Neural Representations of Expected Reward and Risk during Gambling

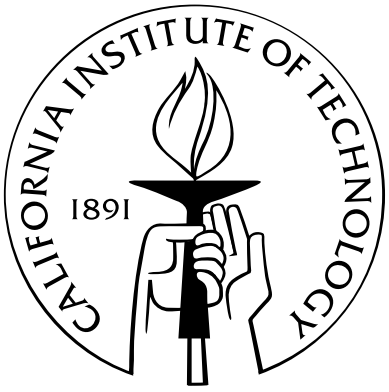
Thesis by

Kerstin Preuschoff

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2007

(Defended November 03, 2006)

Acknowledgements

I would like to thank my advisor, Steve Quartz, for pointing me in the direction of neuroeconomics and his support throughout this thesis. I am greatly in debt to my advisor, Peter Bossaerts, for his continuing support, guidance, and invaluable mentoring. I would also like to thank the members of my advisory committee, John O'Doherty, Shin Shimojo, and Christof Koch.

I am grateful to Tony Bruguier for his collaboration and to Wolfram Schultz, Ulrik Beierholm, Paul Glimcher, and Brian Knutson for many fruitful discussions and lots of encouragement throughout my graduate career.

This work would not have been possible without the continuous support and sympathy of my friends and family, including but not limited to Dirk Walther, Vivek Jayaraman, Johannes and Birgit Graumann, Farshad Moradi, Wolfgang Einhäuser Treyer, Connie Hipp, and of course my parents.

Financial support for this work was provided by the David and Lucile Packard Foundation and the Gordon and Betty Moore Foundation.

Abstract

Organisms continuously monitor the stimuli they encounter and the outcome of their actions. To survive in an uncertain world they aim for rewards and try to avoid punishments. Research in neuroscience, ecology, and economics implies that organisms base their decisions in uncertain situations on expected rewards and risk. Neuroscience focuses on reward prediction learning based on reward prediction errors. In contrast, economic studies emphasize risk in addition to expected reward.

We used functional imaging in humans during gambling tasks to understand how the brain represents expected reward and risk. We find that brain activity in subcortical dopaminoceptive structures can be separated, both spatially and temporally, into signals that correlate with (mathematical) expectation of reward, and with reward variance (risk) – two fundamental parameters in financial decision theory. Our results suggest that the primary function of the dopaminergic system extends beyond its established role in learning, motivation, and salience: it signals different aspects of upcoming stochastic rewards – expected reward and risk.

Based on financial decision theory we then hypothesized neural representations of prediction risk and prediction risk errors. We find that the insula represents both. In analogy with reward representations in subcortical structures, the signals are spatially and temporally differentiated. These findings expand our understanding of the neural basis of decision making under uncertainty by adding prediction risk estimation.

Finally, we investigated where and how expected reward and risk are combined into the neural representation of a gamble's overall value. Using canonical correlation analysis, we find a new predictor that – contrary to expected utility theory – *adds* risk to expected reward. This sum may define a metric of conflict or attention. This metric significantly correlates with activation in the anterior cingulate cortex a structure associated with conflict monitoring.

Drawing on financial theories, we show how the brain represents expected reward and risk. Our results suggest that the earlier understanding of decision making under uncertainty needs to be expanded to include (prediction) risk as measured by variance as well as prediction risk errors. Such integration has far-reaching implications, in particular for pathological decision making.

Contents

List of Figures

List of Tables

Part I

Decision Making under Uncertainty: Overview and Background

Part I Overview

An organism trying to survive in an uncertain world cannot be indifferent to the stimuli it encounters and the outcomes of its actions. Instead, incoming stimuli are assigned a motivational value that guides decision making. The most basic form of assignment is to dissociate rewarding (appetitive) and punishing (aversive) stimuli. Accordingly, the most basic behavior is to seek out rewards and avoid punishments.

In a dynamic and fast-paced environment, fast behavioral reactions can be advantageous. (Accurate) predictions can reduce reaction times. In addition, based on predictions, organisms may evaluate future stimuli and events before they actually occur. Behavioral responses can be chosen and prepared beforehand and thus increase the probability of approaching a rewarding, or avoiding a punishing object.

Early neuroscience research focused on predicting rewarding stimuli and on responses to unexpected appetitive and aversive events. This research was soon related to reinforcement learning, which uses prediction errors to improve future predictions. Since then much of the research within neuroscience has been driven by the reward-prediction-error hypothesis. In the context of learning and motivation, the neural representations of reward, reward predictions and reward prediction errors are now well studied (chapter 1). However, evidence from behavioral ecology shows that values are assigned to stimuli and events not based on predicted reward alone but that consistent preferences are still formed between stimuli of the same predicted reward when the uncertainty associated with the stimuli differs (chapter 2). Most organisms will avoid uncertain rewards yet attitudes toward uncertainty are not as homogenous as those toward predicted rewards.

The mathematical framework for forming preferences and making decisions under uncertainty is often borrowed from economics and finance. Its significance for neuroscience is highlighted by its successful application to behavioral ecology and emerging results on uncertainty-related signals in the brain (chapter 3).

Taken together, previous research in neuroscience, ecology, and economics implies that in order to assign motivational values and decisions in situations of uncertainty, organisms make predictions about expected rewards and uncertainty which are continuously updated using prediction errors.

Chapter 1

Neural Basis of Reward Processing and Reward Learning

The dopaminergic system is the cornerstone of reward processing and reward learning. It is the origin of reward-related input to many brain regions including various limbic regions (caudate nucleus, putamen, nucleus accumbens) and many areas of the neocortex (with a particular focus on prefrontal cortical areas).

Figure 1.1 shows one of the earliest recordings of reward-related signals from dopaminergic midbrain neurons in a behaving monkey. Phasic responses are elicited when a monkey touches a small piece of food that is hidden in a box (Romo and Schultz 1990). The response can be transferred to other somatosensory stimuli (e.g., visual or auditory) if these fully predict the availability of food *before* the monkey reaches into the box (Schultz and Romo 1990). Most importantly, these neurons do *not* respond to fully predicted rewards (Mirenowicz and Schultz 1994). Such conditioning responses have been shown for other conditioning paradigms as well (Schultz et al. 1993). As such, dopaminergic neurons respond phasically to unexpected rewards and reward predicting stimuli.

The resemblance of this response to the error term used in temporal difference (TD) learning, a form of reinforcement learning (RL) (Sutton and Barto 1981), led to the reward-prediction-error hypothesis in which phasic dopamine responses encode a prediction error about the summed future rewards. In TD models, prediction errors are the result of the continuous comparison of predictions of reward with actual rewards. Reward prediction errors are then used to update the future predictions of the model (Montague et al. 1995, 1996; Schultz et al. 1997). The application of the TD model generated many testable predictions, which were first investigated in dopaminergic neurons of the midbrain and

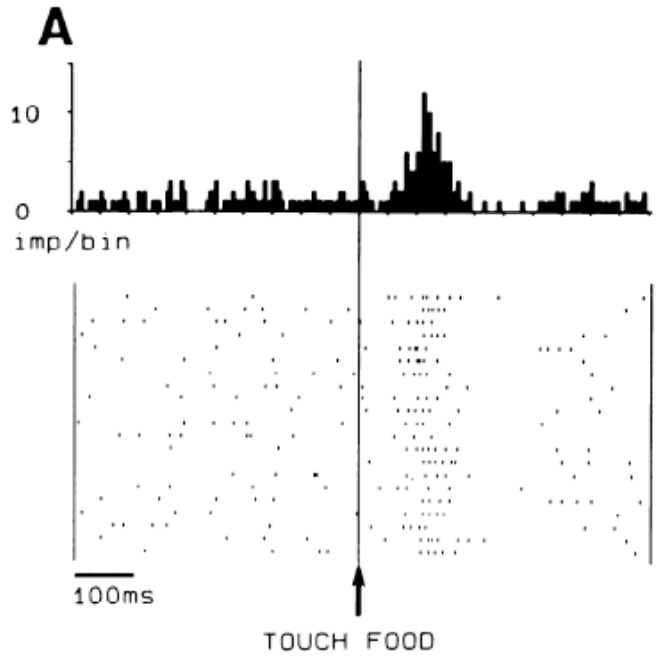


Figure 1.1: Phasic response of a dopaminergic neuron in response to a food reward (Romo and Schultz 1990).

later expanded to include projection targets of these neurons. While the TD model does have limits some of which I will explore in the discussion (chapter 8), the steady growth of the number of structures involved in or modulated by reward processing has emphasized the importance of this model. The following section gives an overview of the structures involved in reward processing and the different aspects of reward that are known to be encoded. As this is a field of intense research, the list is by no means complete and includes mostly very early electrophysiology and fMRI studies. Many of the more recent studies are omitted at this stage as they will be related to the current research in later chapters.

1.1 Reward-processing structures

Brain regions that respond to rewards include ventral tegmental area (Romo and Schultz 1990), the striatum (including caudate, putamen, ventral striatum and nucleus accumbens) (Apicella et al. 1991; Hassani et al. 2001), subthalamic nucleus (Matsumura et al. 1992), pars reticulata of substantia nigra (Sato and Hikosaka 2002), dorsolateral and orbital prefrontal cortex (Tremblay and Schultz 1999), anterior cingulate cortex (Niki and Watanabe 1976),

amygdala (Nishijo et al. 1988), and lateral hypothalamus (Burton et al. 1976). In addition, the activity in numerous brain regions seems to be modulated by reward resulting in higher activity during rewarded vs unrewarded trials. These include many of the areas listed above plus the perirhinal cortex (Liu and Richmond 2000), superior colliculus (Ikeda and Hikosaka 2003), multiple regions within parietal cortex as well as pars compacta of substantia nigra (Waelti et al. 2001). The specific aspect of reward and the response pattern that reward triggers varies across regions.

For instance, reward-detecting neurons in orbitofrontal cortex (OFC) can dissociate between liquid and solid food rewards, discriminate positive and negative reinforcers and reflect the difference in value of (identical) food when presented in states of hunger vs satiation (Critchley and Rolls 1996), all shown in nonhuman primates. In humans, the involvement of OFC in coding stimulus value has been established for rewards associated with sensory modalities (Anderson et al. 2003; O'Doherty et al. 2000) as well as abstract rewards such as money (Elliott et al. 2003).

The phasic response of the dopaminergic neurons to the delivery of unexpected rewards and reward-predicting stimuli increases for some of these neurons with increasing reward probability (Fiorillo et al. 2003).

A reward prediction error has been confirmed in monkey midbrain neurons (Bayer and Glimcher 2005) and is also seen in putamen in human imaging studies (McClure et al. 2003). Signals associated with reward expectation and anticipation have been found in medial prefrontal cortex and in human ventral striatum as well as in the amygdala (Knutson et al. 2001a,b).

Early research in amygdala had suggested that its activity can determine the magnitude of rewards (Pratt and Mizumori 1998) and predict expected outcomes during learning (Schoenbaum et al. 1998). More recently, neuroimaging studies have pointed to an intensity rather than valence encoding (Anderson et al. 2003) in amygdala.

1.2 Summary

Neuroscience research has focused on predicting rewarding stimuli and events and on responses to unexpected appetitive and aversive events. The resemblance of the encoding seen in dopaminergic neurons to the error term used in TD learning has helped to significantly

expand our understanding of neural reward processing.

However, due to the influence of TD learning the vast majority of representations have been related to learning and motivation. In the absence of learning, are predictions about reward made and reward prediction errors still generated? In addition, evidence exists that it is not just expected reward that plays a role in decision making under uncertainty. The next chapter will look at results from behavioral ecology that point to a more complex behavior in decision making under uncertainty.

Chapter 2

Behavior in Stochastic Environments

Foraging behavior is studied in behavioral ecology and provides further support for the reward-prediction-error hypothesis as organisms consistently prefer stimuli with larger expected rewards over those with smaller. However, foraging behavior indicates that when assigning motivational values in situation of uncertainty, the uncertainty itself influences decision making.

When a bee flies over a bed of flowers that it frequently visits, it will prefer flowers which are more likely to have nectar over those that are less likely. That in itself is not surprising. However, the bee does *not always* choose the more likely floral type. Instead, when a flower is twice as likely to contain nectar than another one the bee will choose this type of flower about twice as often (Real 1991). Is the occasional visit to the nectar-poor floral type an error of choice or does bees' choice behavior consistently reflect the statistics of floral fields?

One of the first to formally study choice behavior in stochastic environments was Richard Herrnstein in his Harvard pigeon lab (Logue 2002). Together with his colleagues he studied pigeons' foraging behavior by manipulating the probability of food availability. In a typical experiment, pigeons had to pick one of two levers one of which would be "better," i.e., more likely to result in food pellets, than the other. The pigeons' choice behavior, just like the bee's, revealed a preference for the "better" lever with occasional visits to the other lever. The researchers noticed that the ratio of visits to the levers was not random. Instead, it was best described by the matching law, a quantitative model, that states that pigeons distribute their choices according to the ratio of likelihoods of reward on either lever (Herrnstein 1961,

1974). Foraging behaviors that reflect the average amount of reward available have since been observed in many other species including pigeons (Davison and Hunter 1976), rats (Graft et al. 1977), nonhuman primates (Sugrue et al. 2004) and humans (Baum 1975). It seems irrelevant if the difference in probability is presented in terms of spatial, temporal or quality aspects of the available food or stimulus.

This behavior mirrors the findings of the neural basis of reward processing and is compatible with error based learning. Stimuli that provide larger rewards are assigned higher motivational values and are preferred over those with lower motivational values. In addition, when underlying averages change suddenly, most organisms will adapt over a short period of time (e.g., (Dorris and Glimcher 2004)) in a manner consistent with TD learning. As the predictions are now less accurate, prediction errors will occur more frequently (in the direction of the new averages) and predictions are updated accordingly to reflect the new averages.

In summary, when an organism is presented with two stimuli of which one stimulus on average is more rewarding than the other stimulus, the organism will prefer the more rewarding stimulus. The difference in average is mirrored in behavior such that the larger the difference, the more preferred the "better" stimulus becomes. So when averages are the same, and decisions or preferences are based on averages alone, organisms should be indifferent.

However, when a bee flies over a bed of flowers in which some floral types contain nectar for sure and others have more nectar but its availability is less certain, bees will prefer a constant rate of nectar over varying rates (although the average rate is the same) (Real 1991). Such variance sensitive foraging demonstrates that two stimuli which are on average equally rewarding (based on expected reward) do not have the same value to the organism. Instead, a second parameter comes into play; variance, or more general, uncertainty.

Variance-sensitive foraging behavior is found in many species. Sparrows (Caraco 1983), juncos (Caraco 1981), wasps (Real 1981), rats (Battalio et al. 1985), and shrews (Barnard and Brown 1985) all favor constant food availability over varying one given that the average amount of food is the same. The macaque is the only nonhuman animal for which a consistent preference for varying over constant rewards has been reported (McCoy and Platt 2005). In addition, variance-sensitivity has been successfully manipulated in other species as a function of the organism's energy budget (Caraco 1981; Barnard and Brown 1985). As

such, preferences with respect to variance are not as straightforward as reward-sensitivity.

In decision making under uncertainty, does variance (or uncertainty) play a role similar to that of expected reward? Are there neural correlates of uncertainty in analogy to expected rewards? I.e., are there predictions of uncertainty and errors about these predictions? Can the TD model make use of such predictions? Or is there a separate TD model for uncertainty?

This idea gets strong support from economics and finance, which observe similar behavior in human economic decision making under uncertainty. The mathematical framework that has been developed to model such behavior may be able to inform and guide neuroscience research under uncertainty just as TD learning has guided research on reward prediction and reward processing. For this, at least two things should be established:

(i) Preferences of choice in humans using abstract reinforcers such as money are similar to preferences of choice in both humans and non-humans using primary rewards, such as gustatory stimuli. This is to show that motivational values are assigned in the same way in humans and non-humans and that motivational values are assigned in the same way for primary and abstract rewards.

(ii) Economic theories can be applied successfully to behavioral and neural data.

Assuming (i) and (ii), uncovering the neural basis of basic economic variables and concepts would provide the inputs and therefore be an important step toward understanding decision-making mechanisms with the potential to expand the results to non-humans.

Using simple variables such as magnitude, probability, and utility of reward, economics and finance successfully describe choices under uncertainty in humans, and therefore have helped to better understand behavioral decision processes made by individuals. While the above accounts of foraging behavior and reward processing circuits may imply that uncertainty, just like expected reward, *must* have explicit neural representations, economic theories demonstrate that variance-sensitive behavior can be generated without explicitly encoding variance. This suggests that implicit neural codings have to be considered as well. In the next chapter I will review the two main economic approaches to establishing motivational values. The first is rooted in economics and uses implicit representations of both expected reward and risk. The second is rooted in finance and uses explicit representations.

Chapter 3

Preferences of Choice in Economics and Finance

"Suppose two players, A and B, are playing a three-point game, each having wagered 32 pistoles, and are interrupted after A has two points and B has one. How [many pistoles] should each receive?" (Encyclopaedia Britannica 2006) This problem was posed by Chevalier de Mere in 1654 to Blaise Pascal which led the latter to develop the theory of probability (and expected value).

To arrive at the answer (48 for A, 16 for B) Pascal suggested, that choices are based on the expectation (or expected value), i.e., the sum of the value of all possible events weighted by the probability of their occurrence. Given N possible outcomes x_i (here $N = 2$, corresponding to the two outcomes `"A wins"` and `"B wins"`), each having the probability $p(x_i)$ of occurring, the expectation EV is given by the sum of all possible outcomes x_i, with $i = 1, ..., N$, each weighted by their probability $p(x_i)$

$$EV = \sum_{i=1}^{N} p(x_i) \cdot x_i. \tag{3.1}$$

The idea of maximizing this expectation (or expected value) was the basis of early economic theories of choice. At the time expected value was equivalent to what today is called expected reward. Just like neuroscientists, economists assumed that options with higher expected rewards are always preferred over those with lower expected rewards. Uncertainty first entered the picture in 1738 (Bernoulli 1738), yet it was not further developed and incorporated into theories of decision making until a couple of hundred years later (Knight 1921; von Neumann and Morgenstern 1944).[1] Since then two questions have driven eco-

[1] In economics, risk denotes situations of uncertain outcomes for which the probability of the outcome

nomic theories: How are preferences formed for uncertain prospects? and, How do you mathematically model an uncertain prospect?

Today, there are two major mathematical approaches to preferences of choice under uncertainty, one rooted in economics, the other in finance. They differ primarily in how they deal with choice under uncertainty. In finance, expected reward and risk are modeled explicitly, in economics they are modeled implicitly.

3.1 Economics and the notion of expected utility

To model choices, neoclassical economics describes the value of goods and services in terms of expected utility. Expected utility is derived from a *utility representation* which captures preferences and values in form of a single-dimensional index. A good with higher utility is preferred over one with lower utility and a preferred good has higher utility than all other available goods. *Expected utility* was first introduced by Daniel Bernoulli (Bernoulli 1738) and later axiomatized by John von Neumann and Oskar Morgenstern (von Neumann and Morgenstern 1944). Expected utility expands the idea of utility by including choices with uncertain outcomes. In analogy to utility, an option with higher expected utility is preferred over an option with lower expected utility. And a preferred option has higher (or at least equal) expected utility as all other available options. Expected utility sums the utilities $u(x_i)$ of all possible outcomes $x_i, i = 1, ..., N$ of an action or choice weighted by their probability of occurrence (given that choice).

$$EU = \sum_{i}^{N} p_i u(x_i). \tag{3.2}$$

As utility or expected utility is based on an axiomatic approach, the utility function can take many different forms as long as the axioms are satisfied. When there is no uncertainty, a utility representation $u(x)$ has very little structure, it is simply a preference ordering. As such, a statement such as $u(x_1) = 2u(x_2)$ has no meaning. This representation can easily capture individual preferences. Differences in attitudes toward goods, services, probabilities, risk, ambiguity, and wealth, and even differences that arise from emotional states may all

is known, e.g., betting on a fair coin flip is risky as the probability of winning is 50%. Ambiguity or uncertainty denote uncertain outcomes for which the probabilities are not know. In neuroscience and other fields, uncertainty often (but not always) means risk. In this thesis I will use the term uncertainty in its most general form which includes both ambiguity and risk.

contribute to preference formation. However, in economics it is usually not of interest how the exact shape of an individual's utility curve comes about as long as it correctly predicts the individual's preferences. As such, it is often the goal to find a mathematical function that satisfies all axioms and at the same time can fit accurately many different forms of decision-making behavior under risk. However, how probability and risk are integrated to give rise to that function is of no interest.

This is where neuroscience diverges. Assuming that there is an expected utility function in the brain, how is it put together? Which aspects of a good or option do we perceive? How are these quantities integrated into a utility index? As a result, utility or expected utility is a good framework to determine preferences but often falls short on (or does not attempt) accurately describing phenomena such as risk for instance.

The interesting thing is that economics accounts for risk without explicitly modeling risk. According to the representation above an explicit (neural) representation is not necessary to form preferences under uncertainty to account for things such as risk aversion.

3.2 Finance and the notion of risk

Finance deals with how prices are determined in financial markets. And while this is a fascinating field on its own, neuroscience is only interested in how finance determines preferences of choice under uncertainty. The two key ideas that neuroscience is interested in were both introduced by Harry Markowitz (Markowitz 1952): (i) Preferences can be accurately described using only the expected return (expected reward or *mean*) and risk of a security or portfolio[2] (equation (3.3)). (ii) Risk is measured as *variance*.

$$EU \sim a \cdot E[R] + b \cdot var[R]. \tag{3.3}$$

For neuroscience this translates into two hypotheses: (i) There are explicit neural representations of expected reward and risk, (ii) Expected rewards are measured as the mean, uncertainty or risk is measured as the variance of outcomes.

Examining expected utility (preferences) (equation (3.2)), Harry Markowitz noted that determining a utility function is often not feasible and may require substantial computational power. He proposed an approximation to EU that uses only the means and variances,

[2] the securities held by an investor

the two key parameters that characterize uncertain outcomes. This is in effect a second-order Taylor series expansion. It is generally agreed that this mean-variance scheme is but an approximation of expected utility that tends to incorrectly describe preferences in situations of high skewness and kurtosis (or nonnegligable higher-order terms). And while skewness and kurtosis often pose a problem in finance they are less relevant in natural environments where many processes are well described by Gaussian distributions.

The computational ease of a mean variance approach is of particular interest to neuroscience. Equation (3.2) implies that for a neural system to determine expected utility it is necessary to neurally represent the probability p_i and utility $u(x_i)$ of *every* possible outcome. Intuitively, this seems impractical if there are many possible outcomes. If however, there are explicit neural representations of means (expected reward) and variances (risk) the problem of determining EU can be reduced to the trade-off between the two (equation (3.3)).

3.3 Applications of economics and finance to foraging in stochastic environments

The measures provided by economics and finance have frequently been applied to and often have been successful in describing foraging behavior (Battalio et al. 1985; Real and Caraco 1986). Evidence is limited, though, of successful applications in neuroscience.

A particularly interesting result is the biomechanical derivation of a utility-type function for bumblebees based solely on biomechanical principles (Harder and Real 1987). Instead of measuring a bee's utility function by observing their foraging behavior in situations of varying expected reward and risk, the researchers asked, "why are bumblebees risk-averse?" They defined maximizing the expected rate of net energy intake as the bumblebee's goal (rather than maximizing expected utility). The net energy intake is a nonlinear function of the net nectar volume. It can be derived only from biomechanical principles such as the bee's weight, flight time between flowers, energy per drop of nectar and so on. This function is *not* derived from the bees preferences during foraging yet it exhibits the typical concavity seen in utility functions and therefore accounts for the bees' sensitivity to variance seen in their foraging behavior. E.g., for large nectar volumes the energy intake per volume is smaller than for small nectar volumes and therefore a constant nectar volume should be

preferred over a varying nectar volume of the same mean.

3.4 Risk processing structures

Within neuroscience, uncertainty is less well understood and has primarily been explored in the context of complexity (Lauwereyns et al. 2002; Takikawa et al. 2002; Grinband et al. 2006), entropy (Harrison et al. 2006), or ambiguity (Hsu et al. 2005). In recent years, functional neuroimaging studies have pointed at the involvement of cortical structures such as anterior insula and parietal cortex in risk processing (Huettel et al. 2005a). In addition, there is some evidence that neurons of the ventral tegmental area encode a quantity that correlates with measures of uncertainty (Fiorillo et al. 2003). The specific measure of risk or uncertainty has yet to be established.

