# Decision making, the frontal lobes and foraging behaviour

## Nils Kolling

**A dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Oxford.**

**Wadham College** **Trinity Term 2014**

# Decision making, the frontal lobes and foraging behaviour

## Nils Kolling

**Wadham College** **Trinity Term 2014**

**A dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the University of Oxford.**

**Approximate word count (excluding references): 52500**

## Overall Abstract

The aim of this thesis was to understand the function of the frontal lobes during different types of decisions thusfar mostly neglected in cognitive neuroscience. Namely, I sought to understand how decisions are made when comparisons are not about a simple set of concrete options presented, but rather require a comparison with one specific encounter and a sense of the value of the current environment (**Chapter 2-3**). Additionally, I wanted to understand how decisions between concrete options can be contextualized by the current environment to allow considerations about changing environmental constraints to factor into the decision making process (**Chapter 4-5**). At last, I wanted to test how the potential for future behaviours within an environment has an effect on peoples decisions (**Chapter 6**). In other words, how do people construct prospective value when it requires a sense of own future behaviours? All this work was informed by concepts and models originating from optimal foraging theory, which seeks to understand animal behaviours using computational models for different ecological types of choices. Thus, this thesis offers a perspective on the neural mechanisms underlying human decision making capacities that relates them to common problems faced by animals and presumably humans in ecological environments (**Chapter 1 and 7**). As optimal foraging theory assumes that solving these problems efficiently is highly relevant for survival, it is possible that neural structures evolved in ways to particularly accommodate for the solution of those problems. Therefore, different prefrontal structures might be dedicated to unique ways of solving ecological kinds of decision problems. My thesis as a whole gives some evidence for such a perspective, as dACC and vmPFC were repeatedly identified as constituting unique systems for evaluation according to different reference frames. Their competition within a wider network of areas appeared to ultimately drive decisions under changing contexts. In the future, a better understanding of those changing interactions between these prefrontal areas which generate more complex and adaptive behaviours, will be crucial for understanding more natural choice behaviours. For this temporally resolved neural measurements as well as causal interference will be essential.

This work was funded by a Wellcome Trust 4 Year Prize Studentship.

# Acknowledgments

Firstly, I would like to thank my supervisors continued support that began years prior to my doctoral thesis. Since the first year of my undergraduate studies, Matthew has always been a great mentor to me and I owe most of my scientific approach to imitating him. My good fortune and early successes are due to his amazing tutelage as well as the great environment within the laboratory making collaborations a professional and personal privilege.

I would also like to thank all the other people that were important to my academic growth over the years. Rogier Mars, who supervised my first undergraduate project and since has always had an open door for discussions of any kind. His advice, especially during the early years of my academic career, has been invaluable.

I also want to thank Tim Behrens, who's enthusiasm and computational skill often re-invigorated me when I was stuck.

I want to thank Jerome Sallet for both his scientific help, but importantly, also for his constant good spirits, which gave comedic relief when it was most needed.

Mark Walton, Rei Akaishi and Jill O'Reilly, all helped me to broaden my horizon through many discussions and sage advice.

I also want to thank all my Colleagues in the Laboratory, especially Bolton Chau Marco Wittmann and Urs Schueffelgen who were great Office mates and colleagues over the years.

I also benefited from great conversations with my friends over the years especially, Natalie Nelissen, George Wallis, Kathryn Atherton, Katharina Stumpenhorst, Max Sloan and Anna Huber, who must have been bored of my continued monologues over the years.

I also want to thank my family, who had to endure my monologues for many years longer than anyone else.

Most Importantly, I want to thank Jacqueline Scholl for being the best friend, companion, collaborator and wife, one can wish for. I am extremely privileged to have had the chance to spend both my professional and personal life with a person so insightful, kind and honest.

# Table of Contents

# Chapter 1: The frontal Lobes and Choice

## General Introduction

The question of what motivates human behaviour has always been a key part of philosophical and scientific inquiries. Whereas ethics has focused on understanding what ought to guide us in our decisions, more recently behavioural and cognitive sciences have concerned themselves with answering what factors actually drive human and animal behaviour. Most significantly, the advent of mathematical modelling and quantitative approaches to behaviour have made those questions much more tractable empirically. Today, it is possible to use such models in tandem with modern neurophysiological techniques to generate and test theories about the actual physiological/neural mechanism implementing behaviours as well as inform our models of behaviour through our understanding of physiological plausibility. However, being faced with such a complex system as the human brain and the infinite range of human behaviours, it is paramount to take a principled approach toward both the generation of candidate models of behaviour and the interpretation of complex neural data. In this thesis I am using an approach exploiting insights from classical economics, ecological and behavioural biology, cognitive psychology and modern neuroscience.

More specifically, experiments inspired by problems ubiquitous in the ecological context, and which are explored by optimal foraging theory, are used to understand human decision making both behaviourally and neurally. Neuroeconomics inspired methods are used to interpret the behavioural and neural data, except using categories and concepts derived from optimal foraging theory.

