homemaker	intelligent	lobbyist	naive	patient
glum	homesick	intern	lonely	nark	patriot
God	homosexual	interrogate	lonesome	narrowminded	patrolman
goofoff	honeymoon	interrogation	loser	needle	pauper
gossip	honeymooner	interrupt	lovesick	neglect	peaceful
grab	hoodlum	interview	lunatic	negotiator	peacetime
grade	hooker	interview	luncheon	neighbor	pediatrician
graduate	horny	interviewee	luncheonette	nephew	peeved
graduation	horrified	interviewer	lunchroom	nervous	penalize
grandchild	hostess	intimate	lustful	nestle	penitentiary
granddaughter	hostile	intolerant	mad	newlywed	perceptive
grandfather	hothead	introspective	malcontent	niece	persistent
grandmother	hotheaded	introvert	male	nobody	pessimist
grandparent	hotshot	introverted	male	nonsmoker	pessimistic
grandson	hound	invalid	malign	nostalgic	pest
grasp	houseguest	irked	malingerer	novice	pester
greedy	housewife	irritable	man	nudge	pet
greet	hug	irritated	manager	nurse	petrified
grind	humble	jealous	manageress	nut	petty
groom	humiliated	jerk	masculine	nuzzle	photograph
grouchy	hunk	Jew	masochist	nymphomaniac	physician
grownup	hurry	jock	massage	obedient	pickpocket
guest	husband	joggle	matriarch	obey	pickup
guide	hush	josh	mature	obstruct	pimp
gullible	hussy	jostle	mealtime	office	pizzeria
gunfight	idealistic	joyful	medicate	ogle	placid
gunman	idiot	joyless	meek	old	plainclothesman
guy	ignoramus	jubilant	meeting	opportunist	playful
gym	imaginative	judge	melancholy	oppose	playground
gynecologist	imitate	juror	mentalcase	optimistic	playmate
gyp	immature	kick	merchant	order	poke
hail	immoral	kid	merry	organizer	policeman
Halloween	impatient	kid	millionaire	orgy	politician
halt	implore	kind	mimic	orphan	pompous
handcuff	imprison	kiss	minister	outgoing	poor
handicapped	incarcerate	kitchen	mischievous	outlaw	poorhouse
handyman	inconsiderate	klutz	miser	outraged	popular
harangue	incriminate	knife	miserly	outspoken	pornographer
harass	indecisive	laboratory	mistress	overcharge	praise
hardworking	independent	laborer	mob	overjoyed	preacher
harm	indignant	lackey	mock	overpower	prejudiced
hassle	industrious	lady	modest	overwhelm	priest
hatemonger	infant	lawyer	molest	overwhelmed	priestess
heal	infatuated	lazy	monitor	overwork	principal
healer	inform	lead	moron	pagan	prison
Heaven	informer	leader	mortified	pal	probationer
heckle	inhibited	leave	mother	pamper	prod
hell	injure	lecher	mourner	panicked	professor
helper	innocent	lecture	mournful	paranoid	prompt

prosecute	rob	sincere	supervisor	undergraduate	youth
prostitute	robber	sinner	supporter	understanding	zoo
Protestant	roommate	sister	surgeon	uneasy	
protester	rouse	slap	surprise	unfair	
proud	ruthless	slaughterhouse	suspicious	unfriendly	
psychiatrist	sad	slug	sweatshop	unimaginative	
psychopath	sadist	slut	sweetheart	unpopular	
psychotic	sadistic	sly	swinger	unreliable	
punch	saint	smoker	sympathetic	upset	
punish	saintly	smug	tactful	uptight	
punk	salesclerk	snuggle	taxi	vacationer	
pupil	saleslady	sock	taxpayer	vain	
purchaser	salesman	son	teach	vengeful	
push	sarcastic	soothe	teacher	victim	
quack	satisfied	sorrowful	teammate	victimize	
quarrelsome	sauna	sorry	tease	vigilante	
queer	sawbones	spank	teenager	village	
question	scared	spendthrift	temperamental	villain	
questioner	schizophrenic	spinster	tenant	violent	
quiet	scholar	spiteful	tent	VIP	
quiz	schoolboy	spokesman	terrified	virtuous	
rabbi	schoolgirl	spokeswoman	terrorist	voter	
racist	schoolmate	sponger	test	voyeur	
racketeer	schoolroom	spouse	test	waitress	
rape	schoolteacher	squeeze	thank	warm	
rapist	scientist	stab	thankful	warn	
raunchy	scold	steady	theater	wartime	
ravish	scornful	stepbrother	thoughtless	wash	
reassure	scratch	stepchild	threaten	washroom	
rebellious	scrooge	stepdaughter	thrilled	watch	
rebuff	scrutinize	stepfather	thug	wedding	
rebuke	search	stepmother	tickle	welcome	
receptionist	secretary	stepparent	timid	wheedle	
reckless	seduce	stepsister	toady	whip	
redhead	selfish	stepson	toast	White	
reform	sensitive	stingy	toddler	whore	
regretful	sentence	stop	tolerant	whorehouse	
rehabilitate	sentimental	store	tomboy	widow	
relieved	serenade	stranger	torment	widower	
remind	serene	strangle	tormented	wife	
remorseful	server	street	torture	wild	
renounce	sexist	streetfair	touch	wilderness	
repentant	shaken	strict	touched	windbag	
reproach	sheriff	stubborn	trainee	winkat	
rescue	shocked	stud	traitor	winner	
resentful	shoot	student	traveler	wise	
resort	shopclerk	student	treat	withdrawn	
responsible	shopkeeper	study	troublemaker	witness	
restrain	shoplifter	stupid	truant	woman	
retiree	shopper	subdue	trusting	womanizer	
reverent	shove	submissive	tug	worker	
reward	shrewd	subordinate	tyke	workman	
rib	shrink	subway	unadventurous	workmate	
rich	shy	sue	unambitious	worried	
ridicule	sibling	superior	uncle	yesman	
riot	sickened	supermarket	underachiever	young	
rival	silence	supervise	underdog	youngster	