3.5 Summary

Economics and finance provide two different approaches toward modeling and determining preferences of choice under uncertainty. Finance shows that preferences can be described (close to) accurately when expected reward and risk are modeled explicitly. Economics shows that preferences can be described accurately with a single-dimensional index that captures attitudes toward expected reward and risk implicitly. The neural basis of reward processing is well studied yet the relationships that have been established between neural activity and expected reward, or between neural activity and risk, are usually simply monotonic. The commonly used block designs of fMRI studies in particular can capture only preference orderings rather than specific representations of expected reward, risk, utility or other value functions. Economics and finance can inform neuroscience by suggesting specific models for encoding values and preferences. In return, neuroscience can inform economics and finance by establishing which of the suggested measures have neural representations.

Chapter 4

Toward a Hypothesis of Decision Making under Uncertainty

In decision making under uncertainty, neuroscience has focused on reward and reward prediction learning. However, ecology shows that uncertainty or variance also plays a role.

Integrating research from neuroscience, behavioral ecology, economics and finance suggests two hypotheses of how motivational values can be assessed in situations of uncertainty:

(i) Motivational values are captured by (predictions of) a single-dimensional index. Explicit neural representations of (subjective) probability and utility functions are necessary; explicit representations of risk are not.

(ii) Explicit neural representations of (estimates of) expected reward and risk are combined into a motivational value. Thus, the trade-off between the two is also made explicit. Estimates are updated using prediction errors of both expected reward and risk.

Several questions arise from this. Are expected reward and risk explicitly represented in the brain as suggested by finance? More specifically, are there brain regions whose activity increases linearly in expected reward? Are there brain regions that encode variance? Or is there a value function analogous to expected utility that implicitly generates preferences for high expected rewards and low risks? Are there neural signals that can be used to update estimates of expected reward and risk.

The suggested mathematical framework was originally derived from (and is therefore known to work for) human behavior in mostly financial decision making. It seems best to first explore the implications of these hypotheses in humans using similar tasks. Neural

activity can be accessed using functional magnetic resonance imaging (fMRI). However, in order to extend such studies at a later stage to nonhuman animals it would be advantageous to use a paradigm that is well explored within both neuroscience and ecology and can easily be transferred to other species, such as a foraging task like matching.

4.1 Human behavior in stochastic environments

We conducted two behavioral studies to explore human "foraging" behavior in financial environments. The studies also serve to explore the potential of matching tasks for exploring neural correlates of expected reward and risk using functional imaging in humans. The first is a standard matching task, the second a reversal task. Both have been described and results from previous research have been presented in chapter 2.

4.1.1 Matching task

In the matching task, on each trial, subjects were presented with two visual stimuli (white squares) to the left and right of a fixation cross, exactly one of which would hold the reward. Subjects had to choose a square (left or right); the outcome would be revealed a little later. The probability of a monetary reward ($0.25) to be in either box was fixed; one side was set to $p_1 = 0.7$, the other to $p_2 = 1 - p_1 = 0.3$. Probabilities did not change over the course of a session. Subjects were instructed to maximize their overall reward. Using expected reward, it can be shown that the optimal strategy is to continuously choose the square with the higher probability p_i of winning.

Figure 4.1 shows the data for three different subjects for a single session (100 trials). Each panel plots for each trial i the percentage with which the subject has chosen the left box over the past i trials (solid line), as well as the percentage of the left box winning over the past i trials (dashed line). Thus, the dashed line indicates the probability with which the left square has been rewarded in the past. For the first subject (top panel), the left square had a probability of reward of $p = 0.3$. Toward the end of the experiment, the subject's choices match this probability almost perfectly (solid line). This is consistent with the matching law established in behavioral ecology (Herrnstein 1974).

The typical behavior of a second group of subjects is shown in the middle panel of figure 4.1. The subject continuously chooses the box that is more likely to obtain the reward.

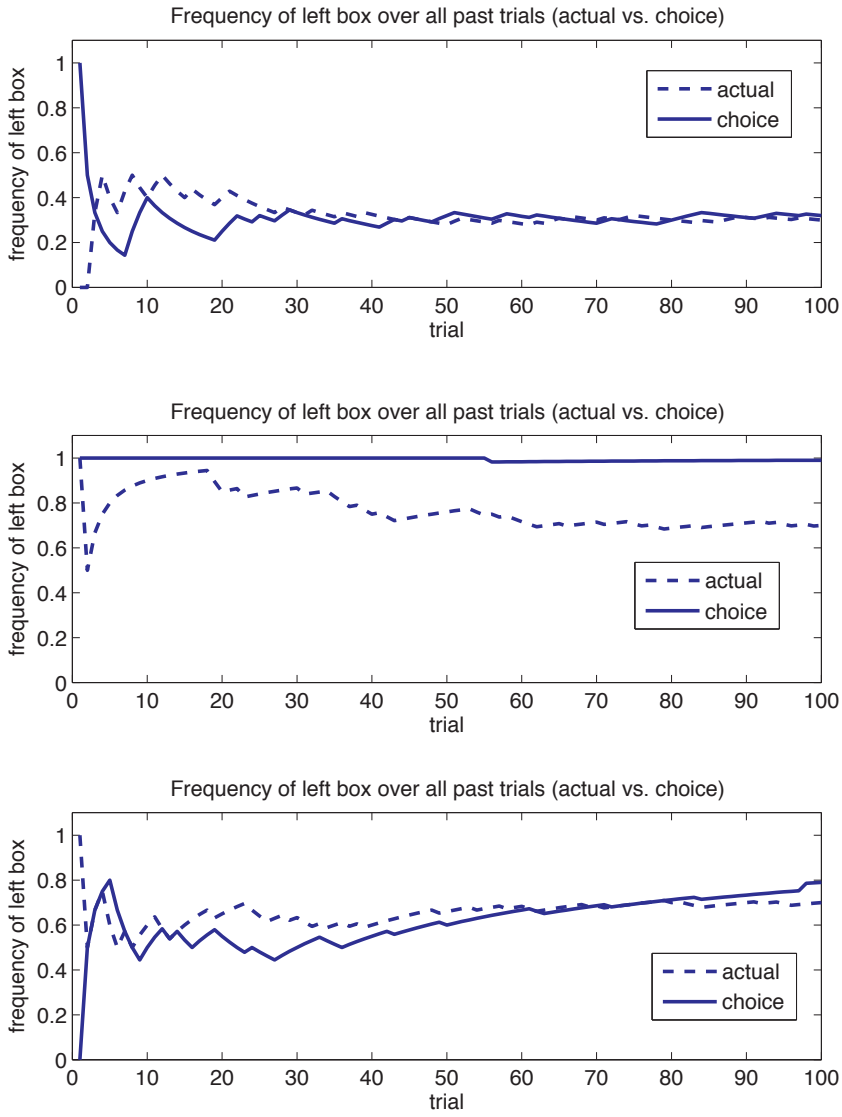


Figure 4.1: Human responses to probabilities. *Upper panel.* Matching behavior. The ratio of choices attributed to the left and right square matches the ratio of probabilities with which a reward is placed behind the left or right square. *Middle panel.* Optimizing behavior. Over the first few trials the subject determines the "better" square and continuously chooses it throughout the rest of the experimental session. *Bottom panel.* The subject starts out matching (as in the upper panel) but converges on the optimal strategy toward the end of the experiment.

This is indeed the optimal strategy which is consistent with choice behavior as predicted by economics and finance. However, this behavior diverges significantly from predictions made by the matching law.

The bottom panel of figure 4.1 shows the behavior for a third subject which starts out matching but slowly learns that there is a better strategy. About halfway through the session this subject switches from matching to optimizing behavior (which is indicated by the overmatching at the end of the experiment). Thus the subject seems to have learned the optimal strategy.

Of the nine subjects that participated in this task five matched, three optimized, and one learned to optimize over the course of the experiment. The exact distribution of these behaviors across the general population is of little significance here and has been studied elsewhere. There are several things to note though: (i) Human choices reflect preferences toward higher expected rewards. (ii) As opposed to nonhuman animals, humans show two significantly different types of behavior, a matching-type behavior in analogy with reports from behavioral ecology (Baum 1975) and an optimized behavior. The split of behaviors makes it difficult to draw conclusions about the general population. More importantly, to explore the neural basis of probability and risk perception a design is needed that will minimize the difference in perception of probabilities and risk between the different groups.

4.1.2 Reversal task

In a second experiment, a reversal task, the same paradigm was used. However, after 20–30 trials the probabilities of reward associated with the two boxes changed. Reversals occurred 5 times throughout one session (approximately 120 trials) resulting in 6 blocks per session. Subjects were informed that probabilities would change several times throughout a session. In order to maximize their reward, subjects had to detect the change of probabilities and adjust their behavior accordingly. Again, we are interested primarily in how human behavior in such tasks compares to nonhuman behavior. Figure 4.2 shows the results for a single subject. The results compare qualitatively to those obtained from reversal tasks in nonhuman primates (Dorris and Glimcher 2004) as changes are detected with a small delay (consistent with TD learning). However, all subjects who participated in the study showed optimized behavior within each block once the change had been detected.

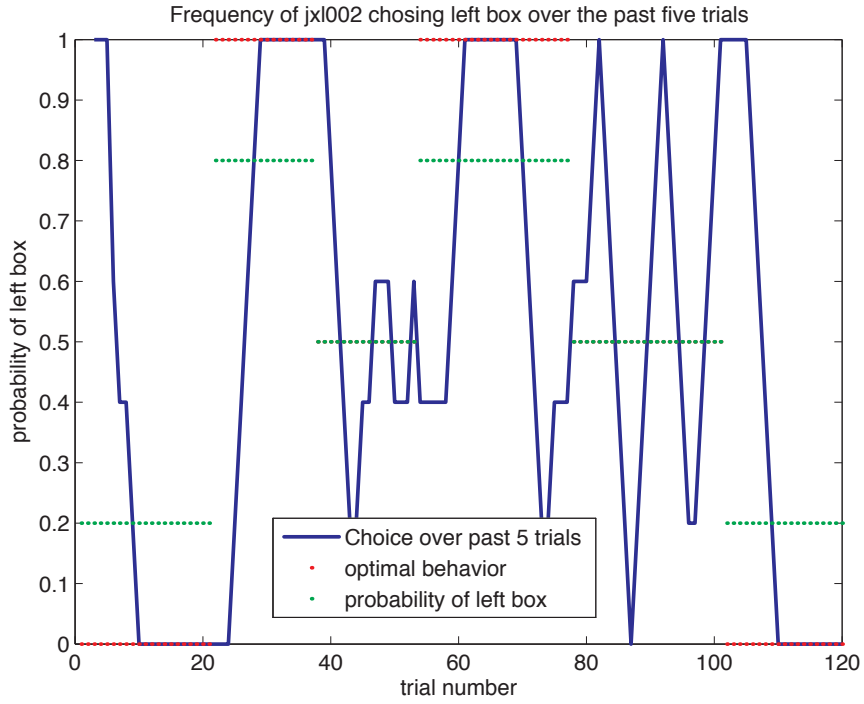


Figure 4.2: Human responses in a matching reversal task in a single subject. The solid blue line indicates the frequency with which the subject chose the left box over the past 5 trials. Dotted green lines indicate the probability of the left box on this trial. Dotted red lines indicate optimal choice behavior as predicted by economics. Detection of change is reflected in subject's behavior several trials after the change occurred. Once the change is detected, the subject follows the optimal strategy within each block.

4.2 Conclusions

While some subjects showed matching behavior similar to that seen in nonhuman species, humans seem to have the capacity to adapt additional strategies which are in fact optimal in an experimental setup. The experiment itself, and especially the dissociation of different player types (i.e., matchers and optimizers) poses problems for an fMRI experiment that is aimed at finding neural representations of expected reward and risk: (i) In optimizers, there is no behavioral measure for the perceived probability and risk. Both the matching and the reversal behavior reflect a simple preference ordering but not necessarily a representation of expected reward. Such representations will be particularly difficult to assess in the reversal task. (ii) For both matchers and optimizers it is difficult to assess at which point subjects have established a stable estimate of probability due to ongoing learning. (iii) The dichotomy between matchers and optimizers is not a strict one as matchers are capable of optimizing as well. As such, there may be an additional process which may eventually result in behavioral switches. The neural basis of this learning is unknown. (iv) Due to (iii) it may also be difficult to dissociate between matchers and optimizers throughout an experiment.

In summary, while matching tasks are great to explore learning about probabilities, they seem too complex to explore basic representations of expected reward and risk. A more controlled experiment is needed that avoids the problem of different behavioral types and that does not involve learning.

Part I Conclusions

Driven by models of TD learning, neuroscience has focused on reward prediction and reward learning to explore decision making under uncertainty. And while ecology shows that estimates of uncertainty are equally important for preference formation, neuroscience provides very limited evidence or neural encodings of uncertainty.

Two approaches to preference formation suggested by economics and finance provide a mathematical framework for exploring neural representations of expected reward, risk and motivational values.

Combining the ideas and research from neuroscience, ecology, and economics, and using fMRI in humans we will test the following hypotheses:

- There are neural representations that are linear in probability and/or expected reward.
- There are neural representations of risk such that risk is measured as variance.
- Representations of both expected reward and risk are estimates or predictions associated with future stimuli and events.
- At the time of the stimulus prediction errors are generated whenever predictions are violated. Such prediction errors can be used to update future estimates of expected reward and risk.
- Responses to expected reward and risk occur in the absence of learning and motivation.
- Expected reward and risk are combined to form motivational values.
- Expected reward and risk can be used to evaluate different forms of uncertainty.

The reminder of this thesis is devoted to exploring these hypotheses. Part II reports on a functional imaging experiment that uses a gambling task to probe for the hypothesized

neural responses. Part III investigates the implications of this work and presents another brief pilot study as a first step toward expanding this work beyond its current scope.

Part II

Neural Representations of Decision Making under Uncertainty: An fMRI Study

Part II Overview

In decision making under uncertainty, neuroscience focuses on expected reward and learning rather than risk (chapter 1). Economic studies emphasize the importance of risk in addition to expected reward (chapter 3). We designed a simple gambling task and conducted a functional imaging study using humans to understand if and how the brain represents and processes both expected and risk in situations of uncertainty. The result is a series of three papers (one published (Preuschoff et al. 2006), two submitted (Preuschoff et al.; Bruguier* et al.)) which are collected in this part of the thesis.

Chapter 5 explores *subcortical* representations of probability and risk. Drawing on financial decision theory, expected reward and risk are modeled as mathematical expectation of reward, and reward variance, respectively. Activations in dopaminoceptive structures correlated with both mathematical parameters. These activations differentiated spatially and temporally. Temporally, the activation related to risk was delayed. Analysis confirmed that the paradigm minimized confounds from learning, motivation, and salience. These results suggest that the primary task of the dopaminergic system is to convey signals of upcoming stochastic rewards, such as expected reward and risk, beyond its role in learning, motivation, and salience. With minor modifications the text in chapter 5 was published in (Preuschoff et al. 2006). The methods have been expanded to provide significantly more details than the original publication. Supplementary materials are included in the main text.

Chapter 6 emphasizes *cortical* representations of risk. Considerations from financial decision theory led us to hypothesize that the brain encodes a prediction risk signal and a prediction risk error signal. An early-onset activation in the human insula correlates significantly with prediction risk error. Its time course is consistent with a role in rapid updating. Activation previously associated with prediction risk emerges with a delay consistent with a role in anticipation. These findings indicate that our understanding of the neural basis of de-

cision making under uncertainty needs to be expanded to include prediction risk estimation. Such integration may have far-reaching implications for our understanding of pathological decision making. In addition, the analysis demonstrates how carefully modeling the temporal aspect of functional imaging data can reveal complex, intertwined processes. With minor modifications the text in chapter 6 has been submitted. The methods provided in this chapter are limited to those parts not used and mentioned in chapter 5. Supplementary materials are included in the main text.

Chapter 7 looks at how the different representations of probability and risk can be combined into a single representation of value. On a behavioral level, expected reward and risk are often competing as a high expected reward may also come with a high risk, i.e., a small chance of not getting anything at all or even losing some. What is the overall value of a gamble after taking into account both expected reward and risk? This chapter addresses the question of where and how these two variables are combined into a single variable that evaluates the overall value of a gamble. Using canonical correlation analysis on the ventral striatum, putamen, and insula signals a new predictor emerges that summed both expected reward and risk. This suggested that contrary to expected utility theory, risk is added to and not subtracted from expected reward to obtain a metric of conflict. The new predictor significantly activates the anterior cingulate cortex (ACC) in accordance with its role in conflict monitoring.

Contributions

Experimental design, data collection, and data analysis for all studies (including pilot studies and final design) were done by me. The results of chapters 5 and 6 inspired the data analysis for chapter 7 for which Tony Bruguier (together with Peter Bossaerts) worked out and applied a canonical correlations analysis capable of exploiting correlations across brain regions. Peter Bossaerts and Steve Quartz coauthored the papers presented in chapters 5, 6, and 7. Tony Bruguier coauthored the paper presented in chapter 7.

Chapter 5

Neural Differentiation of Expected Reward and Risk in Human Subcortical Structures

5.1 Summary

In decision making under uncertainty, economic studies emphasize the importance of risk in addition to expected reward. Studies in neuroscience focus on expected reward and learning rather than risk. We combined functional imaging with a simple gambling task to vary expected reward and risk simultaneously and in an uncorrelated manner. Drawing on financial decision theory, we modeled expected reward as mathematical expectation of reward, and risk as reward variance. Activations in dopaminoceptive structures correlated with both mathematical parameters. These activations differentiated spatially and temporally. Temporally, the activation related to risk was delayed. Analysis confirmed that our paradigm minimized confounds from learning, motivation, and salience. These results suggest that the primary task of the dopaminergic system is to convey signals of upcoming stochastic rewards, such as expected reward and risk, beyond its role in learning, motivation, and salience.

5.2 Introduction

When faced with decision making in an uncertain world, it is fundamental to evaluate both expected rewards and risks. Higher expected rewards are usually preferred over lower expected rewards. But sensitivity to risk is also ubiquitous. For instance, when an investor

has the option of either opening a simple savings account (low expected reward but a known outcome) or investing all of her money into a particular stock (higher expected reward but an uncertain outcome), she may prefer the option with the lower expected reward because of the higher risk of the alternative. Economic studies (Bossaerts and Plott 2004; Holt and Laury 2002) have confirmed that risk considerations, in addition to expected reward, indeed play a role in decision making under uncertainty and in the valuation of risky gambles. This sensitivity to both expected reward and risk is not unique to financial situations. It is also observed in nonhuman primates facing uncertain rewards (Fiorillo et al. 2003; McCoy et al. 2003) and in bees choosing among different flowers (Real 1991).

In neuroscience, evidence has accumulated that brain activation correlates with expected reward. Human fMRI studies have found that subcortical dopaminergic structures such as striatum are involved in reward-related processes. Activity in these structures correlates with reward value of a variety of stimuli, including primary rewards, such as gustatory stimuli (Berns et al. 2001; O'Doherty et al. 2002) and abstract stimuli, such as money (Breiter et al. 2001; Elliott et al. 2000, 2003; Knutson et al. 2001a, 2003, 2000). Studies of nonhuman primate conditioning document a monotonically increasing relationship between phasic activity of midbrain dopamine neurons and reward probability or expected reward (Fiorillo et al. 2003; Tobler et al. 2005).

The correlations found between risk and activation in cortical regions are unambiguous (McCoy et al. 2003; Huettel et al. 2005a, 2006). However, correlations found between risk and activation in subcortical regions are not. In the nonhuman primate brain, delayed firing of dopaminergic neurons was positively correlated with risk when risk was modulated by changing reward probabilities (Fiorillo et al. 2003; Tobler et al. 2005). Correlation was also positive in caudate neurons when risk was modulated by manipulating problem complexity (stimulus recognition uncertainty (Lauwereyns et al. 2002; Takikawa et al. 2002)). For the human brain, however, the findings are inconsistent. When risk was modulated by altering the degree to which one knows the probability of reward, i.e., when manipulating knowledge of the reward probability rather than the probability itself, activity in striatum correlated negatively with risk (Hsu et al. 2005). In contrast, when risk concerned problem complexity (categorization uncertainty (Grinband et al. 2006)), striatal activation correlated positively with risk.

Consequently, to date no studies have examined whether and how subcortical dopaminocep-

tive regions in the human brain code for risk when risk is modulated by changing reward probability. This is the primary type of modulation for decision making under uncertainty, and needs to be understood before modulating knowledge of probabilities and studying learning of probabilities.

Here, we manipulated probabilities so that not only risk changed over the full range, but also expected reward, and in such a way that expected reward and risk varied orthogonally. We then determined whether and how activation in subcortical dopaminoceptive regions correlated with expected reward and risk.

In addition, our study addresses a number of important open issues about the representation of expected reward and risk in subcortical dopaminoceptive regions. First, since neuroscientific studies have examined expected reward for a limited number of values, the precise representation, or mathematical model, of reward expectation in the brain remains unknown. As financial decision theory models reward expectation as mathematical expectation of reward (Knutson et al. 2003), it makes specific predictions regarding the form brain activation must take to represent expected reward. When reward is kept constant across rewarded trials, mathematical expectation of reward increases linearly in the probability of reward. We therefore hypothesized that brain activation increases linearly in reward probability if it is to reflect expected reward. To test this representation hypothesis stemming from financial decision theory, the probability of reward needs to be varied over all probabilities (ranging from $p = 0$ to $p = 1$) with a sufficient number of intermediate values.

Second, the specific form of risk representation in the brain is unknown. As with reward expectation, financial decision theory suggests a specific metric for measuring risk, namely, variance, the mean squared deviation from the expected outcome (Markowitz 1952). When reward magnitude is kept constant across rewarded trials, reward variance is quadratic in reward probability p; variance attains a maximum at $p = 0.5$, and minimums at the extremes, $p = 0$ and $p = 1$ (figure 5.1). Because variance is monotonically increasing for $p < 0.5$ and monotonically decreasing for $p \geqslant 0.5$, care has to be exercised that p varies sufficiently to ensure that the effects of changes in risk and expected reward can be disentangled. Otherwise, an increase in risk may be confounded with a change in expected reward (Critchley et al. 2001; Dreher et al. 2005; Knutson et al. 2003). From a statistical point of view, variance and expected reward become orthogonal if they are varied over the full range of reward probabilities and are mean corrected (i.e., after subtracting their

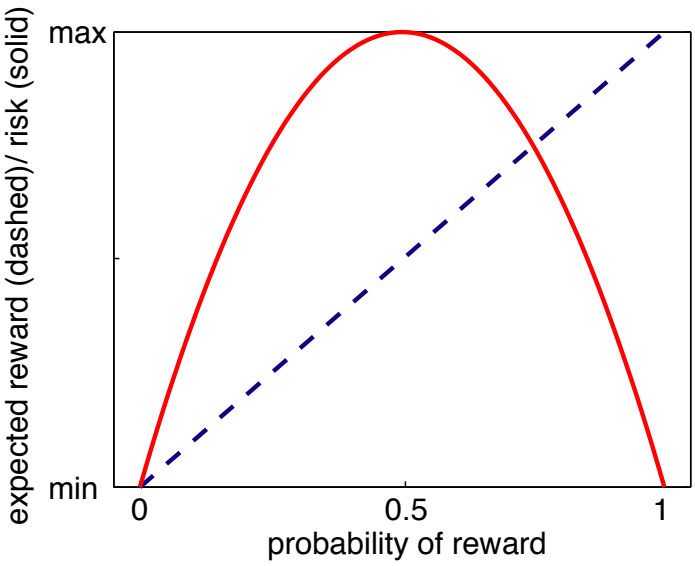


Figure 5.1: Expected reward and risk as a function of the probability of reward. Expected reward, measured as mathematical expectation of reward, increases linearly in the probability of reward p (dashed line). Expected reward is minimal at $p = 0$ and maximal at $p = 1$. Risk, measured as reward variance, is an inversely quadratic function of probability that is minimal at $p = 0$ and $p = 1$ and maximal at $p = 0.5$ (solid line). As such, expected reward and risk are orthogonal over the full range of probabilities, p in $[0, 1]$. When subjects place their bet, the reward probability p is 0.5. After display of card 1, the reward probability changes, depending on whether the subject bet that the second card is higher or lower, and depending on the number on card 1. If the subject bet that the second card is going to be lower, then p increases linearly in the number on card 1; otherwise p decreases linearly in the number on card 1.

average values).