Appendix D

Blogscope concept list

Table D.1: Blogscope concept words

abandon	aunt	brutalize	club	curse	divorce
abortionist	authoritarian	brute	coach	cuss	doctor
abuse	authority	buddy	coach	customer	dogmatic
accommodate	awe-struck	bully	cocky	cynical	domineering
accommodating	baby	bully	coddle	damn	downhearted
accuse	baby	bum	coed	dance	dress
address	bachelor	bureaucrat	coerce	dare	dropout
admonish	back	bus	cold	daring	drunk
adolescent	badger	businessman	colleague	date	dummy
adult	baldy	businesswoman	combat	daughter	dyke
adulterer	banquet	cafe	comfort	debate	eager
adulteress	baptize	cafeteria	command	debrief	earnest
adventurous	bar	cajole	companion	defeat	Easter
advise	bash	calm	compassionate	defend	easygoing
advisor	bathe	camp	compensate	defendant	ecstatic
affectionate	battlefield	campground	competent	defensive	educate
afraid	beach	campus	competitive	defiant	egghead
aggravated	bed	capitalist	competitor	deflated	egotistical
aggressive	bed	capture	compliment	defy	elated
agitated	bedroom	car	compulsive	degrade	elbow
agnostic	beg	careless	conceited	dejected	elder
aid	beginner	caress	concert	delinquent	elevator
aide	belittle	carnival	condemn	demagogue	embarrassed
airplane	berate	casino	confidant	demean	embrace
alarmed	beseech	catch	confident	demote	employ
alcoholic	bill	cathedral	confine	denigrate	employee
aloof	bind	Catholic	confront	denounce	employer
alumnus	bisexual	caution	congratulate	dependable	encourage
ambitious	bite	cautious	conscientious	dependent	enemy
amuse	bitter	celebration	conservative	deprecate	enraged
analyze	Black	celebrity	considerate	deride	entertain
angry	blame	cemetery	console	despondent	enthusiastic
anguished	bless	challenge	consultant	detective	entreat
annoyed	blonde	champion	contemplate	devil	entrepreneur
answer	blue	chapel	contemptuous	direct	envious
antisocial	bootlick	charmed	contented	disappointed	escape
anxious	bootlicker	chase	contradict	disapproving	euphoric
apathetic	boozer	chastise	contrite	disciplinarian	evangelist
applaud	bore	chatterbox	convalescent	discipline	exalt
applicant	boss	cheat	convict	disco	examination
apprehend	bossy	cheer	cooperative	discontented	examine
apprehensive	bouncer	cheerful	cop	discourage	exasperated
apprentice	boy	cheerless	correct	discouraged	excited
approach	boyfriend	chide	counsel	disgusted	excuse
arrest	brat	child	courageous	disheartened	execute
arrogant	brave	childish	courtroom	dismayed	execution
ashamed	bribe	choke	cousin	disobey	executioner
assail	bride	Christmas	cowardly	disparage	executive
assault	bridegroom	chum	criminal	displeased	exonerate
assist	bridesmaid	church	critic	disrobe	exploit
assistant	brief	citizen	criticize	dissatisfied	extol
atheist	bright	classmate	crook	dissuade	extrovert
athlete	brothel	classroom	crowd	distract	extroverted
attack	brother	clergyman	crushed	distressed	eye
attorney	browbeat	client	cuddle	divorce	face
auction	brunette	clinic	cue	divorc	factory