Third, previous neuroscience studies have focused on the learning aspect of reward anticipation (Hollerman and Schultz 1998; Mirenowicz and Schultz 1994; Romo and Schultz 1990), leaving it unclear whether activation related to reward expectation and risk in subcortical structures requires learning and motivation (Knutson et al. 2001a) to be present. Many of these studies have been guided by Temporal Difference (TD) models of learning (Sutton 1988). In the case of risk encoding in dopaminergic neurons, it too has been interpreted in terms of reward learning (Fiorillo et al. 2003; Tobler et al. 2005). The controversy regarding the interpretation of subcortical dopaminergic activation is compounded by studies that suggest that such activation may represent salience (Zink et al. 2004) or nonspecific forms of uncertainty (Aron et al. 2004; Berns et al. 2001). Our hypothesis is that the primary

task of the dopaminergic system is to convey signals of upcoming stochastic rewards, like expected reward and risk, while learning, salience and motivation constitute only secondary, albeit important, tasks. Testing this hypothesis requires disassociating the signaling task from learning, salience and motivation, which requires a perceptual experimental paradigm, unlike the previously employed conditioning tasks. Elimination of learning confounds is especially important because the correlation of sustained activation of dopaminergic neurons with uncertainty (Fiorillo et al. 2003) has been interpreted as the effect of back-propagation of reward prediction errors during learning (Fiorillo et al. 2005; Niv et al. 2005).

Finally, if a single brain system, the dopaminergic system, is to represent two parameters (expected reward; risk) of a single phenomenon (a gamble), the issue of discrimination arises. Discrimination could be achieved spatially, in which different regions could specialize in encoding the different parameters or distinct neural populations within the same region of the brain could encode different parameters. Another possibility is that discrimination could be achieved temporally, in which the same sub-region sequentially encodes the two parameters.

To test how subcortical dopaminergic structures encode these two parameters it is necessary to utilize an experimental design that allows for distinguishing these alternative encoding strategies. Based on recent evidence of activation of dopaminergic neurons in the nonhuman primate brain when risk is modulated by varying probability (Fiorillo et al. 2003), we expected to find an early onset activation in subcortical dopaminoceptive regions that correlated positively with expected reward, while a late onset activation would correlate positively with risk. We therefore allowed for sufficient time between stimulus and outcome and used a statistical analysis of the imaging data that is able to capture potential temporal differentiation.

Nineteen subjects played a gamble where two cards were drawn (without replacement within each trial) from a deck of 10, numbered 1 through 10 (figure 5.2). Before seeing either card, subjects first placed a $1 bet on whether the first or the second card would be higher. Once the bet was placed, subjects saw card 1, followed $\sim$7 s later by card 2. We refer to the time interval between display of card 1 and card 2 as the anticipatory period. Upon display of card 1, the probability of winning changes as a function of the number on card 1. For instance, if the subject bet on "second card higher", the probability of winning is given by the number of cards initially in the deck (always 10) minus the number displayed on the

first card (C) and divided by the number of cards remaining in the deck: $p = (10 - C)/9$.

Since a new deck was used on every trial, subjects had no prior information about the outcome of the gamble, so that on any given trial the initial probability of winning at the time of bet was $p = 0.5$ with maximal risk. As a result, gains and losses were independent of the strategy the subject chose. In addition, there is no role for learning, as any strategy is optimal. Following the presentation of card 2, subjects were asked to report whether they won or lost. In case of an incorrect response subjects lost \$0.25, independent of whether their gamble had paid off. As such, motivation (the degree to which one is willing to work to report whether one won or lost) during the anticipation period should not depend on expected reward or risk.

Reward level was kept constant across all rewarded trials. Because of this, expected reward and risk (variance) upon display of card 1 change only as a function of the probability of winning, as shown in figure 5.1. Altering the reward level would have potentially introduced a confounding factor, namely, varying complexity, which is known to induce activation in subcortical dopaminoceptive structures in itself (Grinband et al. 2006).

5.3 Results

In this section, we first report statistics on task performance. Using a voxel-based analysis, we subsequently document that subjects encoded the task as a reward prediction problem by replicating previously found activation patterns for reward. We then focus on the anticipatory period to find regions of interest (ROIs) whose activity is modulated by expected reward and risk. We determine how activity varies with the probability of reward within the identified ROIs, to verify that activations correlate with mathematical expectation of reward, and reward variance. Finally, we report on tests to determine whether the results were affected by learning, motivation or salience.

5.3.1 Task performance

Participants won on $48.60 \pm 3.95\%$ of all trials and correctly reported the outcome of their bet on $97.8 \pm 2.6\%$ of all trials showing that the gamble was indeed random and subjects kept track of the cards displayed on the screen.

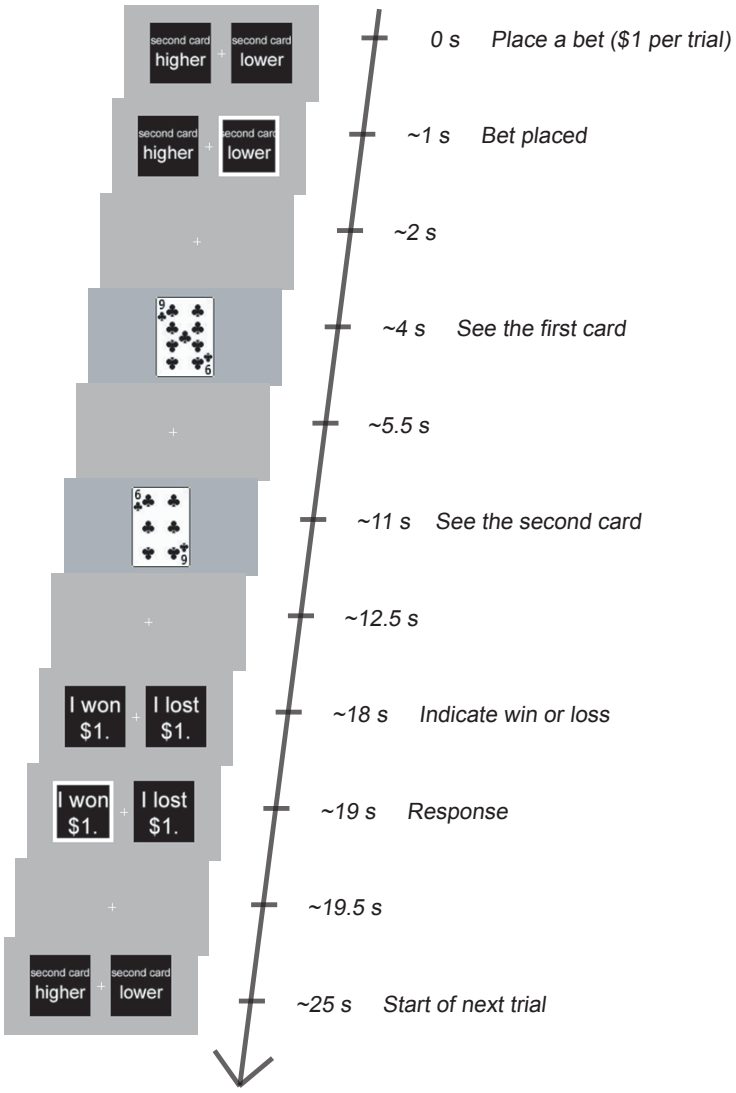


Figure 5.2: On each trial, two cards were drawn (without replacement within each trial) from a deck of 10, numbered 1 through 10. Before seeing either card, subjects first placed a $1 bet on one of two options, "second card higher" or "second card lower" (than first card shown). Subjects could earn $1 if they guessed the right card, and lost $1 if they were wrong. Once the bet was placed, subjects saw card 1, followed ∼7 s later by card 2. At the end of each trial subjects had to indicate whether they won or lost on this trial. A $0.25 penalty was imposed for misreporting, independent of the outcome of the gamble. All times shown are with respect to the onset of the trial.

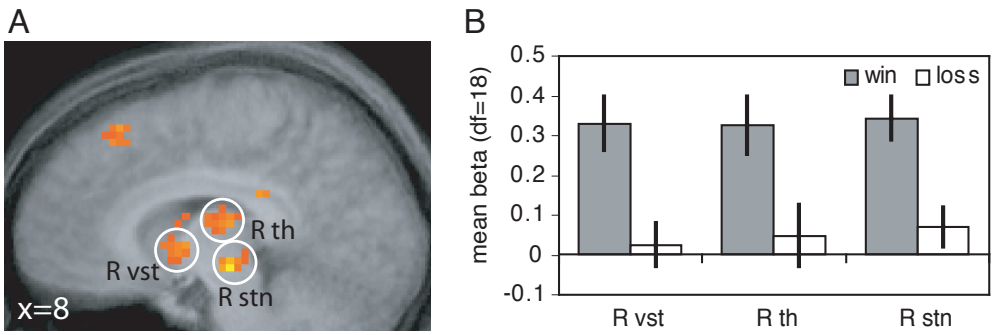


Figure 5.3: Neural activations related to reward (at the display of card 2).

5.3.2 Reward activation

The contrast between wins and losses (i.e., the difference in activation following wins vs. that following losses) revealed significant activation ($p < 0.0001$, figure 5.3) of a subcortical network including caudate, globus pallidus, thalamus and putamen as well as midbrain and cingulate gyrus (table 5.3.2), in agreement with previous reports (Delgado et al. 2000; Elliott et al. 2003; Knutson et al. 2001a, 2003, 2000). The contrast between losses and wins, i.e., "negative reward," revealed no significant activation, which is also supported by prior findings (Knutson et al. 2003). This indicated that subjects were encoding the task as a reward prediction problem and motivated our investigation of decision variables underlying this response.

5.3.3 Anticipatory period activation

Focusing on the anticipatory period, we first used a model to define regions of interest that correlate with expected reward and risk (reward variance) during this period. Based on nonhuman primate evidence that temporally distinct responses of dopaminergic neurons might encode expected reward and risk respectively (Fiorillo et al. 2003), we decomposed the anticipatory period into (i) a response at the onset of card 1 (initial sub-period), followed by (ii) a response until the onset of card 2 (subsequent sub-period). The duration of the initial response (i) was set at only 1s to allow for an onset of the delayed response as early as 1 s after the stimulus. The duration of the subsequent response (ii) was therefore longer ($\sim$6 s), in accordance with the finding in the nonhuman primate brain of sustained neuronal firing that correlates with risk. Because the total length of the anticipatory period is only

region	L/R	mean x	mean y	mean z	cluster size	max stat
lateral dorsal thalamic nucleus	L	−6	−20	15	17	6.21
	R	12	−16	14	63	6.77
subthalamic nucleus	R	14	−18	−2	51	10.64
lateral geniculate	L	−22	−16	−6	50	9.55
caudate head	L	−10	1	1	39	6.72
	R	12	4	2	33	7.03
caudate	R	12	2	13	27	6.09
putamen (posterior part)	R	27	−18	−2	22	6.72
midbrain	L/R	−3	−23	−18	12	6.07
parahippocampal gyrus	L	−33	−36	−9	11	6.73
	R	23	−32	−6	17	6.68
pulvinar	L	−7	−22	−1	15	6.56
posterior cingulate gyrus	L/R	1	−39	29	101	7.2
cingulated gyrus	L/R	−2	−23	28	5	5.8
superior frontal gyrus	L/R	−3	21	49	111	8.54
inferior frontal gyrus	L	−44	22	16	10	5.82
inferior medial frontal gyrus	R	46	18	24	5	6.7
insula	L	−33	−6	11	10	6.12
short insular gyri	L	−26	18	1	8	5.59
precentral gyrus	R	48	1	28	28	6.38
inferior precentral gyrus	L	−46	2	18	14	7.11
middle temporal gyrus	L	−52	−57	4	11	7.49
middle temporal gyrus	L	−33	−73	11	7	9.93
posterior middle temporal gyrus	R	56	−42	9	7	6.04
fusiform gyrus	R	38	−58	−13	12	6.16
precuneus	R	19	−64	22	6	5.82
lingual gyrus	R	4	−85	−3	8	5.97
cerebellum	R	17	−41	−30	9	8.13
cerebellum	L/R	−4	−47	−16	6	6.08

Table 5.1: Activation to reward

$\sim$7 s, whereas hemodynamic responses typically peak at only about 4 s, we did not, however, expect to be able to detect the precise length of the respective responses, and hence, to be able to differentiate between phasic and sustained durations. We therefore only focused on differentiation of the onset of the signal: early (response (i)) vs. later (response (ii)).

We next examined whether the activation we observed conformed to the model of expected reward and risk as specified in financial decision theory. We tested the hypothesis that activation levels relate to reward probability in the way that mathematical expectation of reward and reward variance relate to reward probability (figure 5.1). We changed the specification of our general linear model to compare activation levels at different probabilities within the identified ROIs. To do this, in the new model, one predictor for each individual probability level replaces the predictors for expected reward and risk in the old model.

5.3.4 Modulation of anticipatory period activation by expected reward

Over the initial sub-period (1 s) of the anticipatory period, expected reward was highly correlated with activation in putamen, ventral striatum, globus pallidus, anterior cingulate cortex, midbrain and other regions (figure 5.4 A; table 5.2). We also detected significant activation to expected reward during the subsequent sub-period (6 s) in several foci in the cerebellum and medial temporal gyrus. Although our imaging sequence was not optimized for frontal regions we also found activation in medial orbital gyrus and gyrus rectus (table 5.3).

Based on our a priori hypothesis that subcortical structures encoded expected reward as mathematical expectation of reward, and hence, that activation increased linearly in reward probability, we compared the responses in ventral striatum and putamen for each of the ten reward probabilities that were obtained as a result of the number on card 1 (figure 5.4 B). Activation in bilateral ventral striatum (L vst, R vst) and putamen (L put, R put) showed a linear increase with increasing reward probability; the best linear fit is highly significant and explains a large proportion of the variance of the mean activation levels (L vst: $r^2 = 0.87$, $p < 0.001$; R vst: $r^2 = 0.66$, $p < 0.01$; L put: $r^2 = 0.69$, $p < 0.01$; R put: $r^2 = 0.7$, $p < 0.01$).

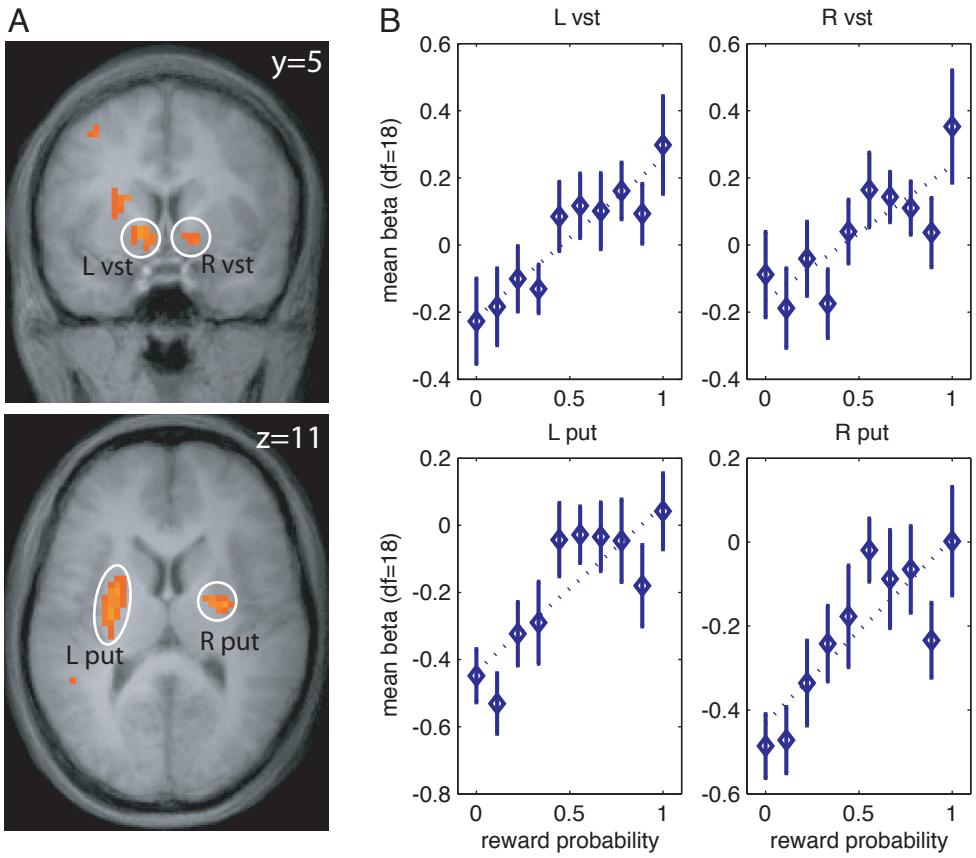


Figure 5.4: Immediate neural correlates of expected reward. A. Neural activations related to expected reward (immediate response within 1 s of display of card 1). Bilateral activity in putamen (L put, R put) and ventral striatum (L vst, R, vst) correlates with the probability of win, and hence, expected reward (random effects, $p < 0.001$). Neural responses are displayed in coronal and axial formats. B. Mean activations (parameter estimates beta with standard error) for ten probabilities. In both left and right ventral striatum (vst) and putamen (put) neural responses increase with increasing probability of win. Dashed line indicates the best linear fit (L vst: $r^2 = 0.87$, $p < 0.001$; R vst: $r^2 = 0.66$, $p < 0.01$; L put: $r^2 = 0.69$, $p < 0.01$; R put: $r^2 = 0.7$, $p < 0.01$).

region	L/R	mean x	mean y	mean z	cluster size	max stat
putamen	L	−26	−9	5	176	7.92
	R	23	−7	11	29	6.01
ventral striatum	L	−12	3	−3	35	5.47
	R	12	5	−3	6	4.45
medial geniculate	R	16	−22	−4	8	4.92
pons	L/R	2	−24	−28	29	5.77
midbrain	L	−5	−19	−16	5	4.43
anterior cingulated	L	−2	32	−2	15	5.41
angular gyrus	L	−45	−59	29	36	6.07
middle frontal gyrus	L	−33	7	47	14	5.07
superior frontal gyrus	R	18	25	50	11	4.86
	L/R	−3	23	53	11	4.36
medial frontal gyrus	L	−12	38	17	7	4.63
superior temporal gyrus	L	−44	−37	4	25	5.88
occipital gyrus	R	22	−75	31	19	5.63
cerebellum	L	−14	−38	−29	14	4.57
	L	−36	−66	−24	8	5.23

Table 5.2: Activation to expected reward (immediate response)

region	L/R	mean x	mean y	mean z	cluster size	max stat
cerebellum	R	23	−82	−21	22	5.39
cerebellum	R	32	−32	−22	11	6.44
lingual gyrus	R	9	−80	−16	10	4.63
middle temporal gyrus	L	−44	8	−24	8	5.35
medial orbital gyrus	L	−24	22	−11	7	5.71
gyrus rectus	L	−5	23	−11	6	4.97
parahippocampal gyrus	L	−16	−23	−13	6	4.74

Table 5.3: Activation to expected reward (delayed response)

5.3.5 Modulation of anticipatory period activation by risk

During the second (6 s) sub-period of the anticipatory period, risk was highly positively correlated with activation in an area extending posterior to and bilateral from the ventral striatum to the subthalamic nucleus as well as mediodorsal thalamic nucleus, midbrain, and bilateral anterior insula (figure 5.5 A). Risk was not significantly correlated over the initial (1 s) sub-period with activation in any of the subcortical regions of interest except for midbrain. Instead, risk correlated significantly over this sub-period with activation in the anterior insula and orbitofrontal cortex (table 5.4). As we are focusing on subcortical structures, we do not elaborate here on the latter finding.

To determine that the risk-related activation over the second (6 s) sub-period of the anticipatory period reflected reward variance, we used the same approach as for expected reward and studied activation separately for each of the ten different reward probability levels (figure 5.5 B). If the activations reflected reward variance, then their relationship with reward probability should be quadratic, with maximum at $p = 0.5$, and minima at $p = 0$ and 1. We found that responses in ventral striatum (L vst, R vst), midbrain (mb), and thalamic nucleus (md) are indeed maximal at intermediate probabilities and minimal at both minimal ($p = 0$) and maximal ($p = 1$) probabilities. Furthermore, the responses in all four regions of interest were shown to correlate with a function that is inversely quadratic (inversely u-shaped) in the probability of winning with a maximum at $p = 0.5$; the best quadratic fit is highly significant and explains a large proportion of the variance of the mean activation levels (L vst: $r^2 = 0.89$, $p < 0.001$; R vst: $r^2 = 0.88$, $p < 0.001$; mb: $r^2 = 0.84$, $p < 0.001$; md: $r^2 = 0.80$, $p < 0.001$). To ensure that the close quadratic fit did not merely result because activation is low at p equal to 0 and 1, while high elsewhere, we performed a standard nonnested hypothesis test (Davidson and MacKinnon 1981) that determined whether a simple model with low activation at $p = 0, 1$ and high for $p \neq 0, 1$ should be rejected in favor of the quadratic model. Across the four brain regions, the quadratic fit was found to be significantly better ($p < 0.001$; results for individual brain regions: L vst: $p < 0.01$; R vst: $p < 0.01$; mb: $p < 0.01$; md: $p = 0.01$).

Finally, if the encoding indeed reflects variance of reward (risk as measured in financial decision theory), we should be able to use activation levels during the anticipatory period to successfully predict activation levels after the bet but before display of card 1. At the

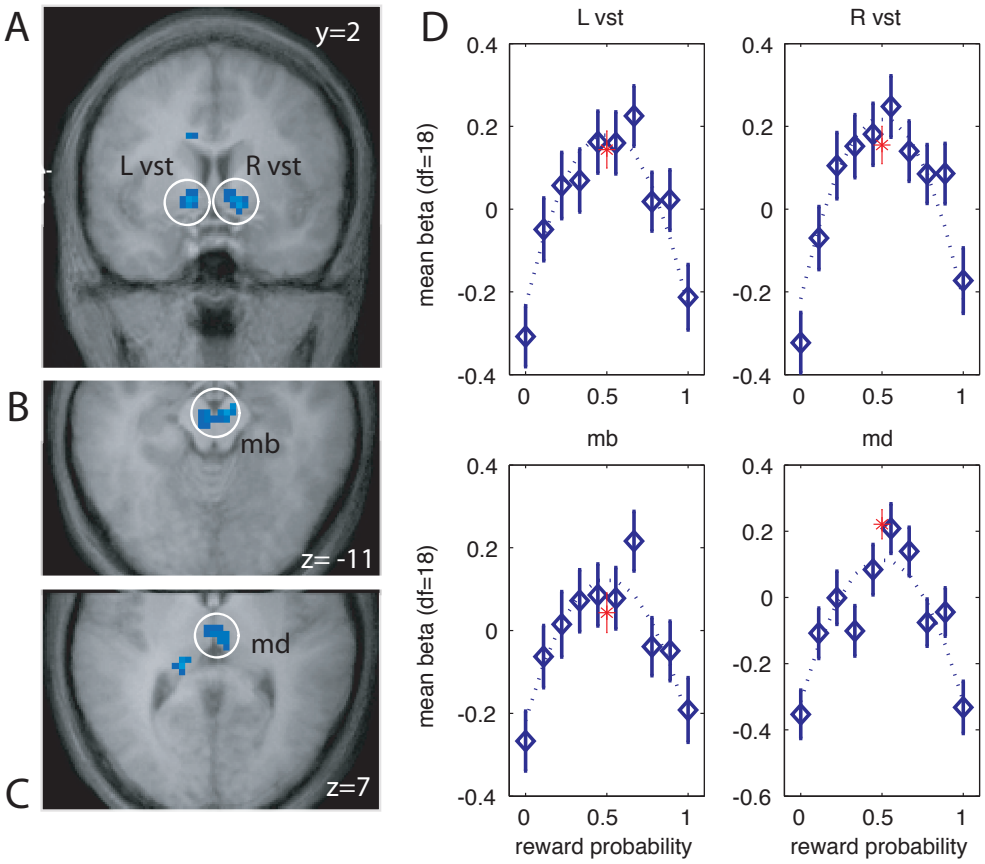


Figure 5.5: Delayed neural correlates of risk. A: Neural activations related to risk (delayed response, after 1 s of display of card 1 and until display of card 2). Brain regions whose activity correlates with reward variance, reflecting risk (random effects, $p < 0.001$), include left and right ventral striatum (L vst, R vst) extending into the subthalamic nucleus, midbrain (mb) and mediodorsal thalamic nucleus (md). B: Mean activations (parameter estimates beta with standard error) for ten probabilities. Neural responses in regions displayed in A increase toward medium probabilities and decrease toward low and high probabilities. Dashed lines indicate best quadratic fit (L vst: $r^2 = 0.89$, $p < 0.001$; R vst: $r^2 = 0.88$, $p < 0.001$; mb: $r^2 = 0.84$, $p < 0.001$; md: $r^2 = 0.80$, $p < 0.001$). Across all four regions, the quadratic fit is significantly better than a model that predicts low activation at $p = 0, 1$ and high activation for $p \neq 0, 1$ ($p < 0.001$; results for individual brain regions: L vst: $p < 0.01$; R vst: $p < 0.01$; mb: $p < 0.01$; md: $p = 0.01$). Red data points (asterisks) at $p = 0.5$ indicate late onset activation levels between the time of bet and card 1 when risk is maximal.

region	L/R	mean x	mean y	mean z	cluster size	max stat
parahippocampal gyrus	L	−17	−29	−17	110	−6.75
	R	18	−22	−14	9	−5.26
transverse temporal gyrus	R	57	−13	10	76	−6.52
	R	38	−25	23	5	−4.94
	L	−53	−8	3	19	−5.16
short insular gyri	L	−32	17	1	68	6.47
short insular gyri	R	34	13	1	9	4.61
midbrain	R	9	−32	−12	44	−8
anterior cingulate	L/R	2	22	−6	14	−4.95
supramarginal gyrus	R	35	−36	45	9	−5.52
	R	40	−26	49	5	−5.06
superior frontal gyrus	L	−9	−29	47	7	−4.61

Table 5.4: Activation to variance (immediate response)

region	L/R	mean x	mean y	mean z	cluster size	max stat
short insular gyri	L	−30	21	9	72	−6.59
	R	31	24	9	11	−4.25
ventral striatum	L	−10	−3	−3	45	−6.59
	R	12	−3	−3	42	−5.76
mediodorsal thalamic nucleus	L/R	1	−16	5	27	−5.18
substantia nigra	L/R	1	−18	−11	17	−5.58

Table 5.5: Activation to variance (delayed response)

time of bet, the probability of win is $p = 0.5$, and risk is maximal (figure 5.1). Therefore, the activation level in ventral striatum must be similar to the levels for $p = 0.5$ during the anticipatory period. Figure 5.5 B shows the level of activation in the same ROIs after the bet and before card 1 is shown. This activation level falls into the confidence interval of when reward variance is maximal. Since the periods before card 1 and 2 are not identical in length, the additional data points have to be evaluated carefully. Nonetheless, they provide corroborating evidence of the hypothesis that these ROIs reflect reward variance.

5.3.6 Modulation of anticipatory period activation by both expected reward and risk

When simultaneously mapping the activation clusters reported for expected reward (activation over the initial 1 s sub-period) and risk (activation over the subsequent 6 s sub-period), a region in left ventral striatum emerged where the clusters overlap (figure 5.6 A-C). We defined this as a region of interest to determine how the different levels of expected reward and risk were reflected in the time courses. We compared average adjusted hemodynamic responses to card 1 for low, medium and high expected reward (figure 5.6 D) and low, medium and high levels of risk (reward variance; figure 5.6 E). Early during the anticipatory period the hemodynamic response increased with the level of expected reward, whereas starting from about 4 s after the onset of the anticipatory period (when card 1 is displayed), the hemodynamic response increased with the level of risk.

5.3.7 Testing for absence of learning, motivation, and salience confounds

As we were interested in determining whether there were subcortical activations that were related to decision-making parameters independently of their previously documented role in reward prediction learning, our task was designed so that learning would not improve the potential outcome of the gamble. Nonetheless, it may be possible that learning-related signals are generated during the task, particularly if the reward prediction learning role of these structures is primary. To test this possibility, we probed our behavioral and imaging data for evidence of learning, salience, and motivation. We distinguished between switch and stay trials, defining a switch trial as one in which a subject chooses a different bet than in the previous trial; in a stay trial, the subject chooses the same bet. For instance, if a

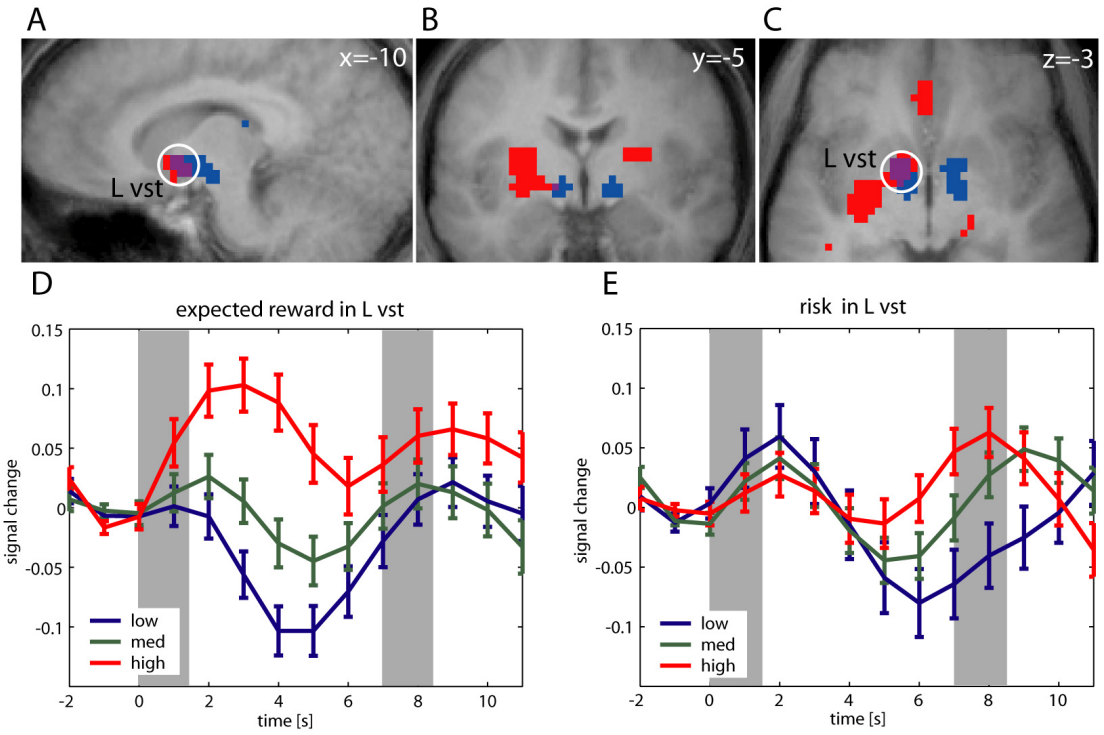


Figure 5.6: Temporal encoding of expected reward and risk. A-C: Immediate (within 1 s of display of card 1) activations related to expected reward (probability of win; red) and delayed (after 1 s of display of card 1) activations related to risk (blue), superimposed on a mean anatomical image. Activations are identical to those shown in figure 5.4 and figure 5.5 but are not pseudo color coded in this map. A, C: Spatial relationship between encoding of expected reward and risk include overlapping regions (displayed in purple and circled) in ventral striatum (vst) and spatially contiguous areas. D, E: Averaged adjusted time courses showing different temporal patterns for representations of expected reward and risk during the anticipatory period ($t = 0, 1, \ldots, 7$ s) in the same sub-region of left ventral striatum (circled purple region in A, C). Separation of time courses for low, medium and high expected reward trials peaks early in the anticipatory period. Separation of time courses for low, medium and high risk trials starts later and peaks around the time card 2 is shown. Gray bars indicate the presentation of card 1 ($t = 0$ s) and card 2 ($t \sim 7$ s).

subject who chose second card higher on the previous trial chooses second card lower in the current trial, then the current trial is a switch trial.

Learning would imply that the likelihood of switching increases after a loss trial, while the likelihood of staying increases after a win trial. There was, however, no significant difference between the number of switches after loss trials vs. win trials or the number of stays following win trials vs. loss trials (figure 5.7 A). We do find an (insignificant) tendency to stay rather than switch regardless of outcome, which is consistent with previous reports on the status quo bias in decision making under uncertainty ((Samuelson and Zeckhauser 1988); it should be added, however, that in our paradigm the status quo bias does not lead to inferior performance; any strategy is optimal, as pointed out before).

Learning would also imply that reaction times to placing the bet would be effected by previous outcomes. Reaction times to placing the bet corroborate the absence of learning effects: reaction times are approximately equal whether the previous trial generated a win or a loss (figure 5.7 B) and they do not differ significantly across stay and switch trials (figure 5.7 C).

To determine whether learning could have caused the reported activations in subcortical structures for expected reward and risk over the two sub-periods of the anticipatory period, we included a variable in our general linear model that indicated whether the immediately preceding trial generated a loss or a gain. Under TD learning, activation for expected reward should be significantly correlated with this indicator variable (while activation for risk could be explained as an effect of back-propagation of prediction errors (Niv et al. 2005; Fiorillo et al. 2005)). We found no significant effect of prior-trial outcome in the regions where activation was found to be reflecting expected reward. The lack of an effect as predicted by TD learning indicates the absence of a learning confound.

We also examined reaction times to determine whether motivation or salience affected our results during the anticipatory period (figure 5.7 D). One could legitimately be concerned that higher expected reward induces higher motivation, while higher risk induces higher salience. The penalty to false reporting of the outcome after the anticipatory period is independent of the outcome, while no reward is given to correct reporting of the outcome. Since both expected reward and risk (reward variance) are related to reward probability, we verified the absence of motivational and salience confounds by plotting reaction times to outcome reporting against reward probability. Figure 5.7 D confirms that there is indeed

no relationship.

While the data presented in figure 5.7 are pooled over all subjects the results reported also hold on an individual subject basis (see also table 5.6). Specifically, differences in strategy (switch or stay) after wins vs. losses are not significant in 17 of 19 subjects ($p > 0.05$, not corrected for multiple comparison). There are no significant differences ($p > 0.05$) for any subject for reaction times after win vs. loss trials and for switch vs. stay trials. No significant linear or quadratic relationship between reaction times and probability of win emerges for any subject. Where applicable we also tested for long-term effects of learning and found no significant results. We see a general trend of decreasing reaction time over time, which does not affect the results reported in figure 5.7.

5.4 Discussion

By utilizing a design in which expected reward and risk as measured in financial decision theory varied orthogonally and across the full range, we tested whether activation in human primary projection targets of midbrain dopaminergic neurons significantly correlated with these two critical decision-theoretic parameters. Further, the paradigm was designed to minimize potential confounds from learning, motivation or salience, allowing us to determine whether these target areas encode expected reward and risk, the primary parameters of financial decision theory, beyond their established role in learning.

We found that during reward anticipation, initial activation in ventral striatum and other subcortical dopaminoceptive structures varied with expected reward, whereas subsequent activation in ventral striatum varied with risk. Activations correlating with expected reward and risk were thus differentiated both spatially and temporally and arose in the absence of learning, motivation, or salience confounds.

5.4.1 Expected reward is reflected in linear response to probability

The response of the ventral striatum and other subcortical structures is highly linear in reward probability. This provides strong support that the phasic responses of ventral striatum and putamen encode the expected reward parameter of financial decision theory and as such goes beyond the monotonicity of encoding expected reward shown in previous studies.

The interpretation of the early response as encoding expected reward is also consistent

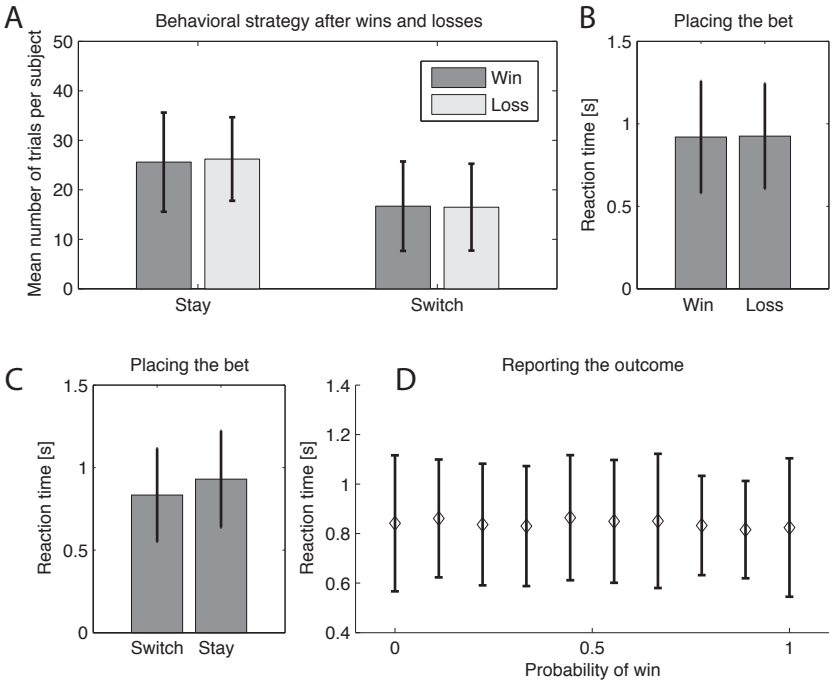


Figure 5.7: Relationship between subject behavior and outcome history. A: Histograms of choice behavior as a function of the outcome (win, loss) in the previous trial. Stay refers to a trial where the subject bets the same as in the previous trial; Switch refers to a trial where the subject bets differently from the previous trial (if subject bet that 2nd card is higher in previous trial, then subject bids that 2nd card is lower in current trial, and vice versa). Evidence of learning would emerge if subjects tend to switch more after a loss, and tend to stay after a win. B: Mean reaction time from trial start to placement of the bet as a function of outcome (win, loss) in prior trial. Evidence of learning would emerge if reaction times tend to be shorter after gains than after losses. C: Mean reaction time from trial start to placement of the bet as a function difference in choice between current and previous trial (switch, stay). Evidence of learning would emerge if reaction times tend to be shorter for stays than for switches. D: Mean reaction times from display of card 2 to reporting of outcome as a function of probability of reward as of display of card 1. Evidence of motivation would emerge if mean reaction times decrease in reward probability. Evidence of salience would emerge if mean reaction times are maximal for maximum risk ($p = 0.5$) and minimal for minimum risk ($p = 0, 1$). See table 5.6 for individual results.

	After win	After loss
Prefer to stay	0	2
Prefer to switch	2	0
Indifferent ($p > 0.05$)	17	17

	After win	After loss
Slower reaction time	0	0
Faster reaction time	0	0
No difference ($p > 0.05$)	19	19

	When switching	When staying
Slower reaction time	0	0
Faster reaction time	0	0
No difference ($p > 0.05$)	19	19

	As probability of winning increases ...
Reaction time increases	0
Reaction time decreases	0
No change in reaction time ($p > 0.05$)[a]	19

[a]Test for both linear and quadratic changes in probability of winning.

Table 5.6: Relationship between subject behavior and outcome history (individual subjects' results). Number of subjects (out of 19) for which the results in figure 5.7 were significant. Differences in strategy (switch or stay) after wins vs. losses are not significant in 17 of 19 subjects ($p < 0.05$), not corrected for multiple comparison). There are no significant differences ($p < 0.05$) for any subject for reaction times after win vs. loss trials and for switch vs. stay trials. No significant linear or quadratic relationship between reaction times and probability of win emerges for any subject. Where applicable we also tested for long-term effects of learning and found no significant results. We see a general trend of decreasing reaction time over time, which does not affect the results reported in figure 5.7.

with theoretical Temporal Difference (TD) models (Montague et al. 1996; Montague and Sejnowski 1994), which have primarily guided the investigation of dopaminergic structures (Knutson et al. 2003), though our results reveal that this signal is generated even in the absence of learning.

5.4.2 Reward variance (risk) is reflected in quadratic response to probability

The quadratic relationship between reward probability and the late response in ventral striatum and other subcortical structures supports the hypothesis that risk is encoded as reward variance in the brain. Variance, however, is just one of several measures of uncertainty that are all maximal at $p = 0.5$. Within neuroscience, entropy (equals minus the weighted sum of the logarithm of the probabilities of each possible outcome), is the most common measure of uncertainty and has been used extensively in information-theoretic analysis of spike trains (Bialek and Rieke 1992). One interpretation based on nonhuman primate electrophysiology in ventral tegmental area has suggested that the sustained response may be encoding entropy (Fiorillo et al. 2003), since entropy is also maximal at $p = 0.5$. Closer inspection of the data demonstrates, however, that the sustained firing of dopaminergic neurons actually correlates with magnitude (Fiorillo et al. 2005). As such, variance (which is sensitive to both probability and magnitude), not entropy, is the right measure of risk. This is consistent with financial decision theory. Financial decision theory sometimes uses additional risk metrics (skewness, kurtosis, etc.), but these appear to be unnecessary to explain valuation when risk is as small as in our experiments (Bossaerts and Plott 2004).

5.4.3 Separation of expected reward and risk through spatial and temporal differentiation

Though some subcortical regions are responsive exclusively to either expected reward or risk, others are responsive to both parameters, leading to the issue of how these signals are differentiated. Our results indicate that the brain differentiates these signals temporally: the initial response reflects expected reward; the subsequent response reflects risk. The distinct hemodynamic responses to expected reward and reward variance in human dopaminoceptive

structures follow a pattern remarkably consistent with that found in nonhuman primate ventral midbrain using electrophysiology and reporting an immediate, phasic encoding of expected reward and a late-onset sustained encoding of risk (Fiorillo et al. 2003). The late onset happened to be too late relative to the duration of the anticipation period, however, to discriminate between phasic and sustained responses as well as one has been able to do in electrophysiological studies. Our results suggest that the downstream effects of temporally differentiated activation in the ventral midbrain result in both an early onset separation of signals correlating with expected reward and a late onset separation of signals correlating with risk in the target (dopaminoceptive) structures.

It is interesting to note that, with the exception of the midbrain, we failed to find an immediate activation in subcortical structures that correlated with risk. Activation in subcortical dopaminoceptive areas that correlated with risk is invariably delayed. This raises a number of issues worth investigating in future research. Is there an immediate signal for risk elsewhere in the brain (our data suggest that insula may play a role)? If so, how are the signals of expected reward and risk combined in order to guide decisions? What is the role of the delayed risk signal in subcortical dopaminoceptive areas? Is it used to improve learning, as suggested in (Tobler et al. 2005)? Interestingly, the late-onset activity in ventral striatum looks similar to delayed activity in parietal cortex reported in (Huettel et al. 2005a, 2006).

5.4.4 Brain activation decomposes along basic financial parameters of monetary gambles

Our investigation was guided by the mathematical model of decision making under uncertainty stemming from financial decision theory. This model specifies the minimal parameters that are necessary for rational choice under uncertainty (expectation and variance of reward). Our study shows that brain activity correlates with these two parameters. In financial decision theory, expectation is balanced against variance, and this trade-off has led to important insights not only about simple animal behavior (e.g., bee foraging (Real 1991)), but also about complex human activity, such as the demand for money and its relation to yields on fixed-income securities (Tobin 1958), or the demand for and pricing of multiple risky securities. For instance, Sharpe (1964), Lintner (1965) and Mossin (1966) demonstrated that expected returns on risky securities should increase not as a function of

their own risk (variance), but only to the extent that they contribute to the risk (variance) of the securities market as a whole. Experiments confirm these predictions (Bossaerts and Plott 2004). Later, Black and Scholes (1973) showed that prices of options (to purchase or sell securities) increase as a function of risk again measured by variance. It is striking that brain activation at the level of subcortical dopaminergic structures reflects the separation of expected reward and risk on which financial decision theory is based.

5.4.5 Objective perception independent of choice, learning and attention

As our results are obtained under purely perceptual conditions, i.e., when no choice is to be made subsequently, the activations we report are related primarily to the assessment of risk and reward in gambles. Many levels of processing intervene between perception and choice, so it is possible that the brain tracks expected reward and risk at the perceptual level, while additional elements, such as contextual factors (e.g., decisions by others (Abel 1990)), modulate choice. As such, perception of reward and risk may continue even if choice is not affected (Bayer and Glimcher 2005). Absent subsequent choice, brain activity may merely reflect information gathering for the case that a choice opportunity would suddenly and unexpectedly arise.

It is important to point out that our goal was to investigate the perception of risk and reward and whether such perception conformed to a specific mathematical model. Our finding that at the perceptual level the brain conforms to this model is entirely consistent with the fact that there may also be subjective representations of decision-making parameters that vary from this model, as they may simply be different levels of representation or generated under different contexts. It is an important issue for future research to investigate when and where subjective representations of these parameters may also be generated and how these signals may be integrated or interact in the generation of choice behavior.

Likewise, our task does not involve conditioning. Behavioral data support the absence of conditioning, and statistical analysis of the brain activation confirms the absence of conditioning confounds. Conditioning paradigms allow one to shed light on factors such as learning, motivation or salience, in addition to perception (Elliott et al. 2003; Knutson et al. 2003; Critchley et al. 2001; Rustichini et al. 2005; Ernst et al. 2004; Zink et al. 2004). Here we show how expected reward and risk correlate with activation in subcortical dopaminoceptive structures when these additional elements are removed.

5.4.6 Implications for pathological decision making under risk

Pathological behaviors ranging from addiction to gambling (Bechara et al. 1997) as well as a variety of mental illnesses, such as bipolar disorder (Minassian et al. 2004), and schizophrenia (Shurman et al. 2005), are partially characterized by risk taking. To date, it is unknown whether such pathological decision making under risk is due to misperception of risk or disruptions in cognitive processes, such as learning, planning, and choice. For example, a bipolar subject during a manic episode may invest in a risky business proposition either because they misperceive the risk to be lower than it actually is, or because they accurately perceive the risk to be high but may have impaired learning, attentional, working memory, or choice processes. To date, studies of pathological decisions making under risk have primarily utilized the Iowa Gambling Task (IGT) (Cavedini et al. 2002; Clark et al. 2001; Shurman et al. 2005), which was designed to assess sensitivity to future negative reinforcers (Bechara et al. 1997). Recent studies (Dunn et al. 2006; Maia and McClelland 2004), however, suggest that impaired performance on the IGT may be due to impairments in reversal learning, working memory, attentional shift, and related high-level cognitive processes rather than misperceptions of risk per se. Since our task was designed to minimize the involvement of these high-level processes, in the future it may be utilized with clinical populations to determine whether alterations in risk perception accompany their changes in risky behavior. This may lead to a better understanding of the relative contributions of risk misperception vs. cognitive impairments in these pathological cases, may suggest different treatment approaches, and may also gauge the impact on and the feedback from higher-level brain regions known to contribute to decision making (e.g., ventromedial prefrontal cortex (Fukui et al. 2005)).

5.4.7 Conclusion

In neuroscience, the investigation of the dopaminergic system in tasks involving uncertainty has emphasized reward prediction learning. Risk perception has been less well studied, yet it is central to decision making under uncertainty, as formalized in financial decision theory. Our results show that brain activity can be separated, both spatially and temporally, into signals that correlate with (mathematical) expectation of reward, and with reward variance (risk) - two fundamental parameters of financial decision theory. The role of

human subcortical dopaminoceptive structures in reward-related processing is extended even further as our results suggest that these structures convey basic signals of upcoming stochastic rewards, like expected reward and risk, beyond these structures' role in learning, salience and motivation.

5.5 Methods

A total of 19 subjects participated in the study (10 male, 9 female; aged 18-30, mean age 21.4 years). All subjects gave full informed consent to participate in the study. The study was approved by the California Institute of Technology Institutional Review Board.

5.5.1 Experimental paradigm

Each subject was given written instructions for the game and completed a brief training session outside the magnet. For each session, subjects were provided with an initial endowment of $25. If no bet was placed, they lost automatically. They also lost $0.25 if they incorrectly reported the outcome of their bet or if they did not respond. Accumulated gains were shown only at the end of each session. Subjects played 3 sessions with 30 trials per session. At the end of the experiment, subjects selected one of the three sessions at random, which determined their final payoff. During scanning, trials were randomly ordered. The first card was pseudorandomized such that each of the 10 cards would be seen 3 times as the first card. The second card was randomized.

5.5.2 Imaging data

Each scanning session included a localizer scan and T1 weighted MPRAGE anatomical scans (256x256 matrix, 176 1mm sagittal slices) followed by the acquisition of functional images while subjects performed the gambling task. Images were acquired using a Siemens TRIO 3.0T full body MRI scanner using T2*-weighted PACE EPI. For each subject three functional runs were collected (392-400 scans each). Imaging parameters for both anatomical and functional data acquisition are summarized in table 5.7.

Visual stimuli were presented using Matlab Psychophysics toolbox. A fixation cross was present on screens with no other stimuli. Cards were presented in the center of the screen (screen resolution 600x800, card size 73x97). A single suit (spades) was used.