failure	God	hit	insecure	lazy	mobster
fairground	gossip	hold	insensitive	lead	mock
fanatic	grab	hombre	insider	leader	modest
father	grade	home	insincere	leave	molest
fearful	graduate	homemaker	inspect	lecher	monitor
felon	graduation	homesick	instruct	lecture	moron
female	grandchild	homosexual	instructor	lecture	mortified
female	granddaughter	honeymoon	insult	lecturer	mother
feminine	grandfather	honeymooner	intelligent	lesbian	mother
feminist	grandmother	hoodlum	intern	liar	mourner
festival	grandparent	hooker	interrogate	liberal	mournful
fianc	grandson	horny	interrogation	librarian	mouthpiece
fiance	grasp	horrified	interrupt	library	moved
fight	graveyard	host	interview	loafer	mug
fight	greedy	hostess	interview	lobbyist	mugger
fine	greet	hostile	interviewee	lonely	murder
fingerprint	grind	hothead	interviewer	lonesome	murderer
finicky	groom	hotheaded	intimate	loser	murderess
firstborn	grouch	hotshot	intolerant	lovesick	museum
flatter	grouchy	hound	introspective	lunatic	nag
flee	grownup	houseguest	introvert	luncheon	naive
flirt	guest	housewife	introverted	luncheonette	nark
flophouse	guide	hug	invalid	lunchroom	narrowminded
flunk	gullible	humble	irate	lustful	needle
flunky	gunfight	humiliated	irked	mad	neglect
flustered	gunman	hunk	irritable	malcontent	negotiator
foe	guy	hurry	irritated	male	neighbor
follow	gym	husband	jail	male	nephew
follower	gymnasium	hush	jail	malign	nervous
fondle	gynecologist	hussy	jealous	malingerer	nestle
foolish	gyp	idealistic	jerk	man	neurotic
foreman	hail	idiot	Jew	manager	newlywed
forget	Halloween	ignoramus	jock	manageress	niece
forgive	halt	ignore	joggle	maniac	nightclub
forgiving	handcuff	imaginative	josh	marry	nobody
freeloader	handicapped	imitate	jostle	masculine	nonsmoker
friend	handyman	immature	joyful	masochist	nostalgic
frightened	happy	immoral	joyless	massage	novice
frisk	harangue	impatient	jubilant	matriarch	nudge
frustrated	harass	implore	judge	mature	nurse
funeral	hardworking	imprison	juror	mealtime	nut
furious	harm	incarcerate	kick	medicate	nuzzle
gangster	hassle	inconsiderate	kid	meek	nymphomaniac
gay	hatemonger	incriminate	kid	meeting	obedient
genius	heal	indecisive	kill	melancholy	obey
gentle	healer	independent	kind	merchant	observe
gentleman	Heaven	indignant	kiss	merry	obstruct
ghetto	heckle	indoctrinate	kitchen	millionaire	office
gigolo	hell	industrious	klutz	mimic	ogle
girl	help	infant	knife	mind	old
girlfriend	helper	infatuated	laboratory	minister	opponent
glad	hero	inform	laborer	mischievous	opportunist
gleeful	heroine	informer	lackey	miser	oppose
gloomy	heterosexual	inhibited	lady	miserly	optimistic
glorify	hideout	injure	laud	mistress	order
glum	hire	innocent	lawyer	mob	organizer

orgy	playful	rebellious	schoolgirl	sorry	tackle
orphan	playground	rebuff	schoolmate	spank	tavern
outgoing	playmate	rebuke	schoolroom	spendthrift	taxi
outlaw	poke	reception	schoolteacher	spinster	taxpayer
outraged	policeman	receptionist	scientist	spiteful	teach
outspoken	politician	reckless	scold	spokesman	teacher
overcharge	pompous	redhead	scornful	spokeswoman	teammate
overjoyed	poor	reform	scratch	sponger	tease
overpower	poorhouse	regretful	scrooge	spouse	tease
overwhelm	popular	rehabilitate	scrutinize	squeeze	teenager
overwhelmed	pornographer	relative	search	stab	temperamental
overwork	praise	relieved	secretary	steady	temple
pagan	preacher	remind	seduce	stepbrother	tenant
pal	prejudiced	remorseful	seize	stepchild	tent
pamper	priest	renounce	selfish	stepdaughter	terrified
panicked	priestess	repentant	seminar	stepfather	terrorist
parade	principal	reprimand	sensitive	stepmother	test
paranoid	prison	reproach	sentence	stepparent	test
parent	probationer	rescue	sentimental	stepsister	thank
parody	prod	resentful	serenade	stepson	thankful
parolee	professor	resort	serene	stingy	theater
partner	prompt	responsible	sermon	stoolpigeon	thoughtless
party	prosecute	restaurant	serve	stop	threaten
passerby	prostitute	restrain	server	store	thrilled
passionate	protect	retiree	sexist	stranger	thug
pastor	protg	reverent	shaken	strangle	tickle
patient	Protestant	reward	sheriff	street	timid
patient	protester	rib	shocked	strict	toady
patriot	proud	rich	shoot	strip	toast
patrolman	psychiatrist	ridicule	shopkeeper	stroke	toddler
pauper	psychopath	riot	shoplifter	stubborn	tolerant
peaceful	psychotic	rival	shopper	stud	tomboy
peacetime	punch	rob	shove	student	torment
pedestrian	punish	robber	shrewd	student	tormented
pediatrician	punk	roommate	shrink	study	torture
peeved	pupil	rouse	shush	stupid	tot
penalize	purchaser	ruthless	shy	subdue	touch
penitentiary	pursue	sad	sibling	submissive	touched
perceptive	push	sadist	sickened	subordinate	train
persistent	quack	sadistic	silence	subway	trainee
pessimist	quarrelsome	saint	sincere	sue	traitor
pessimistic	queer	saintly	sinner	superior	traveler
pest	query	salesclerk	sister	supermarket	treat
pester	question	saleslady	slap	superordinate	troublemaker
pet	questioner	salesman	slaughterhouse	supervise	truant
petrified	quiet	saloon	slug	supervisor	trusting
petty	quiz	salute	slum	supporter	tug
photograph	quiz	sarcastic	slut	surgeon	tutor
physician	rabbi	satisfied	sly	surprise	tyke
pickpocket	racist	sauna	smoker	survivor	unadventurous
pickup	racketeer	save	smug	suspect	unambitious
pimp	rape	sawbones	snuggle	suspicious	uncle
pinch	rapist	scared	sock	sweatshop	underachiever
pizzeria	raunchy	schizophrenic	son	sweetheart	underdog
placid	ravish	scholar	soothe	swinger	undergraduate
plainclothesman	reassure	schoolboy	sorrowful	sympathetic	underpay