Parameter	Anatomical	Functional
Sequence name	mprage	EPI Pace
Slice orientation	sagittal	axial
Number of slices	176	32
Slice thickness [mm]	1	3
Slice resolution [mm^2]	1×1	3.28125×3.28125
Slice resolution [px^2]	256×256	64×64
FOV [mm]	256	210
FOV [%]	100	100
Phase oversampling [%]	0	13
Slice oversampling [%]	9	n/a
Flip angle [deg]	10	90
Bandwidth [Hz/px]	n/a	2694
TR [ms]	1500	2000
TE [ms]	3.05	30
TI [ms]	800	n/a
Acquisition sequence	n/a	interleaved
Phase encoding direction	A $\rightarrow$ P	n/a

Table 5.7: Imaging parameters

5.5.3 Behavioral Data

Throughout the experiment all responses and corresponding reaction times were recorded. The correct timing of unprompted or late button presses was not tracked, however their occurrence and the trial in which they occurred was.

5.5.4 Data processing

Data were processed and analyzed using BrainVoyager v1.26 for Windows running on a Sony Vaio PCG-GRX600P. Preprocessing included motion correction (six-parameter rigid body transformation), slice timing correction, linear drift removal, highpass filtering, normalization to Talairach space, and spatial smoothing with a full width at half-maximum Gaussian kernel of 8 mm.

Coregistration. For each subject functional and anatomical data was aligned using anatomical to anatomical mapping. For this an anatomical volume was created from the first experimental session by averaging over all functional volumes within that session and inverting intensities. The resulting alignment was checked visually and corrected if necessary. The transformation matrices can be used to map activation from all three sessions onto the anatomical volume.

Transformation to Talairach space. Anatomical data was transformed to Talairach space by rotating the anatomical volume into the ACPC plane and applying a standard 9-parameter landmark transformation.

For each subject, a separate linear model was constructed that included the regressors described below as well as visual and motor activation. Regressors modeled the BOLD response to the specified events using a convolution kernel applied to a boxcar function. A first-order autoregressive model was used to correct for temporal autocorrelations. For each subject, contrasts were calculated at every voxel in the brain. In a random-effects analysis, a one-sample t test determined where the average contrast value for the group as a whole (n = 19 subjects) differed significantly from zero. Statistical maps were thresholded for significance ($p < 0.001$) and cluster size ($\geq$5 voxels). The model used to identify regions of interests decomposed the anticipatory period into two consecutive epochs: a short epoch (1 s from card 1) followed by a long epoch (6 s) modeling the remainder of the anticipatory period until card 2. Both epochs were modeled with three predictors, a 0th, 1st, and 2nd order term. The predictor for the 0th order term modeled the anticipatory period. The 1st and 2nd order terms modeled the same period, but predicted a hemodynamic response that scaled linearly and quadratically with the probability of win. Note that all three components were orthogonal with respect to one another. The model also included predictors for visual and motor activation as well as for wins and losses at the time of card 2. The different regressors are summarized in the next section.

To compare activation levels at different probabilities (beta estimates in figures 5.4 and 5.5) within the identified regions of interest, the model was modified to include one predictor for each individual probability instead of the 0th, 1st, and 2nd order terms. For determining risk activation levels, a late-onset predictor between the bet and card 1 was also included to model the (maximal) risk at that time. Adjusted time courses (figures 5.6 D and E) are

time courses corrected for the effects (confounds) in the reduced model (for probability time course, reduced model = full model predictor for 1st order term [probability]; for risk time course, reduced model = full model 2 predictor for 2nd order term [risk]). This ensured that the effects shown were orthogonal to all effects captured by the reduced model. Any effect not included in the reduced model would show up in the adjusted data (error term), while any effect included in the reduced model should not show up. For the probability time course, adjusted data were grouped into low ($p < 0.3$), medium ($0.3 < p < 0.7$), and high ($p > 0.7$) probability trials, averaged over all trials (time-locked to card 1). This event-related average was plotted over time for each group. For the risk time course, adjusted data were grouped into low ($p = 0$ or $p = 1$), medium ($0 < p < 0.3$ or $0.7 < p < 1$), and high ($0.3 < p < 0.7$) risk trials, averaged over all trials (timelocked to card 1). This event-related average of residuals was plotted over time for each group.

Chapter 6

Human Insula Activation Reflects Risk Prediction Errors as well as Risk

6.1 Summary

All decision-makers confront many forms of uncertainty when making adaptive choices. One form occurs when stimulus-reward associations are probabilistic and changing. Understanding how organisms deal with this source of uncertainty has been advanced by a convergence between reinforcement learning models and primate physiology, which demonstrated that the brain encodes a reward prediction error signal. However, organisms must also track the level of risk associated with predictions of probabilistic reward, monitor the errors in those risk predictions, and update these in light of new information. To date, it is not known whether, or how, the brain accomplishes this. Considerations from financial decision theory led us to hypothesize that the brain encodes a prediction risk signal and a prediction risk error signal. Using functional imaging during a simple gambling task, here we show that an early-onset activation in the human insula correlates significantly with prediction risk error and that its time course is consistent with a role in rapid updating. Further, we show that activation previously associated with prediction risk emerges with a delay consistent with a role in anticipation. Our findings indicate that our understanding of the neural basis of decision making under uncertainty needs to be expanded to include prediction risk estimation. Such integration may have far-reaching implications for our understanding of pathological decision making.

6.2 Introduction

Many people listen to weather reports to make predictions about their environments. To choose the right clothing, people must predict not only the average temperature but also how much it will vary throughout the day. In other words, they need to estimate the prediction risk (variance) of the predicted average temperature. In addition, if the prediction risk is misjudged, prediction risk errors arise with which to correctly assess future prediction risk. While this is a simple everyday example, estimation of prediction risk is mathematically sophisticated, as evidenced by recent advances in the theory of finance, where prediction of risk is crucial to correctly value securities with complex future payoffs such as options (Engle 2002, 1982). In fact, keeping track of prediction risk errors and thereby correctly estimating prediction risk is crucial for survival of any organism that must adapt to a changing, uncertain environment.

We hypothesized that there exists a system in the brain that encodes both prediction risk error and prediction risk, and that such a system plays the same role for prediction risk and prediction risk error that subcortical dopamine projection areas play for reward prediction and reward prediction error. We predicted that this system may be insula, as insula activation correlates with a broad range of risk-related characteristics of gambles involving probabilistic rewards, such as complexity (Huettel et al. 2005a; Grinband et al. 2006) ambiguity (Hsu et al. 2005; Huettel et al. 2005b, 2006), and uncertainty (Critchley et al. 2001; Paulus et al. 2003; Elliott et al. 2000; Ernst et al. 2002).

6.3 Task

To test our hypothesis, nineteen subjects played multiple rounds of a simple gambling task (figure 6.1) while their brain activity was recorded using functional magnetic resonance imaging (fMRI). In each round two cards were drawn from a randomly shuffled deck of ten. Before seeing either card, players guessed if the second card would be higher or lower than the first. We then displayed the first card followed $\sim$7 s later by the second card. Within a round, predictions occur before the first and second card, which result in prediction errors at both cards. This is illustrated for three exemplary trials in figure 6.1. When guessing which card would be higher, the player has an estimate of the reward to be expected after seeing the first card and of the corresponding prediction risk. Once the first card is revealed,

these estimates are compared to the actual values, resulting in reward prediction errors and prediction risk errors. The player then estimates the reward at the second card and corresponding prediction risk, which again results in errors once the card is revealed (see section 6.6 for formal definitions). We hypothesized a response to prediction risk errors and to prediction risk in insula following both cards.

6.4 Results

The prediction risk error following both the first and second card indeed correlated significantly with activity in bilateral anterior insula (figure 6.2 A). Increasing prediction risk errors were reflected in increasing activitation in right insula (figure 6.2 B, $p < 0.001$, $r^2 = 0.76$). Furthermore, changes of activation intensity as a function of prediction risk error did not depend on whether they occurred after the first card, or after the second card. I.e., knowing the activation level at the first card, one can predict the activation level at the second card from the prediction risk error.

Previous studies have investigated prediction risk and have found an association with activation in insula. The time courses in these studies suggest that the onset of the risk-related activation is delayed with respect to the risk onset (Huettel et al. 2005a). In accordance with these results, we found late-onset prediction risk signals in bilateral insula areas. These areas were slightly more superior and anterior to those found for prediction risk error (figure 6.3).

In summary, we found two signals in bilateral anterior insula, one for prediction risk error and one for prediction risk. The two signals were not only spatially separated (figure 6.3 top) but also temporally (figure 6.3 bottom). Different levels of prediction risk error were best dissociated immediately after a card is shown, while different levels of prediction risk were best dissociated after a delay.

6.5 Discussion

These findings support our hypothesis that there are two signals in anterior insula, a fast-onset prediction risk error signal and a late-onset prediction risk signal. While previous studies have documented risk-related activation in insula (Huettel et al. 2005a,b, 2006;

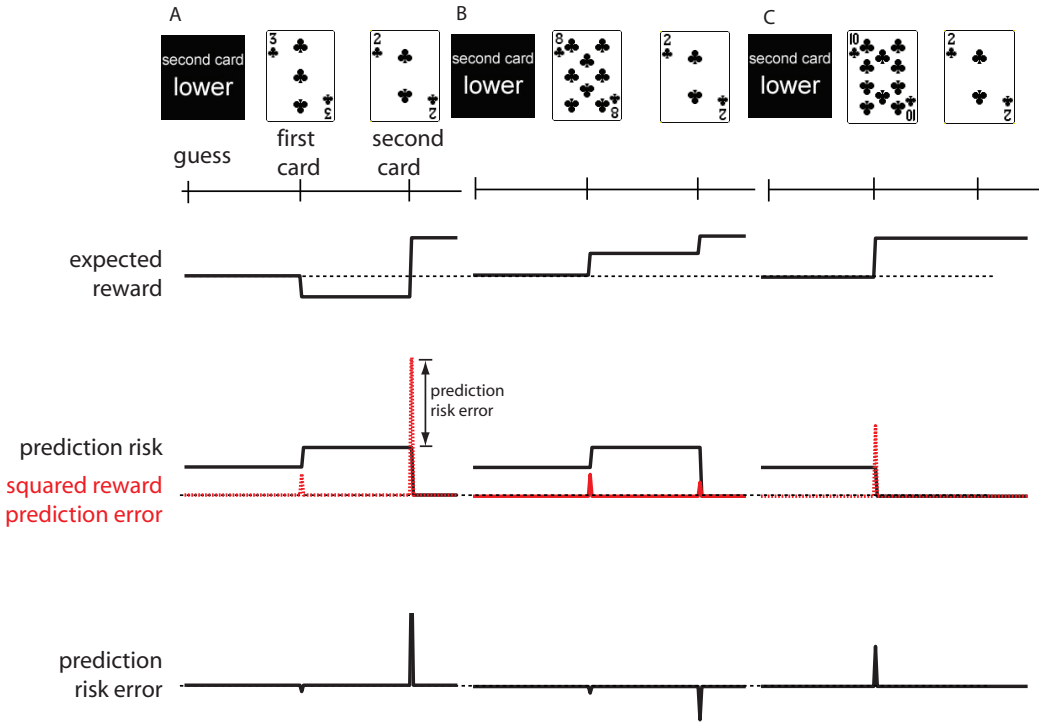


Figure 6.1: Expected reward, prediction risk, and prediction risk errors throughout typical trials. Within a trial, predictions and prediction errors arise twice: when the first card is revealed; and when the second is revealed. As there are 10 cards, numbered 1-10, the expected reward before the first card is \$0; the expected number on the first card is 5.5. Note that this number will never equal the number on the first card, i.e., there will always be a prediction error and it will at least be 0.5. One can compute how large this prediction error will be on average when the first card is shown, by summing the expected rewards for the different possible numbers on the first card multiplied by the probability that these numbers occur. The difference between this average and the actual prediction error results in a prediction risk error. A: The subject guesses that the second card will be lower. The first card drawn is a 3. The second card is a 2 hence the subject wins \$1. The reward prediction error at the first card (red dashed line, second panel) is a little smaller than expected (prediction risk is the expected size of the reward prediction error); therefore the prediction risk error is negative. At the second card, the reward prediction error is much smaller than expected. Therefore, the prediction risk error is positive and large. B: The subject guesses that the second card will be lower. The first card drawn is a 8. The second card is a 2 hence the subject wins \$1. The reward prediction error at the first card is smaller than expected; therefore the prediction risk error is negative. At the second card, the reward prediction error is much smaller than expected. Therefore, the prediction risk error is negative and large. C: The subject guesses that the second card will be lower. The first card drawn is a 10. The second card is a 2 hence the subject wins \$1. The reward prediction error at the first card is much larger than expected; therefore the prediction risk error is positive and large. As no uncertainty remains about the outcome after card 1, there are no prediction errors at the time of card 2.

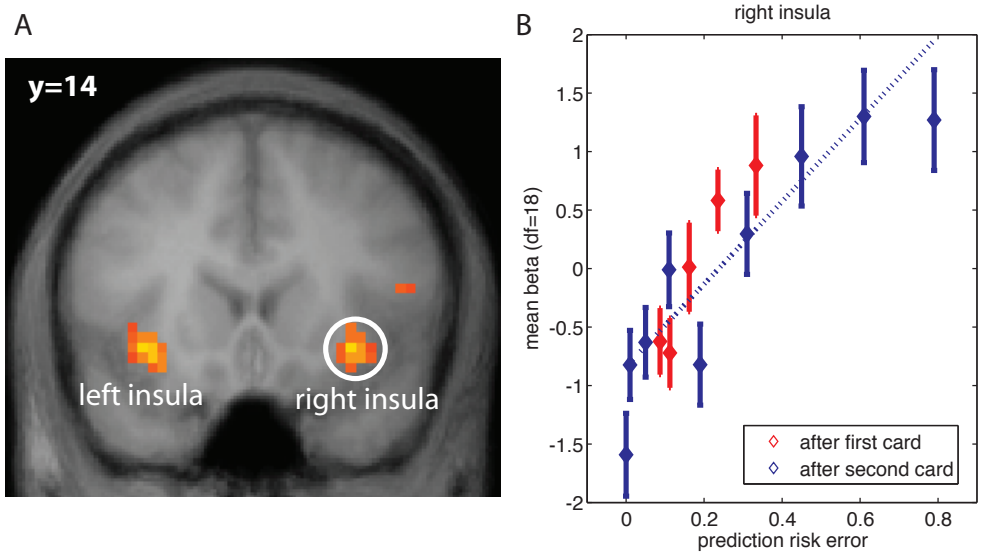


Figure 6.2: A: Activation in bilateral insula correlates positively with a prediction risk error as of display of both the first and second card (random effects, $df = 18$, $p < 0.0005$) B: Activation levels in right insula show a significant linear relationship with the level of prediction risk error at the time of the first card (blue) as well as the second card (red). Furthermore, the functional relationships are comparable at the first and second card.

Brain region	L/R	mean x	mean y	mean z	cluster size	max stat
Anterior insula	R	32	15	−3.3	46	7.3
Anterior insula	L	−31	14	−2.4	41	7.32
Posterior insula	R	49	−11	6.1	95	−6.64
Inferior parietal gyrus	R	57	−27	28	6	−5.63
Angular gyrus	R	53	−52	25	14	7.69
Inferior frontal gyrus	R	48	17	16	12	5.72
Superior temporal gyrus	L	−38	−27	21	58	−7.56
Angular gyrus	L	−43	−60	34	24	7.25
Superior temporal gyrus	L	−56	−11	4.6	24	−6.52

Table 6.1: Activation for early-onset prediction risk error after display of cards 1 and 2.

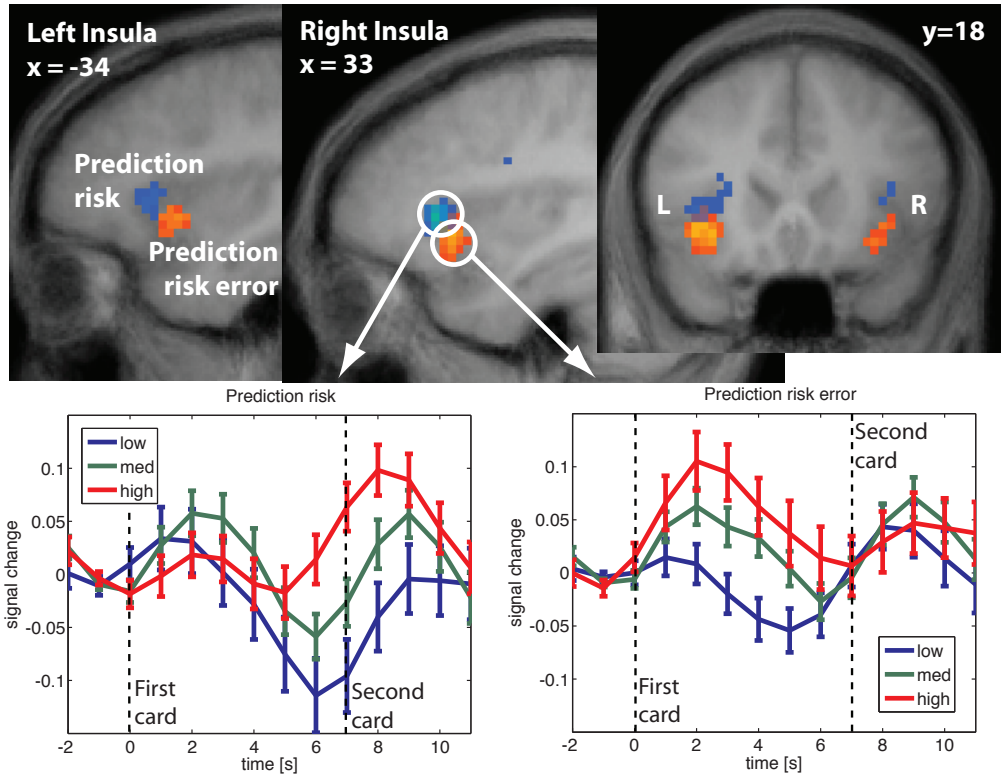


Figure 6.3: A: Activation in bilateral insula correlates with both prediction risk (blue) and prediction risk error (red). Risk is reflected in an area slightly more superior and anterior than prediction risk error. Note, that both the red and blue clusters reflect positive correlations (random effects, $df = 18$, $p < 0.0005$). Different colors were chosen for better visualization. B: Adjusted time courses in right insula at the first card show that different levels of prediction risk errors upon display of the first card are best dissociated immediately after the first card, whereas different levels of prediction risk (anticipated risk of predictions of outcome as of the second card) are best dissociated right before the second card is shown.

Brain regions	R/L	mean x	mean y	mean z	cluster size	max stat
Anterior insula	R	33	21	8	25	−4.94
Subcortical structures*		1	−10	−1.8	258	−8.16
Pulvinar	L	−10	−26	15	4	−4.85
Pulvinar	L	−14	−30	6.4	5	−4.7
Anterior insula	L	−31	22	7.7	82	−7.54
Precentral gyrus	L	−38	−8.8	32	10	−5.25
Precentral gyrus	L	−40	−12	42	6	−4.44

Table 6.2: Activation for late-onset prediction risk during anticipation period between display of cards 1 and 2. *Structures include bilateral ventral striatum, substantia nigra, and thalamic nuclei. This result is reported in chapter 5.

Grinband et al. 2006; Hsu et al. 2005; Critchley et al. 2001; Paulus et al. 2003; Elliott et al. 2000; Ernst et al. 2002), the neural response to prediction risk errors has not been reported, nor has its occurrence been differentiated from that of the prediction risk signals.

The distinct time courses of these signals support the hypothesis that they also play a distinct role in risk processing. The prediction risk error may mediate learning and therefore be important for quickly adapting to rapidly changing uncertain environments. Physiologically, this is supported by the fact that the response to prediction risk errors emerges immediately after the risk cue. In contrast, prediction risk may act as an anticipatory signal before risk is realized. This is supported by the fact that the response to prediction risk is delayed after the risk cue.

There is an intriguing parallel between the activations in insula for prediction risk error and those in subcortical dopaminoceptive structures such as ventral striatum for reward prediction error. Both (error) activations have an early onset. In addition, prior studies report neither early-onset activation in ventral striatum that correlated with risk (Preuschoff et al. 2006) nor early-onset activation in insula that correlated with reward expectation (see tables 6.1 and 6.2), and our data confirm this. The pairing suggests that insula monitors risk and subcortical dopaminoceptive structures monitor expected reward. One could conjecture that insula activation correlates with behavior that primarily involves risk, while activations in subcortical dopaminoceptive structures correlate with behavior modulated mainly by expected reward. A recent study confirms this (Kuhnen and Knutson 2005): insula activation predicts risk avoidance, while activation in nucleus accumbens predicts reward-seeking behavior.

Previous studies of insula activations in the context of uncertainty only reported signals that increased in the level of uncertainty. These studies do not report a prediction risk error, not only because it was not hypothesized, but also because of its peculiar timing. To determine whether our activations were consistent with earlier accounts of uncertainty-related activations in insula, we needed to replace our model with one similar to previous studies. Such models only include prediction risk and no prediction risk error, and they pay little attention to the exact onset of the activations. Insula activations in response to risk became far less significant ($p < 0.05$, uncorrected), whereas prediction risk errors could of course no longer be detected. However, the detectable correlation with risk was positive and therefore consistent with accounts in the literature.

This demonstrates that the standard model (block design) for prediction risk (uncertainty) masked the signal for prediction risk error. We therefore suggest that block designs should be used carefully when modeling complex, intertwined processes. A block design may make it impossible to pick up subtle differentiations of brain signals across sub-regions and over time.

The late-onset activations in insula that correlate with prediction risk exhibit an intriguing pattern. In particular, we discovered that the timing of onset and peak of activations in insula is strikingly similar to those of risk-related activation in ventral striatum (see figure 5.6). Since the signal in ventral striatum is understood to have a dopaminergic origin, and since dopaminergic neurons project to insula as well, we conjecture that the signal in insula may also have a dopaminergic origin. Future research should establish the validity of our conjecture.

The theory of finance dissociates estimation of reward and estimation of prediction risk (Engle 2002, 1982). Our results indicate that the brain implements the same dissociation. Among others, subcortical dopaminoceptive structures encode quantities needed for estimation of expected reward, namely, reward expectation errors. In parallel, insula encodes quantities needed for estimation of prediction risk namely, prediction risk errors.

While forms of reward prediction learning have been found across a diversity of species, including bees (Real 1991) and nonhuman primates (Schultz 2004), the extent of prediction risk learning across species is not known, in part because it has not been explicitly considered. One possibility is that prediction risk estimation may be a relatively late adaptation, perhaps emerging during great ape evolution, which was characterized by/ during heightened habitat instability (Potts 1996). Ripe-fruit frugivory under these conditions would have involved complex spatial and temporal prediction problems, resource variability on both seasonal and longer timescales, and other risk prediction problems that may have contributed to the expansion of brain enlargement and enhanced cognition. By emphasizing the adaptive importance of prediction risk estimation, we suggest that an important future research question will be to assess the extent of its use across species.

In summary, we hypothesized and showed that insula activations in the context of a monetary gamble reflected both prediction risk errors and prediction risk. These are crucial inputs for assessment of risk in a rapidly changing, uncertain world. Most importantly, our results suggest that the earlier understanding that insula is involved in uncertainty-related

phenomena such as complexity, ambiguity and risk needs to be expanded to allow for the possibility that insula encodes prediction risk errors. The suggested expansion is rendered all the more significant by the timing of the prediction risk error signals, which indicates that they may play a primary role in rapid updating. It appears that our understanding of reward anticipation in the dopaminergic system developed in an analogous way: it likewise started with the idea of encoding levels (of rewards), but it later needed to accommodate prediction errors. Finally, just as the idea of a reward prediction error (Montague et al. 2004) has led to new insights into addiction, mental illnesses, and pathological decision making, the notion of risk prediction errors and possible disruptions in risk prediction learning may also have significant clinical implications.

6.6 Mathematical details of reward prediction error, prediction risk and prediction risk error

Referring to the gamble in our experimental paradigm, define the following: P_1 denotes the expected reward conditional on the number on card 1, P_2 denotes the actual reward, revealed upon display of card 2. Before display of card 1, the task is to predict P_1; after display of card 1 and before display of card 2, the task is to predict P_2.

Let P_0 be the prediction of P_1, i.e., $P_0 = E[P_1]$. The prediction error (as of display of card 1) equals $P_1 - P_0$; the prediction risk is the expected size of this prediction error, namely, $E[(P_1 - P_0)^2]$. The prediction risk error is the actual minus the expected size: $(P_1 - P_0)^2 - E[(P_1 - P_0)^2]$.