understanding	vacation	virtuous	watch	wife	worker
undress	vacationer	visitor	wedding	wild	workman
uneasy	vain	voter	weekend	wilderness	workmate
unfair	vengeful	voyeur	welcome	windbag	worried
unfriendly	victim	waiter	wheedle	winner	young
unimaginative	victimize	waitress	whip	wise	youngster
unpopular	vigilante	warm	White	withdrawn	youth
unreliable	village	warn	whore	witness	zoo
upbraid	villain	wartime	whorehouse	woman	
upset	violent	wash	widow	womanizer	
uptight	VIP	washroom	widower	work	

Appendix E

AlphaBet experiment screens

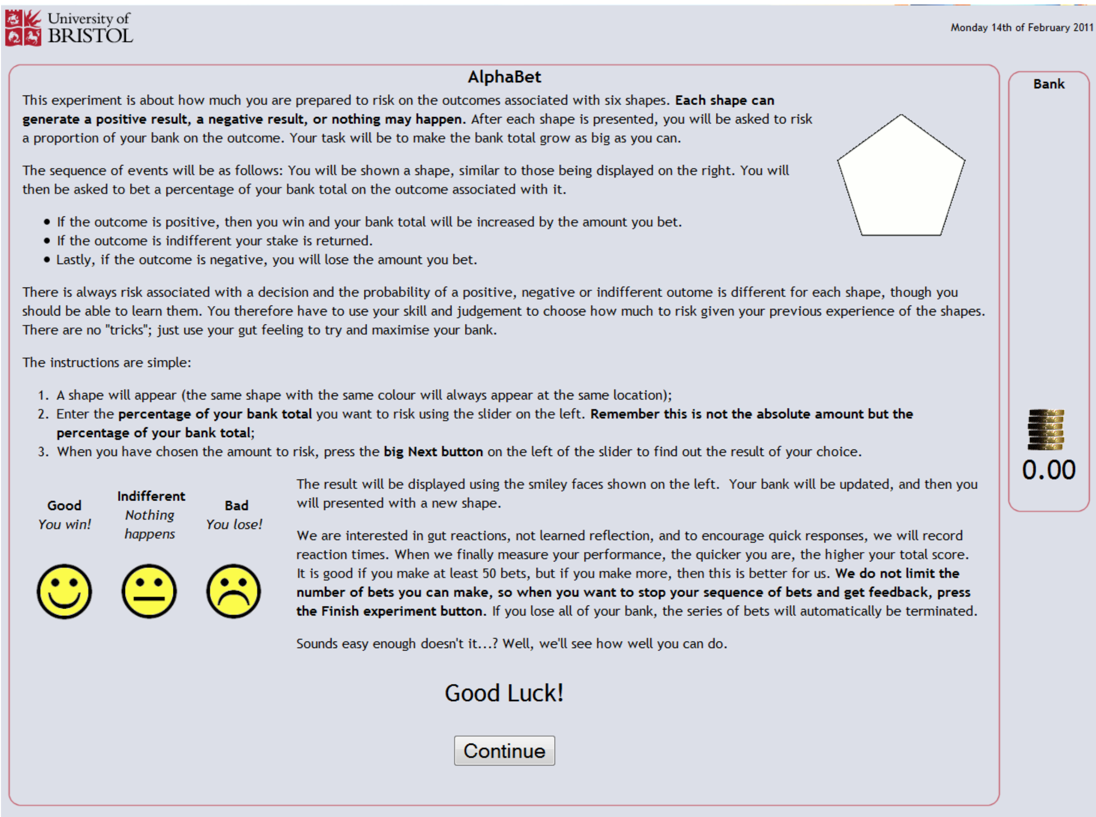


Figure E.1: The instructions/information page presented to participants.

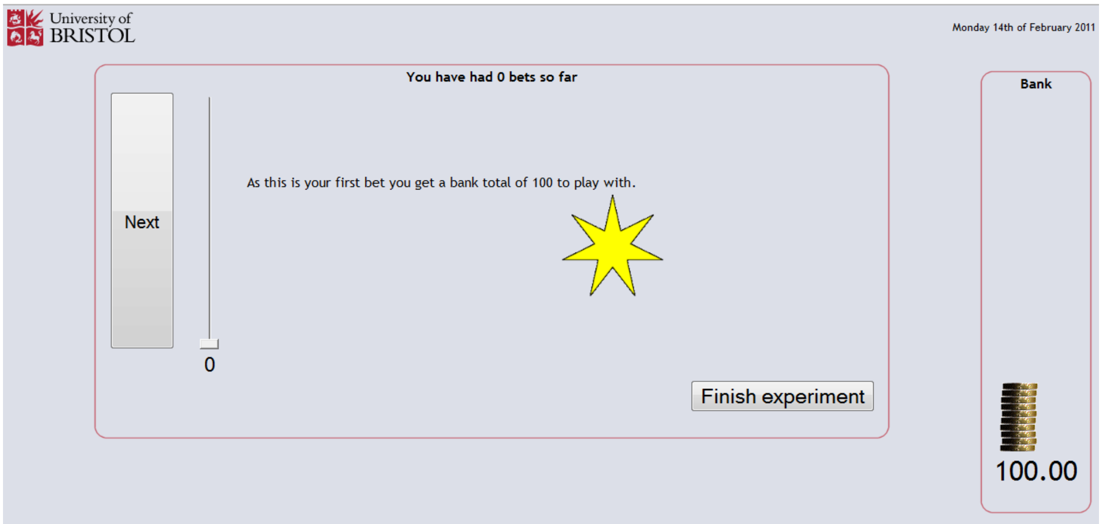


Figure E.2: The first bet page presented to participants with a randomly generated shape and a banked total initialised at 100.

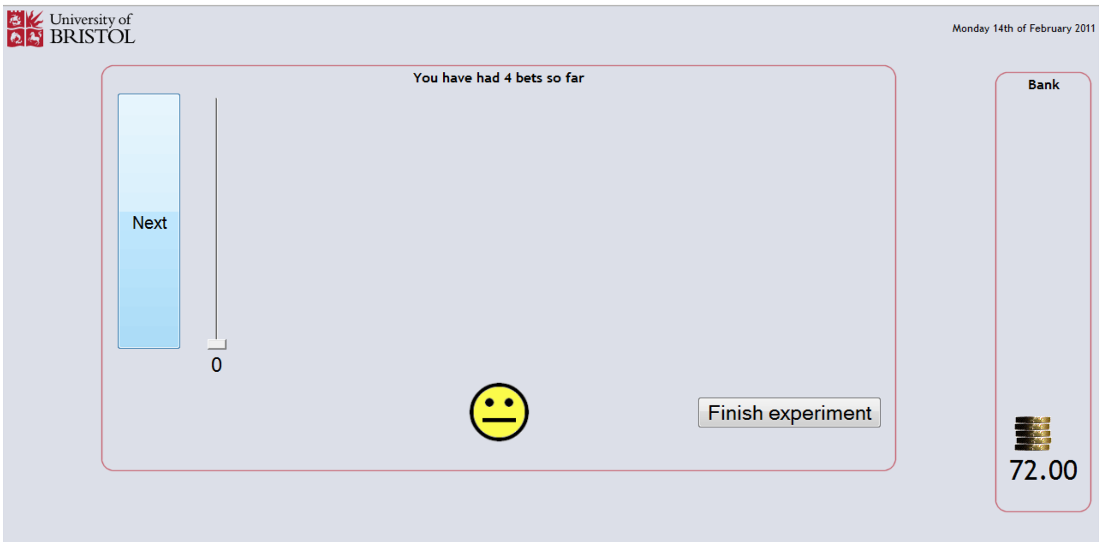


Figure E.3: The feedback page presented to participants showing a 'smiley' face at the bottom and in the centre. In this case the feedback being given is that an indifferent event occurred.