Analogously, after display of card 1, P_1 is the prediction of P_2: $P_1 = E[P_2]$. the prediction error at the time card 2 is displayed equals $P_2 - P_1$. The prediction risk is the expected size of this prediction error, $E[(P_2 - P_1)^2]$. The prediction risk error is the actual minus the expected size: $(P_2 - P_1)^2 - E[(P_2 - P_1)^2]$.

This is summarized in Table 6.3.

Term	Definition
Reward prediction	$E[P_2]$
Reward	P_2
Reward prediction error	$P_2 - E[P_2]$
Prediction risk	$E[(P_2 - E[P_2])^2]$
Prediction risk error	$(P_2 - E[P_2]) - E[(P_2 - E[P_2])^2]$

Table 6.3: Formal definitions

Chapter 7

Brain Signals of Risk and Reward Jointly Integrate into Anterior Cingulate Cortex

7.1 Abstract

Recent studies show that the expected (anticipated) reward of a gamble and its risk (reward variance) are encoded in separate areas of the human brain. Effective perception of a gamble supposes that these two parameters be integrated. Here, we used canonical correlation analysis to determine what signal the brain constructs. Joint activation in the ventral striatum, putamen, and insula revealed a signal that is increasing in expected reward and in risk. This metric can be interpreted alternatively as reflecting attention or conflict (between risk and reward). In line with these interpretations, the new predictor significantly activates the anterior cingulate cortex.

7.2 Main text

The foundation of modern decision theory under uncertainty is that the anticipated value of a decision can be represented by a single-dimensional index. In 1738, Bernoulli first conjectured that such a metric reflects the trade-off between the expected (anticipated) value of an outcome and its risk (variance), as a lower expected reward (e.g., putting your money into a savings account) is sometimes preferred over a higher expected reward (e.g., investing money into the stock market) simply because it involves lower risk (Bernoulli 1738). Modern decision theory rests on the axiomatic treatment that von Neumann and Morgenstern (1944)

first gave for this trade-off between expected reward and risk. This resulted in Expected Utility for objective probabilities, a single-dimensional decision metric. Expected Utility was subsequently axiomatized by Savage for subjective probabilities (Friedman and Savage 1952). Alternatives to Expected Utility have since been proposed, such as the value function of prospect theory (Kahneman and Tversky 1979). However, they mirror the Expected Utility approach of a single-dimensional index that is some function of expected reward and risk.

To date, it remains unknown whether, or how, the brain constructs a single-dimensional representation of the anticipated value of a decision outcome. Accumulating evidence suggests that both expected reward and risk are encoded in the human brain. Several studies indicate that activation in the ventral striatum and putamen encodes expected reward (Knutson et al. 2001a; Preuschoff et al. 2006), while activation in the insula reflects risk (Preuschoff et al. 2006; Huettel et al. 2005a). This evidence indicates that these two decision parameters are encoded in different areas of the brain, yet it leaves open the question of how they are subsequently integrated to form a single-dimensional index of the anticipated value of a decision outcome, as modern decision theory predicts.

We hypothesized that the brain combines the striatum's expected reward signal and the insula's risk signal into a single metric that is reflected in activation of higher-level cortical structures. On functional imaging (fMRI) data this hypothesis cannot be tested using the standard linear model (GLM) approach (Anderson 1984), as the GLM detects the neural responses to a set of predictors, such as expected reward and risk, by minimizing the error between the variables and activation in a *single* location in the brain. We therefore developed a multivariate approach based on canonical correlation analysis (CCA) (Hotelling 1936), which extends the GLM approach for a set of *multiple* locations in the brain. That is, CCA explains the joint activation across multiple locations in terms of a set of predictors.

We recorded brain activity during a simple gambling task (figure 5.2). The task was chosen as it elicits strong responses to expected reward (in ventral striatum and putamen) and risk (in insula) (Preuschoff et al. 2006). In this task, nineteen subjects placed a bet on whether the first or second of two cards drawn from a deck would be larger. The two cards were drawn without replacement from a randomly shuffled deck of ten cards number 1 to 10. We were interested in the brain activity after the first card is shown but before the second card is revealed. During this phase, expected reward and risk vary with the probability of

Weight	p-value	Parameter
32.63	$< 10^{-7}$	Expected Reward
63.44	$< 10^{-5}$	Reward Variance (Risk)

Weight	p-value	ROI	Talairach (x, y, z)
0.25	0.0017	Putamen	(−22, −8, 8)
0.44	$< 10^{-7}$	Ventral striatum	(−12, 5, −3)
−0.22	$< 10^{-4}$	Insula	(−31, 21, 9)

Table 7.1: CCA model of joint activation in ROIs in terms of prediction model with two predictors: expected reward and risk. Weights are fixed across subjects.

winning (figure 5.1) which depends on the first card and on whether the subject bet that the second card was going to be higher.

Using CCA, we determined how joint activity of regions that encode expected reward (ventral striatum, putamen) and risk (insula), can be explained in terms of a single metric consisting of the combination of two predictors, expected reward and risk. We used standard tests to gauge the significance of the correlation between the joint brain activations and the predictors ("canonical correlation") as well as of the estimated weights ("canonical loadings") in the predictor model (Anderson 1984).

A group analysis on all subjects yielded a highly significant canonical correlation between the estimated joint brain activation and the estimated predictor model ($p < 10^{-7}$; computed with Bartlett's correction (Johnson and Wichern 2002)). More importantly, the weights on both expected reward and risk were positive and highly significant (table 7.1). As such, the signal reflected in the estimated predictor model is increasing in both expected reward and risk.

To determine if these results hold for all subjects, we repeated the analysis for each subject. We obtained significant canonicial correlations in 16 out of 19 subjects ($p < 0.1$). For 15 out of these 16 subjects, the weights for both expected reward and risk in the estimated predictor model are positive ($p < 0.0003$, table 7.2).

The canonical correlation is best visualized by comparing the estimated predictor model and the joint activation of the ROIs at the peak of the hemodynamic response function, 5 s after the first card is revealed (figure 7.1). for each of the ten probabilities the value predicted by the model falls into the 95% confidence interval of the actual (measured) joint

	Number of Subjects
No significant canonical correlation ($p \leq 0.10$)	3
Significant canonical correlation ($p < 0.10$)	16
Of which positive weights on both expected reward and risk	15
Total number of subjects	19

Table 7.2: Subject-by-subject analysis of the sign of the weights (loadings) of the estimated decision metric.

activity.

We next determined the projection area for the signal generated by the joint activations of ventral striatum, putamen and insula. This signal is increasing in both expected reward and risk, and hence, is high when either expected reward and risk or both are high, which suggests that it may be an attention metric. Perhaps more accurately, it could be interpreted as measuring the amount of conflict there is between two decision variables, expected reward and risk. Therefore, we hypothesized that it would activate neural regions involved in processing attention or conflict.

The interpretation of the estimated predictor model in terms of an attention or conflict metric suggested that we look for a neural response in the anterior cingulate cortex (CCA), a brain region known to be involved in attention redirection, especially in the context of cognitive conflict (Botvinick et al. 2004). In accordance with the previously reported activation patterns, we hypothesized that the ACC would correlate with the absolute change in conflict. We performed a group analysis on the 15 subjects for which the CCA yielded significant results.

A sub-region of ACC showed highly significant activation in response to the signal generated by the estimated predictor model (figure 7.2; $p(uncorrected) < 0.0014$; peak at $(1, 13, 50)$ in Talairach coordinates; see also table 7.3). The location and spatial extent of the activation matched those reported for other conflict situations in the literature (Botvinick et al. 2004). Qualitatively similar results were obtained when the predictor model was estimated with constant weights across all subjects (figure 7.3; table 7.4).

7.3 Discussion

Using CCA, we found strong evidence that the joint activations reflecting expected reward in ventral striatum and putamen and risk in insula reflect a single metric. Strikingly, the

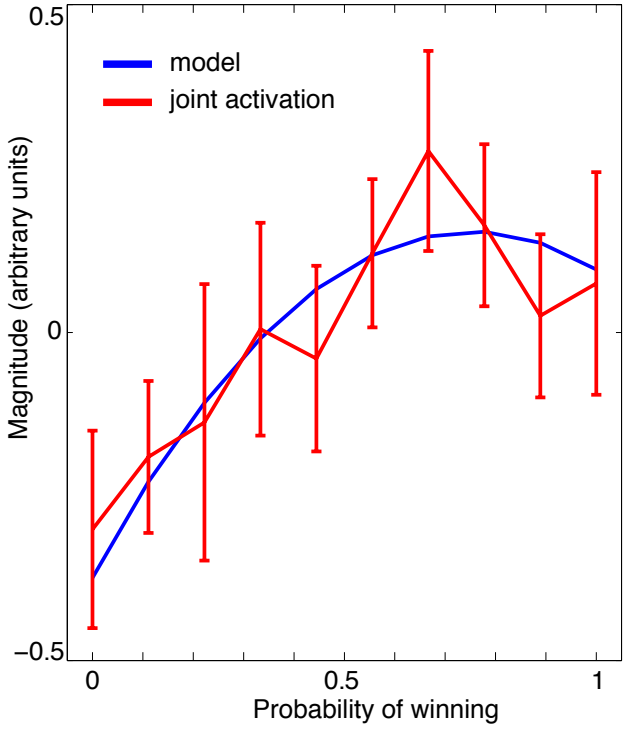


Figure 7.1: Correlation of the activity predicted by the conflict measure U and the combined activity V of ventral striatum, putamen, and insula. Combination of brain activations in three ROIs that is maximally correlated with U is superimposed along with 95% confidence intervals. U and V are measured 5s after the display of the first card, to correct for the hrf delay.

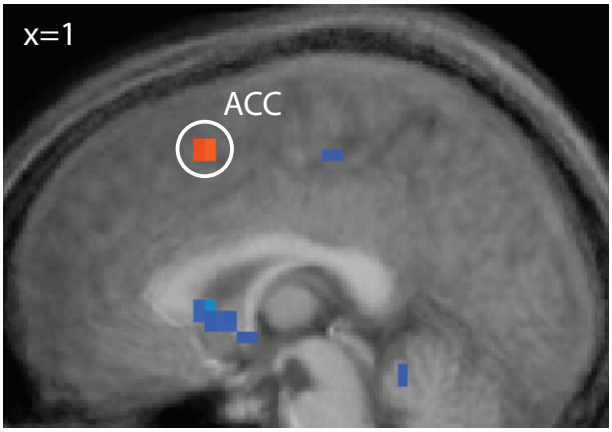


Figure 7.2: ACC responds to changes in the signal predicted from joint brain activations in ventral striatum, putamen, and insula. Sagittal view of activation in response to the size (absolute value) of the change in this signal before and after display of card 1. Map threshold: $p(uncorrected) < 0.0014$.

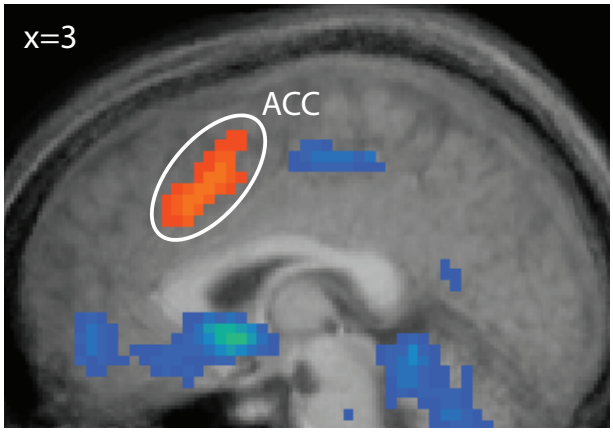


Figure 7.3: ACC responds to changes in the conflict metric estimated form joint brain activations in ventral striatum, putamen, and insula. Weights in conflict metric are forced to be fixed across subjects. Sagittal view of activation in response to the size (absolute value) of the change in conflict before and after display of card 1. Map threshold: $p(FDR) < 0.02$.

Predictor	Talairach (x, y ,z)	t_{14}-stat	cluster size
$\|U - U(0.5)\|$	(1, 13, 50)	4.57	15

Table 7.3: ACC activation in response to change in the estimated conflict metric before ($U(0.5)$) and after (U) display of card 1 (random effects GLM). Coordinates refer to the center of the cluster activated at $P < 0.0014$. The t_{14}-statistic corresponds to the most significant voxel.

Predictor	Talairach (x, y, z)	z-stat	cluster size
$\|U - U(0.5)\|$	(3, 16, 40)	4.644	126

Table 7.4: ACC activation in response to change in the estimated conflict metric before ($U(0.5)$) and after (U) display of card 1 (fixed effects GLM).

metric we recover generates a signal that is increasing in both expected reward and risk, suggesting that it conveys conflict. It is reasonable to posit that humans are risk averse, and hence, if this metric had been reflecting expected utility, while increasing in expected reward, it would have been decreasing in risk (Levy and Markowitz 1979). Supporting our interpretation of the joint activation, the signal correlated significantly with activation in a sub-region of ACC, which is known to be engaged in monitoring conflict. As such, ACC appears to play a crucial role in communicating how much conflict there is between expected reward and risk, the two main parameters that characterize risky gambles.

Our results significantly extend the role of ACC in monitoring of conflict, from the Stroop and related tasks (MacDonald et al. 2000; Pardo et al. 1990) to valuation of probabilistic rewards with varying degrees of expected reward and risk, consistent with recent evidence that ACC is crucial in valuing options based on multiple decision criteria, including risk and expected payoff (Kennerley et al. 2006). Our extension is significant in that we manipulated conflict cardinally, unlike in previous studies, where the level of conflict is only ordinal (generally binomial, and at most multinomial).

Our findings suggest that the metric we observe in ACC may be interpreted as a conflict metric in accordance with the theory of games of conflict (Esteban and Debraj 1999). In this interpretation, sub-cortical regions such as ventral striatum and putamen play a game of conflict with the cortical insula region. One can suppose that the sub-cortical regions have preferences that are increasing in expected reward, while the cortical region has preferences that are increasing in risk. Each region lobbies a third region, the ACC, and one could measure the intensity of the lobbying in terms of the signal strength reflected in the activations of the respective regions. The theory of games of conflict demonstrates that an effective metric of the total amount of conflict is increasing in the contributions of the parties, i.e., in the lobbying efforts. In our context, a metric of conflict between the regions of interest should therefore be increasing in the expected reward and the risk signals that they generate. The predictor model we estimate from the joint activations does precisely that, leading to our interpretation of it as a conflict metric.

That the human brain may engage in monitoring conflict as in games of conflict leads us to a new conjecture that the human brain may solve problems of decision-making under uncertainty through conflict resolution. This contrasts with a view that the human brain would simply aggregate expected reward and risk to construct an expected utility index with

which to facilitate choice and/or determine actions. However, it is important to note that our results do not exclude the possibility that expected utility may be encoded elsewhere in the brain, or at different times. Still, recent lesion studies of ACC in non-human primates (Kennerley et al. 2006) suggests that the ACC signal may be necessary at least in some contexts for adaptive decision-making under uncertainty, and it remains an open issue whether and under what conditions the brain generates additional decision metrics.

As it is increasing in both expected reward and risk, the response in ACC may also be related to attentional processing. The higher expected reward, risk, or both, the more attention should be spent to ensure receipt of the reward (higher expected reward), to avoid risk if the opportunity is provided (higher risk), or to evaluate the trade-off between risk and expected reward (when both are high).

Behind our ability to extract a single metric in expected reward and risk from simultaneous activation in several brain regions lies a novel organizational principle regarding signal processing in the human brain. Indeed, CCA exploits correlations in baseline activations. Such correlations are known to exist between single neurons from different regions with common projection targets (Salinas and Sejnowski 2001); they have also been detected between multiple brain regions using fMRI (Zeki et al. 2003). Aside from an association with attention (Salinas and Sejnowski 2001), the goals of these correlations are unknown. Our conjecture is that they improve re-combination of signals from different brain regions.

Expected reward and risk are two features of random outcomes that need to be jointly considered in order to effectively evaluate gambles. They are separately encoded in the brain, raising the question as to what they are re-integrated into, if at all. Our study uncovered a conflict signal in the joint activations of brain regions that encode expected reward and risk. This signal provides a quantitative measure of conflict, and hence, our finding that it correlates with activation of anterior cingulate cortex significantly expands our understanding of this structure, from ordinal assessment of change in conflict, to cardinal monitoring. Since most humans are risk averse, one could also be surprised that the signal reflected in joint activations of regions that encode expected reward and risk is not an expected utility index. While increasing in expected reward, expected utility is decreasing in risk.

Part II Summary

Following a reward predicting cue, initial activation in ventral striatum and other subcortical dopaminoceptive structures varies with expected reward, whereas subsequent activation in ventral striatum varies with risk. Responses to expected reward are linear in probability; responses to risk are quadratic in probability implying that risk or uncertainty is encoded as variance. The responses arise in the absence of learning, motivation, or salience confounds.

In addition, insula activations reflects both prediction risk errors and prediction risk. These are crucial inputs for assessment of risk in a rapidly changing, uncertain world. This suggests that our understanding of insula needs to be expanded to allow for the possibility that insula encodes prediction risk errors in addition to the previously reported uncertainty-related phenomena such as complexity, ambiguity and risk.

Furthermore, the brain seems to combine the joint activations that reflect expected reward in ventral striatum and putamen and risk in insula into a single metric. This metric is increasing in both expected reward and risk and correlates significantly with activation in a sub-region of ACC, which is known to be engaged in monitoring conflict.

The final part of this thesis will discuss the implications of this.

Part III

Discussion and Future Directions

Chapter 8

Discussion

In the context of learning, in particular learning about future rewards using reward prediction errors, two results presented in this thesis are particularly interesting: (i) reward prediction errors are generated in the absence of learning, (ii) signals of prediction risk and prediction risk errors are generated. The first poses a problem to standard Temporal Difference (TD) models. The second provides a solution.

8.1 Temporal difference learning and prediction risk

TD models, a form of reinforcement learning (RL) models, have successfully explained reward-related neural activity of the dopaminergic system and thus contributed to understanding the neural basis of learning to anticipate uncertain rewards. They provide evidence for the crucial role of the dopaminergic system in such learning and the nature of the underlying learning algorithms. TD models have accounted for a wide range of behavioral and neural phenomena such as foraging behavior in bees (Montague et al. 1995) and dopaminergic response patterns in nonhuman primates (Montague et al. 1996).

The TD model is an extension of the Rescorla-Wagner (RW) rule which tries to predict future rewards by continuously comparing incoming rewards r_t with predictions x_t. The resulting error $\delta_t = r_t - x_t$ is used to update predictions about future rewards $x_{t+1} = x_t + \lambda\delta_t$. The learning rate λ is usually kept constant. However, if the learning rate is fixed, learning should occur whenever a reward prediction error occurs. Consider the following situation:

A fair coin is tossed for a \$1 win or loss. The prediction x_t of the reward p_t on trial t will be \$0 (as this is a random gamble). If the player wins on the next coin toss ($r_t = 1$), then according to TD learning and with a constant learning rate $\lambda > 0$, the player's prediction

x_{t+1} for the next coin toss will be larger than x_t, i.e., $x_{t+1} > 0$. However, most players will not (and should not) update their future predictions, i.e., $x_{t+1} = x_t$. To arrive at $x_{t+1} = x_t$ using the RW rule the learning rate λ should be zero.

The key here is that this is a situation of (known) nonzero risk. Therefore, the prediction risk v_t is non-zero, i.e., prediction errors are expected and updates of future estimates should not be too sensitive to outcomes. In other words, while the prediction x_t is wrong and therefore a prediction error $\delta_t > 0$ is generated, the error itself is predicted. For the TD model to arrive at this solution the learning rate λ should be a function of prediction risk such that large prediction risk decreases the learning rate and small prediction risk increases the learning rate. In addition, note that the best prediction the player could have used, i.e., $x_t = 0\ \forall t$, is *not* correlated with past prediction errors. The covariance between prediction and past prediction errors is zero. According to least squares learning theory, the learning constant λ should be the coefficient in a projection of predictions on past prediction errors. The projection coefficient is defined to be the ratio of a covariance and a variance. In our setting,

$$\lambda_t = \frac{cov(x_{t+1}, e_t)}{v_t(e_t)},$$

where the prediction risk is given by the variance of the prediction error, $v_t = \sqrt{var(e_t, t)}$; v_t cannot be constant over time. During learning, v_t is expected to be larger for early trials than late trials.

Least squares learning can be defended against more powerful learning rules such as Bayes' law because it is agnostic about the model that generates the stimuli (on which predictions are to be based) and the rewards. In other words, it is *model free*.

This formulation suggests that a system that implements simple RL, needs to engage in three tasks:

(i) Track prediction errors.

(ii) Track prediction risk in order to scale prediction errors appropriately.

(iii) Evaluate the covariance between outcomes and scaled prediction errors, i.e., track prediction risk errors.

Evidence has converged over the last ten years that activation of the dopaminergic system in the primate brain reflects prediction errors. In this thesis, evidence is provided

that the brain is also involved in the other two tasks. The prediction risk signal seen in insula is suitable to modulate the learning rate in the manner described above. In addition, insula keeps track of prediction risk errors, such that prediction risk can be evaluated and updated itself.

Is there any other evidence to substantiate the claim that the brain encodes prediction risk and prediction risk errors? And where would such signals originate? This issue has not been addressed directly in the neuroscience literature. On the one side, there is indirect evidence that the signals related to reward prediction and prediction risk both originate in the dopaminergic system. On the other side, there are computational approaches toward modeling uncertainty that suggest the neuromodulators norepinephrine and acetylcholine for mediating different forms of uncertainty.

8.2 The dopaminergic system as a mediator for prediction risk and prediction risk errors

Figure 8.1 reproduces a result from the dopaminergic midbrain neurons of the nonhuman primate brain (Fiorillo et al. 2003). Dopamine neurons in the vental tegmental area of the nonhuman primate brain display a gradual increase in their firing rate in the anticipation period between cue presentation and outcome revelation (reward/no reward). The slope of the increase increases with prediction risk. This effect is referred to as "ramping." In the experiment, prediction risk (when measured as reward variance) is maximal when the reward probability equals 0.5; it is minimal for probabilities equal to zero or one. Correspondingly, ramping increases in probability for probabilities up to 0.5, and decreases for higher probabilities. The researchers interpreted the effect as related to uncertainty, although they were not able to establish the exact measure of uncertainty. The possibility of prediction risk and prediction errors was not explored.

Controversy exists about the origin of this correlation between prediction risk and ramping in dopaminergic neurons or delayed activation in dopaminoceptive areas. Such correlation could spuriously emerge as a result of the averaging of activation across trials on which figures 8.1 and 8.2 are based (Niv et al. 2005). The spurious correlation emerges even in a standard TD learning model, with a constant learning rate, and hence, no account for prediction risk. Still, this explanation relies on specific aspects of the TD learning model

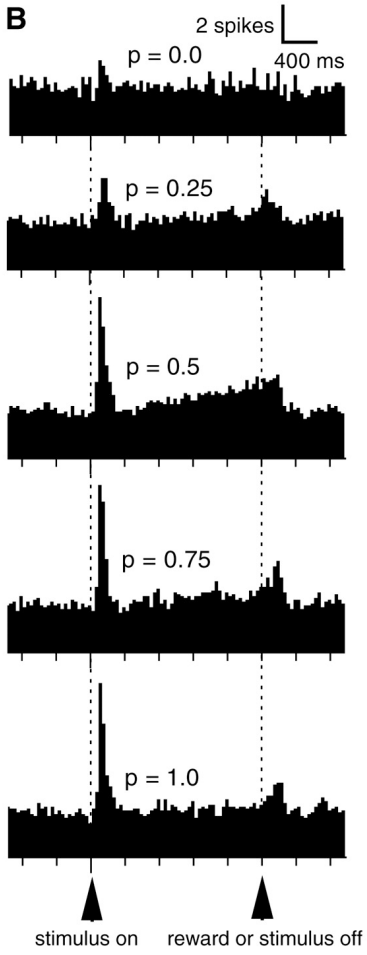


Figure 8.1: Relationship between firing of dopamine neurons in Ventral Tegmental Area and probability of reward. Firing increases gradually ("ramping") in the period of anticipation of reward; the increase is more pronounced the higher prediction risk (measured as reward variance) is. Prediction risk is highest for reward probability p equal to 0.5, and lowest for $p = 0, 1$ (Fiorillo et al. 2003).

(back-propagation of prediction errors in the anticipation period) for which there exists little physiological support. In addition, ramping is observed in single trials for single neurons as well (Fiorillo et al. 2005).