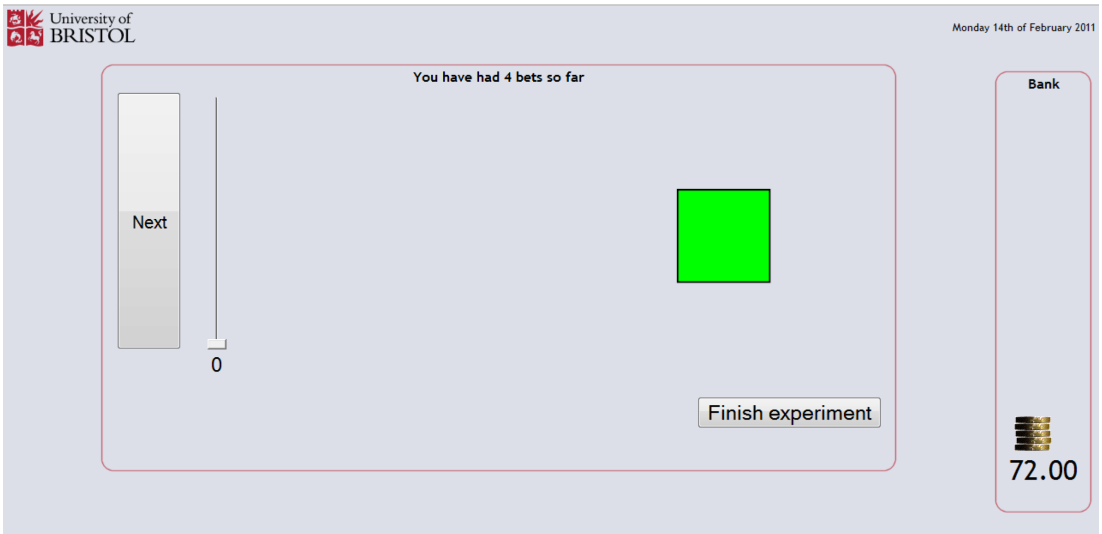
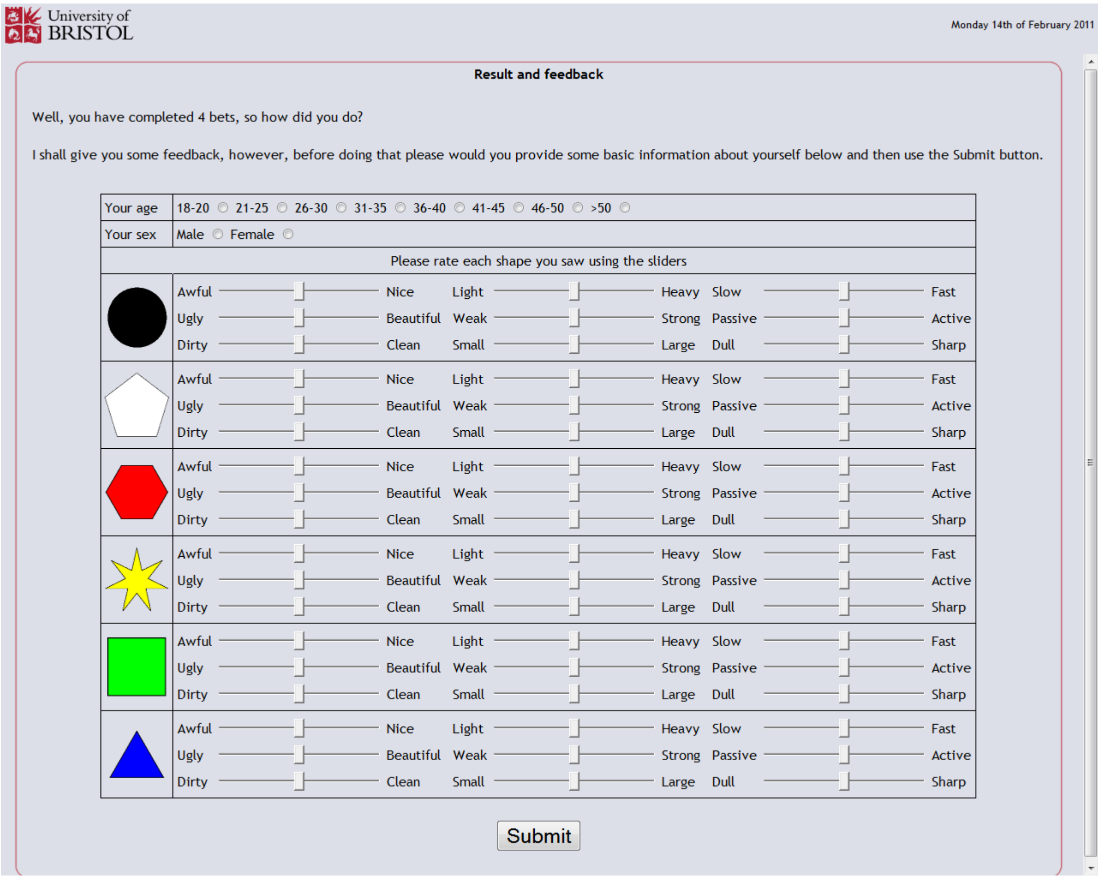


Figure E.4: A subsequent bet page presented to participants with a randomly generated shape, updated banked total and a count of the number of bets made (at the top of the page).



University of BRISTOL

Monday 14th of February 2011

Result and feedback

Well, you have completed 4 bets, so how did you do?

I shall give you some feedback, however, before doing that please would you provide some basic information about yourself below and then use the Submit button.

Your age	18-20 21-25 26-30 31-35 36-40 41-45 46-50 >50		
Your sex	Male Female		
Please rate each shape you saw using the sliders			
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp
	Awful — Nice	Light — Heavy	Slow — Fast
	Ugly — Beautiful	Weak — Strong	Passive — Active
	Dirty — Clean	Small — Large	Dull — Sharp

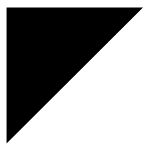
Submit

Figure E.5: The shape rating page consisting of two questions collecting demographic information, then the six experimental shapes each with 9 sliders for participants to indicate their ratings between the extremes indicated by the two opposite words.

Appendix F

Triangles Semantic Differential rating scales

Table F.1: An example of the Semantic Differential ratings sheet used by participants for each triangle. The example below shows the rating sheet for the triangle presented in the upper left corner as depicted at the top of the sheet.



	1	2	3	4	5	6	7	
cruel	☐	☐	☐	☐	☐	☐	☐	kind
curved	☐	☐	☐	☐	☐	☐	☐	straight
masculine	☐	☐	☐	☐	☐	☐	☐	feminine
untimely	☐	☐	☐	☐	☐	☐	☐	timely
active	☐	☐	☐	☐	☐	☐	☐	passive
savory	☐	☐	☐	☐	☐	☐	☐	tasteless
unsuccessful	☐	☐	☐	☐	☐	☐	☐	successful
hard	☐	☐	☐	☐	☐	☐	☐	soft
wise	☐	☐	☐	☐	☐	☐	☐	foolish
new	☐	☐	☐	☐	☐	☐	☐	old
good	☐	☐	☐	☐	☐	☐	☐	bad
weak	☐	☐	☐	☐	☐	☐	☐	strong
important	☐	☐	☐	☐	☐	☐	☐	unimportant
angular	☐	☐	☐	☐	☐	☐	☐	rounded
calm	☐	☐	☐	☐	☐	☐	☐	excitable
false	☐	☐	☐	☐	☐	☐	☐	true
colorless	☐	☐	☐	☐	☐	☐	☐	colorful
usual	☐	☐	☐	☐	☐	☐	☐	unusual
beautiful	☐	☐	☐	☐	☐	☐	☐	ugly
slow	☐	☐	☐	☐	☐	☐	☐	fast

Appendix G

The Problem of Points

The problem of points, a gambling problem which is also known as division of the stakes, is a problem said to have been posed to mathematician Blaise Pascal by French writer, Chevalier de Méré (Antoine Gombaud); the problem was discussed by Pascal and Fermat in a series of letters. The problem that was posed is as follows:

> Suppose two gamblers play a coin tossing game where the first person to get two heads or tails wins, and the game ends after at most three tosses. Suppose also that the gamblers, after placing their stakes, for some reason must stop after one winning head. What would be a rational division of the money?

Intuitively, it seems that the money should not be divided equally because the gambler with the first winning head from the first toss has a greater chance of winning overall. But how much more should this player receive?

If the game continued, it seems that it could end in three different ways: a) by tossing a head, in which case the heads player will win; b) by tossing tails then tails, in which case the tails player will win; c) by tossing tails then heads, in which case the heads player will win: Therefore the stakes should be divided $\frac{2}{3}$ to $\frac{1}{3}$. The conceptual insight, however, is recognising that there are really four possibilities. The extra (imaginary) toss after the second toss comes up heads must also be considered, this makes the complete list of possibilities: a) heads, heads and the heads player wins; b) heads, tails and the heads player wins; c) tails, tails and the tails player wins; d) tails, heads and the heads player wins: Therefore, the rational division of the stakes is $\frac{3}{4}$ to $\frac{1}{4}$.

Pascal (1665) showed that in a game where one gambler needs r points to win and the other needs s points to win, the correct division of the stakes is in the ratio of:

$$\sum_{k=0}^{s-1} \binom{r+s-1}{k} \; to \; \sum_{k=s}^{r+s-1} \binom{r+s-1}{k} \tag{G.1}$$