The result of a late emerging (prediction) risk signal in both ventral striatum and insula, together with the prediction risk error signal may help to settle the issue. Prediction risk signals alone are prone to an argument similar to the one above, the prediction risk error signal is not. Assume for a moment that only a prediction risk or risk signal had emerged in the brain. If the ramping effect was a spurious one, it may still show up in the BOLD response as such does not prove the functional significance of the signal. The fact, that a similar signal is seen in ventral striatum and insula, both projection areas of the dopaminergic system, suggests that the signal (or the spurious effect) is relayed to these regions. Still this does not resolve the issue. If however, a signal can be established that is derived from the signal seen in ventral tegmental area, this would provide strong evidence against the spurious interpretation. To achieve this, two things are needed: (i) a solid mathematical measure of the effect, (ii) another measure derived from this measure. (Prediction) risk, as presented in chapters 5 and 6 address (i) by showing that risk is encoded as variance. The prediction risk error is a linear function of this variance and therefore addresses (ii). The prediction risk error is directly related to (and relies on) a neural representation of prediction risk providing strong support for the interpretation of the effect in figure 8.1 as an uncertainty-related signal.

More evidence for prediction risk encoding in the dopaminergic system comes from encoding of prediction errors at reward delivery (figure 8.2). Upon reward delivery, the firing rate does not change with the size of the prediction error, as when the prediction error is normalized with the prediction risk. The firing rate changes only as a function of the scaled prediction error, i.e., as a function of $\frac{e_t}{v_t}$. Unlike neuronal firing that reflects reward anticipation at stimulus onset, firing that correlates with prediction errors at reward delivery appears to be scaled with prediction risk (Tobler et al. 2005).

Recent analysis in Wolfram Schultz laboratory of imaging data of ventral striatum confirm this finding for the human brain (unpublished data). As such, there is evidence of encoding prediction risk in the dopamine system of the primate brain both during the anticipation period and at reward delivery. That an equivalent signal is seen in the ventral striatum, a primary projection target of the dopaminergic midbrain neurons, supports this

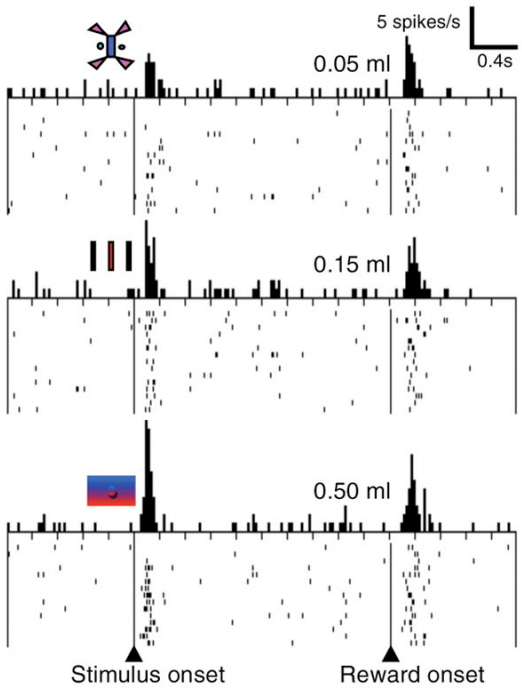


Figure 8.2: Firing of dopamine neurons in ventral tegmental area of nonhuman primate brain, as a function of the size of the prediction risk: low (top), medium (middle), high (bottom). Upon reward delivery, the prediction error encoded in neuronal firing is scaled so that it is independent of prediction risk size (Tobler et al. 2005).

hypothesis.

8.3 Norepinephrine, acetylcholine and the idea of expected and unexpected uncertainty

The results and conclusions presented earlier nicely complement computational work done by Peter Dayan and Angela Yu on the subject of uncertainty. They originally suggested that acetylcholine helps to balance top-down and bottom-up processing in perceptual inference by signaling the uncertainty associated with top-down information. High uncertainty signals that predictions are unreliable and more attention should be paid to bottom-up information (Yu and Dayan 2002).

They later recognized that at least two forms of uncertainty need to be dissociated, expected and unexpected uncertainty. Expected uncertainty arises when the occurence of stimuli or events is known to be unreliable. Unexpected uncertainty however reports on "strong violation[s] of top-down predictions that are expected to be correct." (Yu and Dayan 2005). In the updated model, expected uncertainty is signaled by acetylcholine whereas unexpected uncertainty is mediated by norepinephrine which previously had been implied to mediate attention. Yu and Dayan make a strong case for this distinction and encoding by reviewing a large amount of both modeling and experimental data.

The two forms of uncertainty are closely related to prediction risk and prediction risk errors. Expected uncertainty is equivalent to prediction risk; both signal the (estimated) correlation between predictions and outcomes. Unexpected uncertainty can be viewed as the sum of prediction risk and prediction risk errors. More intuitively, prediction risk errors are the difference between unexpected and expected uncertainty.

Dayan and Yu suggest that acetylcholine and norepinephrine are uncertainty signals encoding these two types of uncertainty which might provide new insights into exploration vs. exploitation mechanisms.

8.4 Risk vs. ambiguity

Another approach toward different forms of uncertainty is once again provided by economics, which makes the distinction between risk and ambiguity (which could also be

termed “known” and “unknown” uncertainty). While this definition is different from Yu and Dayan’s expected and unexpected uncertainty, it clearly addresses the same problem. Sometimes we are aware of the risk and uncertainty associated with our environment and sometimes we are caught by surprise. To point out future directions and approaches toward neural representations of uncertainty we present another study in appendix A that explores decision-making behavior under ambiguity. While at this stage, the study provides mostly pilot data it is included nonetheless because it nicely demonstrates a behavioral approach to modeling choices under ambiguity that can be expanded to functional imaging as well.

8.5 Summary

Compared to our understanding of processing expected rewards in the brain, our understanding of processing uncertainty is in its infancy. The ideas presented above are just some of many possible approaches to dealing with uncertainty. As this is the first account of prediction risk and prediction risk errors in the brain, a lot of research lies ahead, much of which is likely to mirror the path that reward prediction learning has taken, ranging from computational modeling to neuroanatomy, to new learning paradigms that can track learning as a function of prediction risk errors.

Appendix

Appendix A

Decision Making under Ambiguity

Ecologic and economic theories of decision making under uncertainty emphasize the importance of correctly evaluating expected reward and risk. Within neuroscience, the idea obtains strong support from the fact that the human brain reflects both mathematical expectation of reward (mean) and risk (variance) (part II) in situations under pure risk. Pure risk is often dissociated from "true" uncertainty or ambiguity in which probabilities of outcomes are unknown. As such, the outcome of a fair coin flip is risky but not uncertain (ambiguous). Weather reports however are ambiguous. And so are most natural stochastic environments.

A.1 Introduction

In situations that involve neither risk nor ambiguity, organisms prefer higher expected rewards over lower expected rewards. In situations that involve risk but no ambiguity, most people prefer low risk options over high risk options if the options have the same expected reward. Preferences across individuals become less homogenous when both expected reward and risk are varied, i.e., when a trade-off between the two is introduced. Behaviors range from high risk aversion via risk neutrality to risk seeking. Most people are risk averse, though to different degrees. Analogously, in situations of ambiguity most people prefer low ambiguity options over high ambiguity options. In other words, just as increasing risk lowers the overall value of an option so does increasing ambiguity. Again, most people are ambiguity averse (to different degrees) rather than ambiguity neutral or seeking. Attitudes toward both risk and ambiguity have been manipulated by changing the organisms current state, such as the energy budget (Caraco 1981) or financial budget (Rode et al. 1999).

Behavioral ecology treats most situations of uncertainty as ones of risk by talking about *subjective* (perceived) probabilities rather than the actual *objective* (underlying) probabilities. As such any situation in which objective probabilities are not explicitly stated or known can be represented in terms of expected reward and risk using the organism's *subjective* probabilities. In general, ecology is not interested in how *objective* probabilities are transformed into *subjective* probabilities occurs. As such, it does not explicitly deal with the difference between behavior under ambiguity and behavior under risk.

In economics, the dissociation between risk and ambiguity is highly controversial and theories of how to account for different forms of uncertainty are abound (Kreps 1988). And while ambiguity aversion is well documented (Ellsberg 1961) it is not well studied behaviorally.

The study presented in this chapter proposes an approach to studying both behavior and neural circuits of decision making under ambiguity. In this first stage, the study tries to quantify behavior in addition to determining preferences under ambiguity. More precisely, on a behavioral level, how are ambiguous situations treated differently from risky ones? Can theories of decision making under risk as those presented in earlier chapters (be expanded to) capture decision-making behavior under ambiguity? Are risk and ambiguity treated differently? So far, neither neuroscience nor ecology have explicitly dealt with the distinction between risk and ambiguity. Therefore, the mathematical framework will once again be provided by economics and finance.

While this is only an extended pilot study, the long term goal of this work is to understand the neural circuits involved in evaluating situations of all forms of uncertainty.

A.2 A mathematical approach to choices under ambiguity

Economics accounts for choices under ambiguity by expanding the unidimensional index, expected utility, to capture not only attitudes toward expected reward and risk but ambiguity as well. For this, the ambiguity aversion coefficient $\alpha \in [0, 1]$ describes preference for ambiguity, with $\alpha = 1$ for a highly ambiguity averse person, $\alpha = 0.5$ for an ambiguity neutral person, and $\alpha = 0$ for a highly ambiguity seeking person. In analogy to risk, ambiguity is not explicitly modeled in economics but captured by the utility $u(x_i)$. Again, the expected utility of an option X is determined by the utility u of each possible outcome x_i

weighted by the probability p_i of that outcome:

$$EU(X) = \sum_i p_i u(x_i). \tag{A.1}$$

Under the assumption that $u(x)$ is known, expected utility theory estimates the unknown probabilities p_i using the minimum expected and maximum expected utility of an option. This approach uses a weighted sum of the maximum and minimum expected utility, weighted by the ambiguity aversion coefficient α (α max min-preferences (Ghirardato et al. 2004)).

$$EU(X) = \alpha \min_p E[U(x)] + (1-\alpha) \max_p E[U(x)]. \tag{A.2}$$

The idea of α max min-preferences is frequently used to determine preferences in situations of uncertainty (including both risk and ambiguity). To approach stems from the idea that in any decision situation there is a best possible outcome, $minE[U(x)]$, and a worst possible outcome, $maxE[U(x)]$. The coefficient α reflects to which degree a person believes nature to be malevolent or benevolent. For a highly ambiguity averse person ($\alpha = 1$) for instance, $EU(X)$ is based solely on the minimum available utility, $\min E[U(x)]$. In other words, the person will ignore what *may* be available (reflected by $\max E[U(x)]$).

Finance uses the same approaches but again replaces expected utility by its mean-variance approximation:

$$EU \sim \alpha \min_p (E[R] - \frac{1}{2} var[R]) + (1-\alpha) \max_p (E[R] - \frac{1}{2} var[R]). \tag{A.3}$$

Note that the expected utility approach, equation (A.1) is a sum over all possible outcomes, i.e., EU requires a separate (neural) representation for each probability and each utility. Equation (A.3) takes an estimate of the mean and variance (of all possible options together) rather than computing individual probabilities and utilities. As such, the mean variance approximation is not a function of the number of options as we have shown that there are neural representations of mean and variance. However, this approximation relies on an assumption about the underlying expected utility function (here: logarithmic) which determines the weights of $E[R]$ and $var[R]$.

In this first experimental stage, the questions we would like to answer are: Do people make consistent choices under ambiguity such that future choices can be predicted from

past ones? Can equations (A.2) and (A.3) capture such preferences under ambiguity? How do the two equations compare in predicting choices?

We used a gambling task in which subjects had to choose between two ambiguous gambles. To approach the questions above, the goal was to (i) determine subjects' choices under ambiguity (ii) look for patterns of choice, (iii) look for patterns that deviate from (ambiguity neutral) decision making under risk. If ambiguity is treated the same as risk, (close to) all subjects should turn out to be ambiguity neutral, i.e., $\alpha = 0.5$. If however, α turns out to be different from 0.5 then the question arises how and why ambiguous gambles are evaluated differently.

A.3 Task

On each trial two colors are presented to the player. A ball is then drawn from an urn. If the color of the ball matches one of the two colors presented the subject wins \$1, otherwise the player loses \$1. Subjects cannot choose the colors they bet on. But they can choose the urn. On each trial, two urns are shown. Both urns provide different information about the distribution of balls in it. Figure A.1 shows two such comparisons. Each urn has red, green, and blue balls in it. In the given example, for both urns the exact proportion of red balls is known. Also, a minimum number of blue and green balls is known. However, in each of the two urns there is a subset of balls which can be either green or blue. This is the ambiguous part. Subjects are now told that they will bet on either blue or red and are asked if they would like a ball to be drawn from the first or the second urn. They pick an urn, a ball will be drawn from that urn. If it matches either color they win, if not, they lose. For the experiment, the urns are represented as a circle with three colored segments. The size of a segment corresponds to the number of balls with the corresponding color in the urn. Subjects never see the exact composition of an urn because the circle is partly covered. The size and position of this gray cover changes over time. The exact composition of the urns is determined before the experiment and may change from one trial to another. No hints are given about what is underneath the cover. Outcomes in one trial do not tell anything useful for subsequent trials.

Once an urn is chosen, a ball is drawn from that urn. The probability for the ambiguous colors is determined by picking at random (from a uniform distribution) a position

underneath the covered part of the circle.

The true distribution under the cover is uniform although subjects are neither told, nor are they shown the actual distribution under the cover after each trial. Thus we avoid that subjects learn over time what the underlying distribution is. In this case the gamble would turn from an ambiguous into a risky gamble over time as in (Huettel et al. 2006).

A total of 23 subjects participated in the study. All subjects gave full informed consent to participate in the study. The study was approved by the California Institute of Technology Institutional Review Board. Each subject was given written instructions for the game and completed a brief training session outside the magnet.

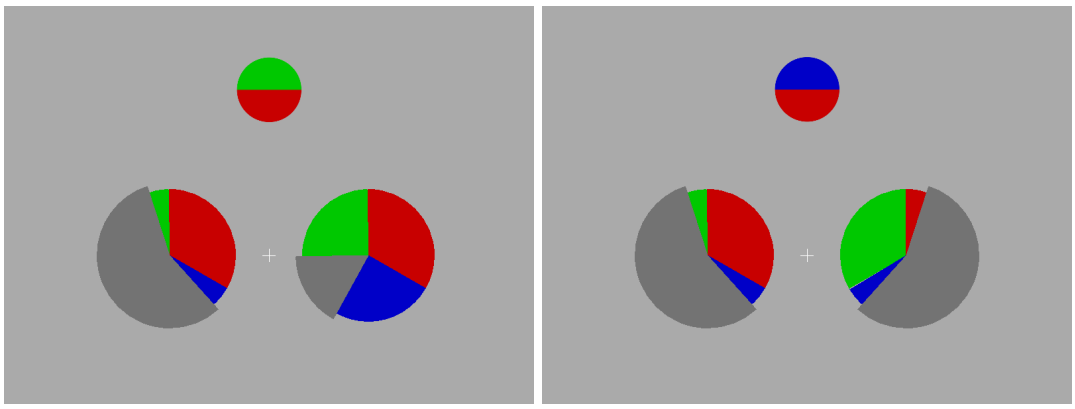


Figure A.1: Decision screen. Left panel - comparing two ambiguous gambles. The two colored ball at the top of the screen indicates that the subject will win if either green or red are drawn from an urn. Two urns are presented. The fraction of the circle covered by each color represents the fraction of balls of that color in the urn. In the left panel, both urns contain exactly one third red balls indicated by one third of the circle being red. In the left urn at least 5% of the balls are green and at least 5% are blue with the remaining balls unknown (either blue or green). In the right urn, at least 25% of the balls are green and 25% are blue. This means that for the left urn the minimum probability of drawing either red or green is given by ($p_{min} = 0.33 + 0.05$), whereas the maximum probability is given as ($p_{max} = 1 - 0.05$). In the right urn, the minimum and maximum probabilities are given by ($p_{min} = 0.33 + 0.25$) and ($p_{max} = 1 - 0.25$), respectively. Both urns are ambiguous since the exact probability of winning is not known for either one. Right panel - comparing one ambiguous and one risky gamble. Both urns are evaluated in the same manner as in the urns in the left panel. Note, that the right urn turns out to be risky rather than ambiguous because the minimum and maximum probabilities are equal and given by ($p_{min} = p_{max} = 1 - 0.33$). As such the subject has to choose between a risky and an ambiguous urn which in this case are matched for both (objective) expected reward and risk.

Trial type	# of trials	Description
A	4	risky
B	4	risky
C	8	ambiguous
D	4	ambiguous
E	4	ambiguous
F	8	ambiguous
G	8	ambiguous
H	4	ambiguous
I	4	ambiguous
J	8	ambiguous
K	8	ambiguous

Table A.1: Classification of trials. The number of trials in each group is needed to properly balance the design when building constrast.

A.4 Trial types

Two groups of trials were used. In the first group subjects were presented with two ambiguous choices. In the second group subjects had to choose between a risky and an ambiguous gamble. In each session, there are 56 trials in the first group, and 8 in the second, summing to 64 trials per session. The trials in the second group are chosen as a control to look at choices involving pure risk. The other 56 trials are chosen to best determine the ambiguity type of the subjects.

The 64 trials can be grouped into 11 types, which are then balanced for colors and laterality. Table A.1 lists the 11 types and the number of trials in each category.

A.5 Ambiguity types

We first predicted the choices for each of the 64 gambles as a function of α using both expected utility (equation (A.2)) and the mean variance approximation (equation (A.3)). That is, different ambiguity types are predicted to make different choices.

For expected utility, equation (A.1) yields

$$EU = \tilde{p}u(win) + (1-\tilde{p})u(loss) \quad (A.4)$$

$$= \tilde{p}. \quad (A.5)$$

The utilities u for the two different outcomes (win or loss) are set arbitrarily to $u(win) = 1$ and $u(loss) = 0$. This can be done as the utility function in itself is not specified and is a number up to a linear transformation. To determine the perceived probability $\tilde{p}$, let p_i be the probability of the risky color, p_a be the probability of the ambiguous color that you bet on, and p_c be the probability of the (ambiguous) color that you do not bet on. The reward R equals 1 or -1. With that,

$$\begin{aligned} \tilde{p} &= \alpha \arg\min_p E[U(x)] + (1-\alpha) \arg\max_p E[U(x)] \\ &= \alpha(p_i + p_a) + (1-\alpha)(1-p_c). \end{aligned}$$

Here, $\arg\min_p E[U(x)]$ denotes the value p that yields the minimum expected utility for a given gamble.

For the finance approach, the mean and variance of each gamble are determined by

$$\begin{aligned} E[R] &= < R >= \Sigma_i p_i x_i \\ &= px + (1-p)(-x) \\ &= 2px - x \end{aligned}$$

and

$$\begin{aligned} V[R] &= < (R - \bar{R})^2 >=< R^2 > - < R >^2 \\ &= \Sigma_i p_i x_i^2 - < R >^2 \\ &= px^2 + (1-p)(-x)^2 - (2px - x)^2 \\ &= px^2 + (1-p)x^2 - (4p^2x^2 - 4px^2 + x^2) \\ &= 4px^2 - 4p^2x^2. \end{aligned}$$

With that, the expected utility is approximated to

$$EU \sim \alpha \min_p \left(E[R] - \frac{1}{2}V[R]\right) + (1-\alpha) \max_p \left(E[R] - \frac{1}{2}V[R]\right), \tag{A.6}$$

where α is again the ambiguity aversion coefficient.

A.6 Results

Based on a uniform distribution underneath the cover, the average expected probability of win is 0.67 as there is on average a 2/3 chance of winning on each trial. Subjects won on $67.3 \pm 1.54\%$ of all trials.

For each subject, the ambiguity type was determined based on their choices. Predicted choices for $\alpha = 0$, $\alpha = 0.5$, and $\alpha = 1$ were compared to the actual choices the subject made. For each type the *hitrate*, that is the number of choices that matched the predicted choices, was computed (figures A.2 A.3). The ambiguity type was then determined using a maximum criterion (highest hit rate). (To compute hit rates all trials in which subjects were predicted to be indifferent were omitted.)

Figures A.2 and A.3 show the behavioral data for 23 subjects, each playing 3 sessions of 64 trials. Subjects played the same 64 trials in each session but in different randomized order. All subjects played the same three sessions although in randomized order.

For most subjects hit rates consistently differ from 50% (chance level) indicating that choices were not random and subjects evaluated the different gambles. Two out of 23 subjects show no correlation with any ambiguity type. This indicates a strategy unrelated to the information provided and *not* ambiguity neutrality. Ambiguity neutrality would be reflected in a large hit rate for $\alpha = 0.5$.

Especially for the mean variance approximation but also for expected utility most subjects can be classified as either highly ($\alpha = 1$) or moderately ($\alpha = 0.75$) ambiguity averse. Choices are *not* best predicted by ambiguity neutrality ($\alpha = 0$) indicating that the evaluation of the gambles is not simply based on objective measures of probability and risk. Nonetheless, the evaluation of the ambiguous gambles seems to be systematic as subjects' choices are consistent across session.

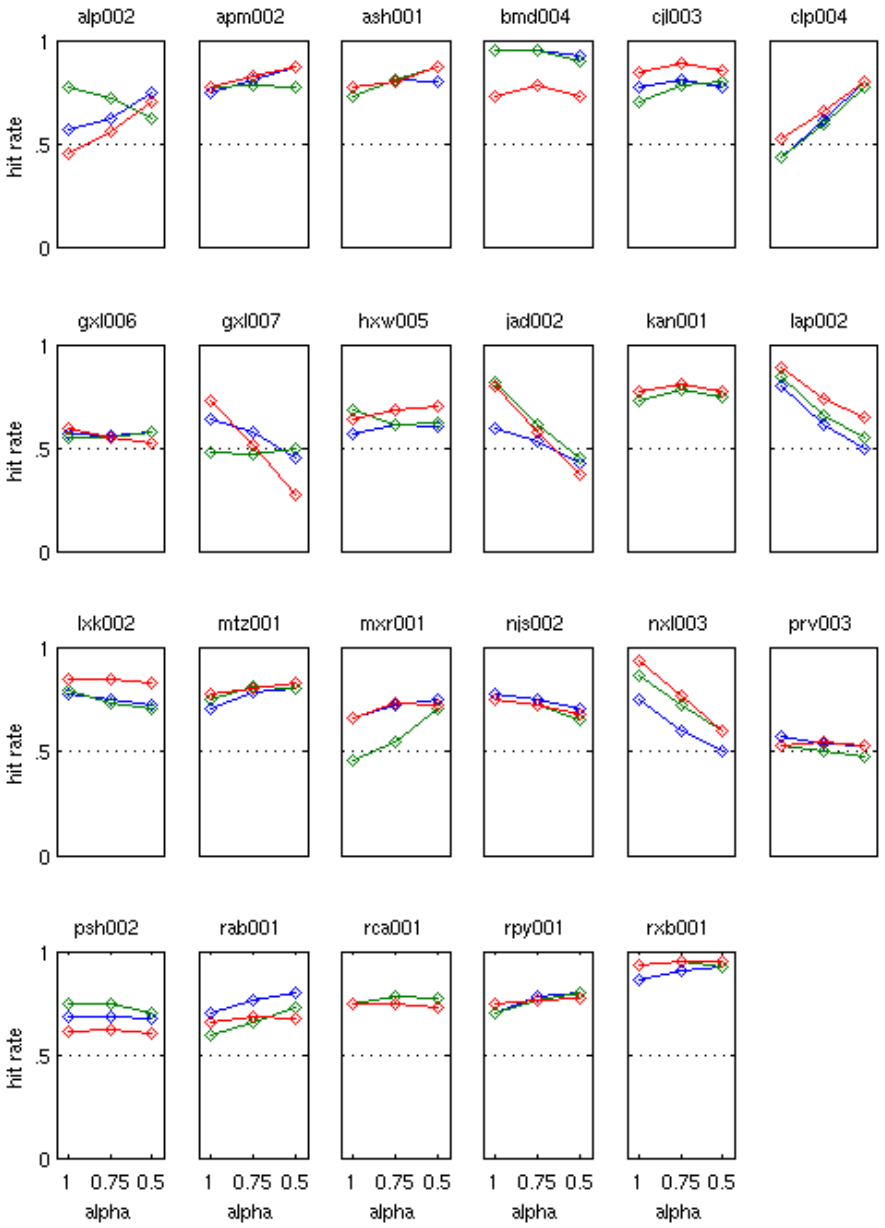


Figure A.2: Ambiguity types based on expected utility (equation (A.2))

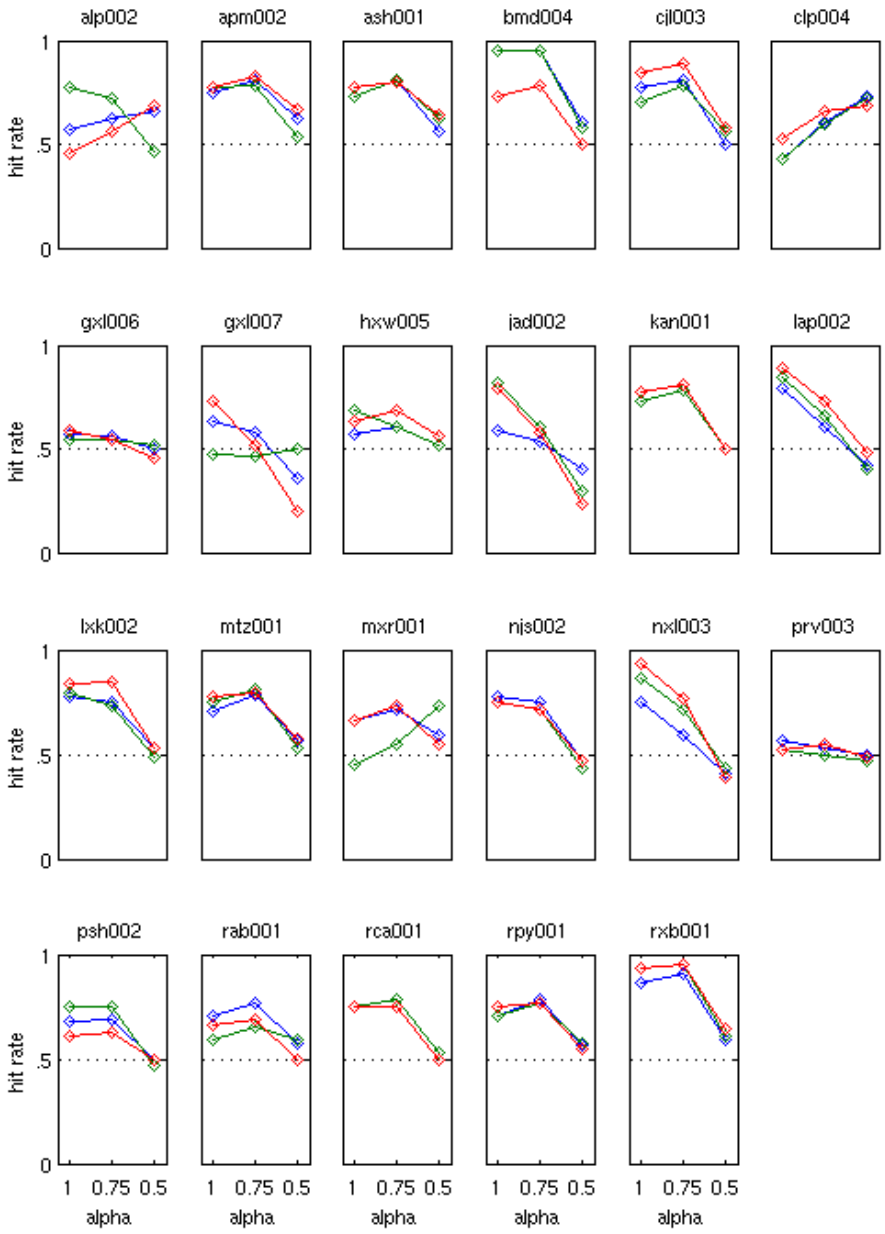


Figure A.3: Ambiguity types based on the mean variance approach to expected utility (equation (A.3)).