Appendix H

Image examples

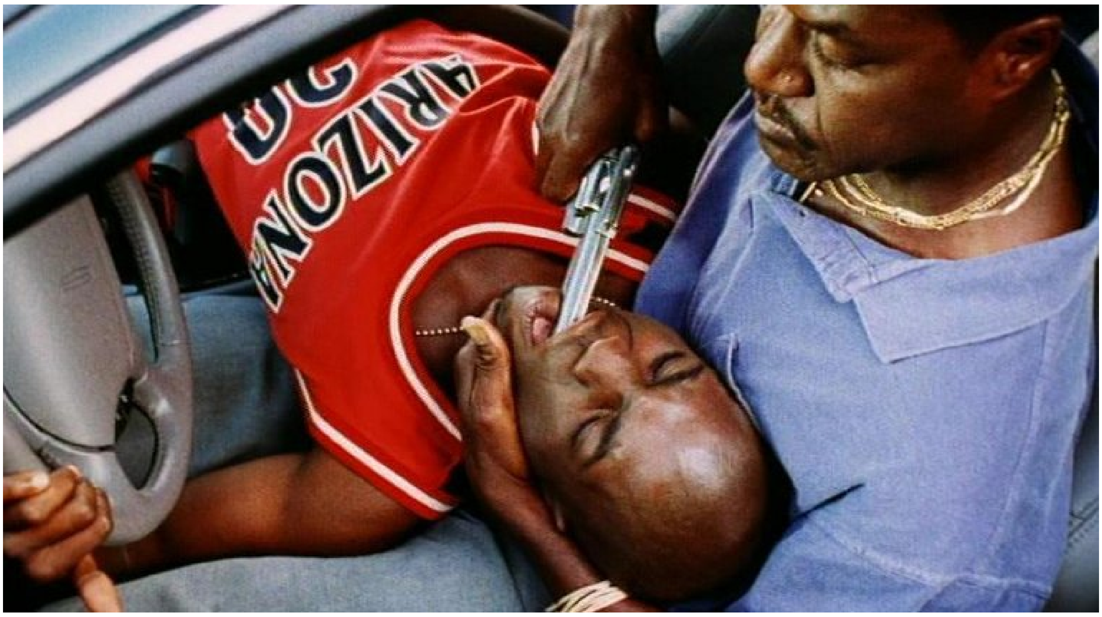


Figure H.1: Image representing the action genre.

Figure H.2: Image representing the classical art genre.

Figure H.3: Image representing the surreal art genre.

Figure H.4: Image representing the nature genre.

Figure H.5: Image representing the romance genre.

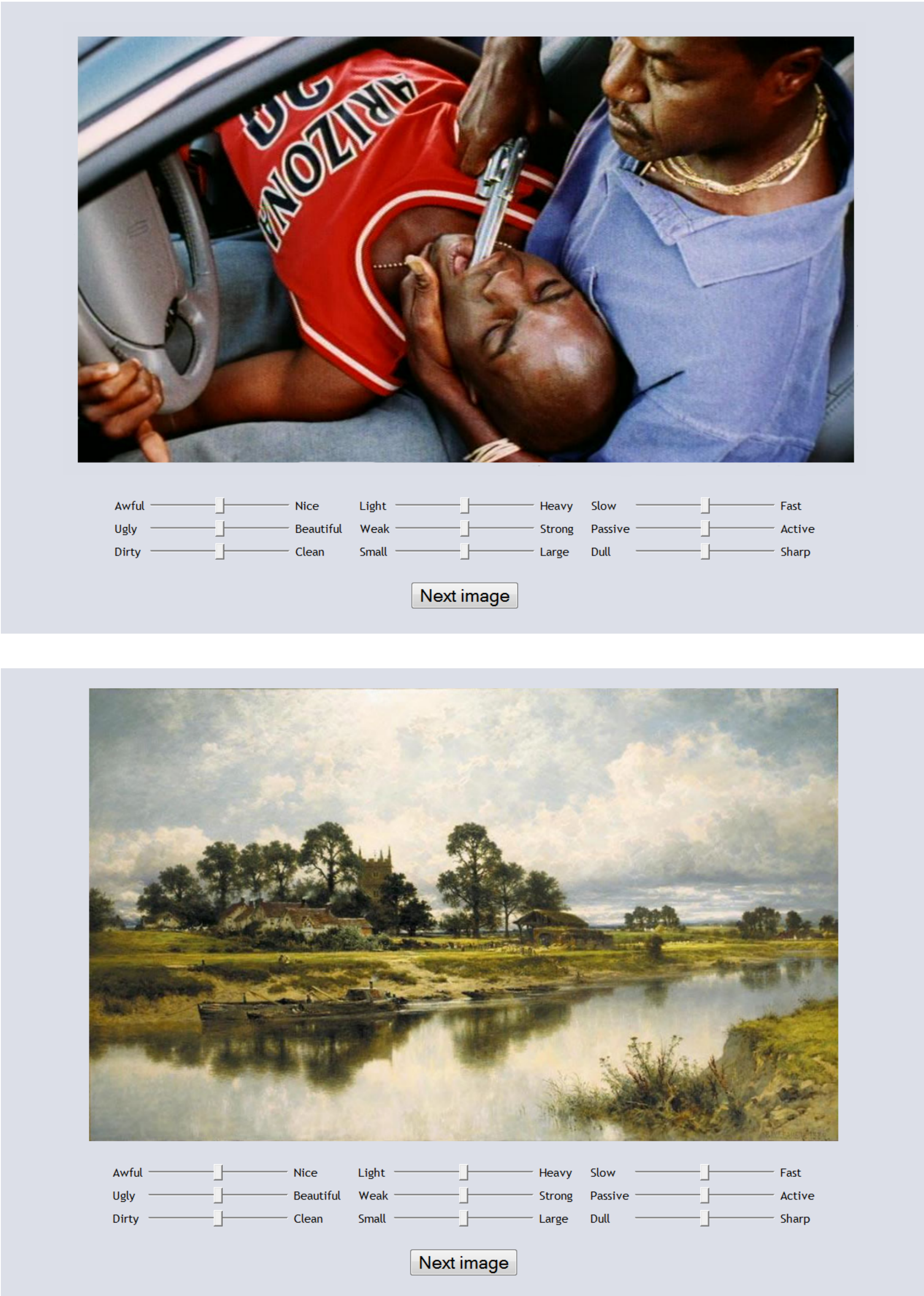


Figure H.6: Presentation of an image on a web page and method of rating using slider bars.

Figure H.7: Example of images used to build the Gaussian mixture model.