A.7 Discussion

While this study is still in its early stages it nonetheless makes a number of interesting points. In decision making under ambiguity subjects systematically evaluate ambiguous gambles to arrive at consistent choices. As such, there seems to be a mechanism underlying choice behavior to systematically evaluate ambiguous situations. The question arises if ambiguity should be treated separately from risk as is often suggested in economics (Knight 1921) or if ambiguity and risk are two facets of a more general form of uncertainty. Both, the expected utility approach as well as the mean variance approximation were able to capture preferences under ambiguity. And both approaches can easily be reduced to the pure risk case.

It can be argued that situations of pure risk are only generated in laboratory experiments and certain types of gambles. As such it is doubtful that the brain has evolved to separately deal with situations of risk and ambiguity. Instead, ambiguity - although within economics it was "discovered" later - may be the only one of these two forms of uncertainty which is encountered in nature. In this view, risk becomes a special case of ambiguity rather than ambiguity being an extension of risk.

From a modeling point of view, ambiguous situations are more complex than situations of pure risk since more parameters have to be determined. Ambiguity aversion as well as estimates of expected reward and risk may all differ across individuals. Determining the correct measures and individual differences poses a challenge to future behavioral as well as neuroscientific studies. In addition, how individuals trade off mean and variances may also vary.

A.8 Future work

The long term goal of this work is to understand the neural circuits involved in evaluating situations of all forms of uncertainty. Previous research has identified structures involved in ambiguity evaluation (Hsu et al. 2005; Huettel et al. 2006). Here, we would like to understand *how* these structures are involved. The hypothesis underlying this work is that in situations of (any form of) uncertainty organisms estimate both means and variances of future stimuli and events. The estimates are then traded off to establish the value of a stimulus, action, or event. This trade-off does not necessarily differ across individuals.

While a phenomenon such as risk aversion or ambiguity aversion may be captured by a risk aversion or ambiguity aversion coefficient α, it may also be reflected in the estimate of risk itself. In other words, a highly risk averse person will grossly overestimate risk rather than giving it a larger weight when combining expected reward and risk.

In the next stage, the estimates of expected reward and risk computed for different ambiguity types may be used to determine ambiguity types on brain activity alone. This would strengthen the idea of a mean variance representation as well as the idea that different types of uncertainty are evaluated in a similar manner.

1

[1]To get a better grasp at the underlying neural representations it may be helpful to change the comparisons between different gambles used in this study to include more risky gambles. As representations of mean and variance are already established under pure risk, such gambles will provide a better baseline for determining neural activity in ambiguous situations. It would also allow to better measure individual attitudes toward risk. However, the overall number of trials is limited in imaging experiments as single trials tend to be longer than in purely behavioral experiments.

It remains to be seen if the two approaches of expected utility and its mean variance approximation can easily be compared. While from figures A.2 and A.3 it may seem as if the expected utility representation has less discriminative power than the its mean-variance approximation this is (at least partially) a result of the experimental setup.

References

A. B. Abel. Asset prices under habit formation and catching up with the Joneses. *American Economic Review*, 80(2):38–42, 1990.

A. K. Anderson, K. Christoff, I. Stappen, D. Panitz, D. G. Ghahremani, G. Glover, J. D. E. Gabrieli, and N. Sobel. Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6(2):196–202, 2003.

T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. Wiley-Interscience, New York, 2nd edition, 1984.

P. Apicella, T. Ljungberg, E. Scarnati, and W. Schultz. Responses to reward in monkey dorsal and ventral striatum. *Experimental Brain Research*, 85(3):491–500, 1991.

A. R. Aron, D. Shohamy, J. Clark, C. Myers, M. A. Gluck, and R. A. Poldrack. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, 92(2):1144–52, 2004.

C. J. Barnard and C. A. J. Brown. Risk-sensitive foraging in common shrews (sorex-araneus l). *Behavioral Ecology and Sociobiology*, 16(2):161–164, 1985.

R. C. Battalio, J. H. Kagel, and D. N. Macdonald. Animals choices over uncertain outcomes - some initial experimental results. *American Economic Review*, 75(4):597–613, 1985.

W. M. Baum. Time allocation in human vigilance. *Journal of the Experimental Analysis of Behavior*, 23(1):45–53, 1975.

H. M. Bayer and P. W. Glimcher. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*, 47(1):129–141, 2005.

A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio. Deciding advantageously before knowing the advantageous strategy. *Science*, 275(5304):1293–1295, 1997.

D. Bernoulli. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, V:175–192, 1738.

G. S. Berns, S. M. McClure, G. Pagnoni, and P. R. Montague. Predictability modulates human brain response to reward. *Journal of Neuroscience*, 21(8):2793–8, 2001.

W. Bialek and F. Rieke. Reliability and information transmission in spiking neurons. *Trends in Neuroscience*, 15(11):428–34, 1992.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

P. Bossaerts and C. Plott. Basic principles of asset pricing theory: Evidence from large-scale experimental financial markets. *Review of Finance*, 8:135–169, 2004.

M. M. Botvinick, J. D. Cohen, and C. S. Carter. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Science*, 8(12):539–46, 2004.

H. C. Breiter, I. Aharon, D. Kahneman, A. Dale, and P. Shizgal. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*, 30 (2):619–39, 2001.

T. Bruguier*, K. Preuschoff*, S. R. Quartz, and P. Bossaerts. Decision making under uncertainty as conflict resolution: Brain signals of risk and reward jointly activate anterior cingulate cortex. *submitted*.

M. J. Burton, E. T. Rolls, and F. Mora. Effects of hunger on responses of neurons in lateral hypothalamus to sight and taste of food. *Experimental Neurology*, 51(3):668–677, 1976.

T. Caraco. Energy budgets, risk and foraging preferences in dark-eyed juncos (junco-hyemalis). *Behavioral Ecology and Sociobiology*, 8(3):213–217, 1981.

T. Caraco. White-crowned sparrows (zonotrichia-leucophrys) - foraging preferences in a risky environment. *Behavioral Ecology and Sociobiology*, 12(1):63–69, 1983.

P. Cavedini, G. Riboldi, R. Keller, A. D'Annucci, and L. Bellodi. Frontal lobe dysfunction in pathological gambling patients. *Biological Psychiatry*, 51(4):334–341, 2002.

L. Clark, S. D. Iversen, and G. M. Goodwin. A neuropsychological investigation of prefrontal cortex involvement in acute mania. *American Journal of Psychiatry*, 158(10):1605–1611, 2001.

H. D. Critchley and E. T. Rolls. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *Journal of Neurophysiology*, 75(4): 1673–1686, 1996.

H. D. Critchley, C. J. Mathias, and R. J. Dolan. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*, 29(2):537–45, 2001.

R. Davidson and J. G. MacKinnon. Several tests for model specification in the presence of alternative hypotheses. *Econometrica*, 49(3):781–93, 1981.

M. C. Davison and I. W. Hunter. Performance on variable-interval schedules arranged singly and concurrently. *Journal of the Experimental Analysis of Behavior*, 25(3):335–345, 1976.

M. R. Delgado, L. E. Nystrom, C. Fissell, D. C. Noll, and J. A. Fiez. Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84(6):3072–7, 2000.

M. C. Dorris and P. W. Glimcher. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*, 44(2):365–378, 2004.

J. C. Dreher, P. Kohn, and K. F. Berman. Neural coding of distinct statistical properties of reward information in humans. *Cerebral Cortex*, 2005.

B. D. Dunn, T. Dalgleish, and A. D. Lawrence. The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews*, 30(2):239–271, 2006.

R. Elliott, K. J. Friston, and R. J. Dolan. Dissociable neural responses in human reward systems. *Journal of Neuroscience*, 20(16):6159–65, 2000.

R. Elliott, J. L. Newman, O. A. Longe, and J. F. Deakin. Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience*, 23(1):303–7, 2003.

D. Ellsberg. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75: 643–669, 1961.

Encyclopaedia Britannica. *"probability and statistics"*. Encyclopaedia Britannica Online, http://search.eb.com/eb/article-248185, 25 Oct., 2006.

R. F. Engle. New frontiers for arch models. *Journal of Applied Econometrics*, 17:425–446, 2002.

R. F. Engle. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50:987–1008, 1982.

M. Ernst, K. Bolla, M. Mouratidis, C. Contoreggi, J. A. Matochik, V. Kurian, J. L. Cadet, A. S. Kimes, and E. D. London. Decision-making in a risk-taking task: A pet study. *Neuropsychopharmacology*, 26(5):682–691, 2002.

M. Ernst, E. E. Nelson, E. B. McClure, C. S. Monk, S. Munson, N. Eshel, E. Zarahn, E. Leibenluft, A. Zametkin, K. Towbin, J. Blair, D. Charney, and D. S. Pine. Choice selection and reward anticipation: An fmri study. *Neuropsychologia*, 42(12):1585–97, 2004.

J. Esteban and R. Debraj. Conflict and distribution. *Journal of Economic Theory*, 87: 379–415, 1999.

C. D. Fiorillo, P. N. Tobler, and W. Schultz. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299(5614):1898–902, 2003.

C. D. Fiorillo, P. N. Tobler, and W. Schultz. Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating td errors. *Behavioral Brain Function*, 1(1):7, 2005.

M. Friedman and L. J. Savage. The expected-utility hypothesis and the measurability of utility. *The Journal of Political Economy*, 60(6):463–474, 1952.

H. Fukui, T. Murai, H. Fukuyama, T. Hayashi, and T. Hanakawa. Functional activity related to risk anticipation during performance of the iowa gambling task. *Neuroimage*, 24(1):253–9, 2005.

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

D. A. Graft, S. E. G. Lea, and T. L. Whitworth. Matching law in and within groups of rats. *Journal of the Experimental Analysis of Behavior*, 27(1):183–194, 1977.

J. Grinband, J. Hirsch, and V. P. Ferrera. A neural representation uncertainty in the of categorization human brain. *Neuron*, 49(5):757–763, 2006.

L. D. Harder and L. A. Real. Why are bumble bees risk averse. *Ecology*, 68(4):1104–1108, 1987.

L. M. Harrison, A. Duggins, and K. J. Friston. Encoding uncertainty in the hippocampus. *Neural Networks*, 19(5):535–546, 2006.

O. K. Hassani, H. C. Cromwell, and W. Schultz. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *Journal of Neurophysiology*, 85(6): 2477–2489, 2001.

R. J. Herrnstein. Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4(3):267–272, 1961.

R. J. Herrnstein. Formal properties of matching law. *Journal of the Experimental Analysis of Behavior*, 21(1):159–164, 1974.

J. R. Hollerman and W. Schultz. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience*, 1(4):304–309, 1998.

C. A. Holt and S. K. Laury. Risk aversion and incentive effects. *American Economic Review*, 92(5):1644–1655, 2002.

H. Hotelling. Relations between two sets of variates. *Biometrika*, 28(3/4):321–377, 1936.

M. Hsu, M. Bhatt, R. Adolphs, D. Tranel, and C. F. Camerer. Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310(5754):1680–1683, 2005.

S. A. Huettel, A. W. Song, and G. McCarthy. Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience*, 25(13):3304–3311, 2005a.

S. A. Huettel, C. Stowe, M. Platt, E. Gordon, and B. Warner. Choices between gambles: Effects of certainty, risk, and ambiguity upon brain systems for decision making and reward evaluation. *Journal of Cognitive Neuroscience*, pages 221–221, 2005b.

S. A. Huettel, C. J. Stowe, E. M. Gordon, B. T. Warner, and M. L. Platt. Neural signatures of economic preferences for risk and ambiguity. *Neuron*, 49(5):765–775, 2006.

T. Ikeda and O. Hikosaka. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*, 39(4):693–700, 2003.

R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, 5th edition, 2002.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

S. W. Kennerley, M. E. Walton, T. E. Behrens, M. J. Buckley, and M. F. Rushworth. Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, 9(7): 940–7, 2006.

F. H. Knight. *Risk, Uncertainty, and Profit*. Houghton & Mifflin, New York, 1921.

B. Knutson, A. Westdorp, E. Kaiser, and D. Hommer. Fmri visualization of brain activity during a monetary incentive delay task. *Neuroimage*, 12(1):20–7, 2000.

B. Knutson, C. M. Adams, G. W. Fong, and D. Hommer. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21(16): RC159, 2001a.

B. Knutson, G. W. Fong, C. M. Adams, J. L. Varner, and D. Hommer. Dissociation of reward anticipation and outcome with event-related fmri. *Neuroreport*, 12(17):3683–7, 2001b.

B. Knutson, G. W. Fong, S. M. Bennett, C. M. Adams, and D. Hommer. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fmri. *Neuroimage*, 18(2):263–72, 2003.

D. M. Kreps. *Notes on the Theory of Choice*. Westview Press, 1988.

C. M. Kuhnen and B. Knutson. The neural basis of financial risk taking. *Neuron*, 47(5): 763–770, 2005.

J. Lauwereyns, K. Watanabe, B. Coe, and O. Hikosaka. A neural correlate of response bias in monkey caudate nucleus. *Nature*, 418(6896):413–417, 2002.

H. Levy and H. M. Markowitz. Approximating expected utility by a function of mean and variance. *American Economic Review*, 69(3):308–317, 1979.

J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economic Statistics*, 47(1):13–37, 1965.

Z. Liu and B. J. Richmond. Response differences in monkey TE and perirhinal cortex: Stimulus association related to reward schedules. *Journal of Neurophysiology*, 83(3): 1677–1692, 2000.

A. W. Logue. The living legacy of the harvard pigeon lab: Quantitative analysis in the wide world. *Journal of the Experimental Analysis of Behavior*, 77(3):357–366, 2002.

A. W. r. MacDonald, J. D. Cohen, V. A. Stenger, and C. S. Carter. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, 288(5472):1835–8, 2000.

T. V. Maia and J. L. McClelland. A reexamination of the evidence for the somatic marker hypothesis: What participants really know in the Iowa gambling task. *Proceedings of the National Academy of Sciences of the United States of America*, 101(45):16075–16080, 2004.

H. Markowitz. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

M. Matsumura, J. Kojima, T. W. Gardiner, and O. Hikosaka. Visual and oculomotor functions of monkey subthalamic nucleus. *Journal of Neurophysiology*, 67(6):1615–1632, 1992.

S. M. McClure, G. S. Berns, and P. R. Montague. Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, 38(2):339–46, 2003.

A. N. McCoy and M. L. Platt. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience*, 8(9):1220–1227, 2005.

A. N. McCoy, J. C. Crowley, G. Haghighian, H. L. Dean, and M. L. Platt. Saccade reward signals in posterior cingulate cortex. *Neuron*, 40(5):1031–1040, 2003.

A. Minassian, M. P. Paulus, and W. Perry. Increased sensitivity to error during decision-making in bipolar disorder patients with acute mania. *Journal of Affective Disorders*, 82 (2):203–8, 2004.

J. Mirenowicz and W. Schultz. Importance of unpredictability for reward responses in primate dopamine neurons. *Journal of Neurophysiology*, 72(2):1024–7, 1994.

P. R. Montague and T. J. Sejnowski. The predictive brain: Temporal coincidence and temporal order in synaptic learning mechanisms. *Learning and Memory*, 1(1):1–33, 1994.

P. R. Montague, P. Dayan, C. Person, and T. J. Sejnowski. Bee foraging in uncertain environments using predictive hebbian learning. *Nature*, 377(6551):725–728, 1995.

P. R. Montague, P. Dayan, and T. J. Sejnowski. A framework for mesencephalic dopamine systems based on predictive hebbian learning. *Journal of Neuroscience*, 16(5):1936–47, 1996.

P. R. Montague, S. E. Hyman, and J. D. Cohen. Computational roles for dopamine in behavioural control. *Nature*, 431(7010):760–767, 2004.

J. Mossin. Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 34(4):768–783, 1966.

H. Niki and M. Watanabe. Cingulate unit-activity and delayed response. *Brain Research*, 110(2):381–386, 1976.

H. Nishijo, T. Ono, and H. Nishino. Single neuron responses in amygdala of alert monkey during complex sensory stimulation with affective significance. *Journal of Neuroscience*, 8(10):3570–3583, 1988.

Y. Niv, M. O. Duff, and P. Dayan. Dopamine, uncertainty and td learning. *Behavioral Brain Function*, 1:6, 2005.

J. O'Doherty, E. T. Rolls, S. Francis, R. Bowtell, F. McGlone, G. Kobal, B. Renner, and G. Ahne. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport*, 11(2):399–403, 2000.

J. P. O'Doherty, R. Deichmann, H. D. Critchley, and R. J. Dolan. Neural responses during anticipation of a primary taste reward. *Neuron*, 33(5):815–26, 2002.

J. V. Pardo, P. J. Pardo, K. W. Janer, and M. E. Raichle. The anterior cingulate cortex mediates processing selection in the stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences of the United States of America*, 87(1):256–9, 1990.

M. P. Paulus, C. Rogalsky, A. Simmons, J. S. Feinstein, and M. B. Stein. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage*, 19(4):1439–1448, 2003.

R. Potts. *Humanity's Decent: The Consequences of Ecological Instability*. William Morrow & Co., New York, 1st edition, 1996.

W. E. Pratt and S. J. Y. Mizumori. Characteristics of basolateral amygdala neuronal firing on a spatial memory task involving differential reward. *Behavioral Neuroscience*, 112(3): 554–570, 1998.

K. Preuschoff, S. R. Quartz, and P. Bossaerts. Human insula activation reflects prediction risk errors as well as risk. *submitted*.

K. Preuschoff, P. Bossaerts, and S. R. Quartz. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, 51(3):381–390, 2006.

L. A. Real. Uncertainty and pollinator-plant interactions - the foraging behavior of bees and wasps on artificial flowers. *Ecology*, 62(1):20–26, 1981.

L. A. Real. Animal choice behavior and the evolution of cognitive architecture. *Science*, 253(5023):980–6, 1991.

L. A. Real and T. Caraco. Risk and foraging in stochastic environments. *Annual Review of Ecology and Systematics*, 17:371–390, 1986.

C. Rode, L. Cosmides, W. Hell, and J. Tooby. When and why do people avoid unknown probabilities in decisions under uncertainty? Testing some predictions from optimal foraging theory. *Cognition*, 72:269–304, 1999.

R. Romo and W. Schultz. Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *Journal of Neurophysiology*, 63(3):592–606, 1990.

A. Rustichini, J. Dickhaut, P. Ghirardato, K. Smith, and J. V. Pardo. A brain imaging study of choice procedure. *Games and Economic Behavior*, 52(2):257–282, 2005.

E. Salinas and T. J. Sejnowski. Correlated neuronal activity and the flow of neural information. *Nature Reviews Neuroscience*, 2(8):539–550, 2001.

W. Samuelson and R. Zeckhauser. Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7, 1988.

M. Sato and O. Hikosaka. Role of primate substantia nigra pars reticulata in reward-oriented saccadic eye movement. *Journal of Neuroscience*, 22(6):2363–2373, 2002.

G. Schoenbaum, A. A. Chiba, and M. Gallagher. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, 1(2):155–159, 1998.

W. Schultz. Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, 14(2):139–47, 2004.

W. Schultz and R. Romo. Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *Journal of Neurophysiology*, 63(3):607–24, 1990.

W. Schultz, P. Apicella, and T. Ljungberg. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, 13(3):900–13, 1993.

W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

B. Shurman, W. P. Horan, and K. H. Nuechterlein. Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the iowa gambling task. *Schizophrenia Research*, 72(2-3):215–24, 2005.

L. P. Sugrue, G. S. Corrado, and W. T. Newsome. Matching behavior and the representation of value in the parietal cortex. *Science*, 304(5678):1782–7, 2004.

R. S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, 1988.

R. S. Sutton and A. G. Barto. Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, 88(2):135–170, 1981.

Y. Takikawa, R. Kawagoe, and O. Hikosaka. Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. *Journal of Neurophysiology*, 87(1):508–515, 2002.

J. Tobin. Liquidity preference as behavior towards risk. *The Review of Economic Studies*, 25(2):65–86, 1958.

P. N. Tobler, C. D. Fiorillo, and W. Schultz. Adaptive coding of reward value by dopamine neurons. *Science*, 307(5715):1642–5, 2005.

L. Tremblay and W. Schultz. Relative reward preference in primate orbitofrontal cortex. *Nature*, 398(6729):704–708, 1999.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944. by John Von Neumann, and Oskar Morgenstern. diagrs. 25 cm. "Corrigenda": leaf inserted. Includes indexes.

P. Waelti, A. Dickinson, and W. Schultz. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, 412(6842):43–48, 2001.

A. J. Yu and P. Dayan. Acetylcholine in cortical inference. *Neural Networks*, 15(4-6): 719–730, 2002.

A. J. Yu and P. Dayan. Uncertainty, neuromodulation, and attention. *Neuron*, 46(4): 681–692, 2005.

S. Zeki, R. J. Perry, and A. Bartels. The processing of kinetic contours in the brain. *Cerebral Cortex*, 13(2):189–202, 2003.

C. F. Zink, G. Pagnoni, M. E. Martin-Skurski, J. C. Chappelow, and G. S. Berns. Human striatal responses to monetary reward depend on saliency. *Neuron*, 42(3):509–17, 2004.