Conceptually, I am targeting three key issues in this thesis in particular:

1) The dynamics of foraging behaviours in human participants including their neural mechanisms. In other words, what factors drive the decisions of whether to engage with a particular offer or forage elsewhere instead and what neural activity underlies it? Here the dorsal part of anterior cingulate cortex (**dACC**) is highlighted in particular as driving foraging style choices and contrasted with ventromedial prefrontal cortex (**vmPFC**), which appears to be involved in the comparison between concrete offers.

2) Adaptive risk taking behaviours under changing environmental constraints. Namely whether people take into account a changing need to take risks when making gambling decisions and the neural mechanisms underlying this flexible value-driven comparison process. This is particularly inspired by the risk sensitive foraging literature, which states that the risk an animal is willing to take is a direct consequence of the risk pressure it finds itself under. The neural systems that appear particularly relevant for this rapid adaptation to changing constraints on decision making are the dACC and the vmPFC, potentially competing with each other, while each increasingly connected to more motor posterior cingulate cortex (**PCC**), when decisions are made that it appears to favour. Concurrently, the inferior prefrontal gyrus (**IFG**) activity predicts a bias in activity toward increased relative connectivity of dACC with PCC over vmPFC. Another candidate region for driving the strategic risk taking behaviour observed is the lateral frontal pole (**lFP**), which continuously tracked information relevant for adaptive risk taking.

3) The ability of people to show prospective foraging behaviour across sequential forage opportunities. These more sophisticated foraging strategies require people to simulate their own future decision strategies in order to

generate accurate value functions to guide their choices. Although participants were apt at exploiting prospective aspects of value, they also showed consistent biases such as over-committing to continue foraging, if they had already made such decisions earlier. The basic pattern of behaviour in this task comprises prospectively driven long term value decisions and the bias in behaviour that was observed may be related to the concept of effort fallacy. The paradigm can be used to relate such cognitive processes to underlying neural mechanisms, proposed to lie in lateral prefrontal areas such as the lateral frontal pole (**lFP**), dorsolateral prefrontal cortex (**dlPFC**), as well as the dorsomedial prefrontal cortex (**dmPFC**) and value driven areas such as the dACC.

# Theories of Choice

## The Expected Value Revolution

Since Blaise Pascal and Daniel Bernoulli, a very powerful tool for understanding human choice has been the concept of expected value. They described how decisions could be made by assigning each potential choice a specific numeric value of exactly how desirable any outcomes resulting from such an choice would be; After discounting each possible outcome X associated with a choice with its likelihood of occurrence by simply multiplying such a value with its respective probability p(X) and summing the product, one can theoretically derive a certain average expected value for any specific option.

$$\Sigma\ (p(X) * \text{Value}(X)) = \text{expected average Value of } X$$

An agent can then base his decision simply on selecting the option with the higher expected value. Equally when trying to understand behaviour, mathematical equations can be used to compute expected values for any number of options and predictions made about future behaviours by finding the choice associated with the highest expected value. Applying this selection criterion is sometimes referred to as using a “hard max” rule, because it assumes that the agent picks the more valuable option every time. Alternatively, it has been proposed that biological agents use more stochastic decision making strategies often called soft max decision rule. This could be to avoid complete predictability (Glimcher, 2004) or reflect an inability to compute expected values perfectly. Hunt and colleagues have recently discussed how choice stochasticity might emerge as a consequence of the features of a biophysically plausible choice evaluation network (Hunt, 2014).

Regardless of whether humans are deterministic or stochastic in their use of expected value, the general approach has enabled a very powerful formalism for describing behaviour. It postulates that by reducing any decision-making problem into a common currency, evaluation becomes the simple process of comparing two or more expected values and selecting the highest one. The only remaining issues would then be how humans are able to compute such values, i.e. evaluate their options, as well as how they compare any arbitrary set of options with differing expected values.

To distinguish between how humans ought to value specific options and their actual valuation, as inferred from their own subjective or revealed preferences (Kahneman & Tversky, 2000), the concept of subjective value has been introduced. This allows for modification of value functions due to for example risk aversion, delay discounting (Ainslie, 1975) and other factors such as ambiguity (Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005).

Indeed, with more sophisticated modelling it has been possible to look at more complex factors that affect valuation, such as the use of environmental meta-parameters such as volatility on the learning of reward values (Behrens, Woolrich, Walton, & Rushworth,

2007) and other Bayesian parameters such as Bayesian uncertainty (Yoshida & Ishii, 2006) or framing (De Martino, Kumaran, Seymour, & Dolan, 2006).

Fundamentally this approach still relies on the assumption that valuation might be difficult, complex, incomplete, stochastic or biased, but decision making is a simply a matter of selecting the option with the higher expected value or universal currency.

## Basic Neuroeconomic Theory

An economic perspective and its models have dominated decision neuroscience for the last fifteen years starting with the use of economic variables to explain neural recordings in macaques (M L Platt & Glimcher, 1999) and a subsequent special issue suggesting its generalized use in cognitive neuroscience (Neuron special issue 2002, edited by Glimcher; (Glimcher, 2002)). With it, many more quantitative description of the neural dynamics underlying decision making in the brain, have emerged (Boorman, Behrens, Woolrich, & Rushworth, 2009; De Martino et al., 2006; Hsu et al., 2005).

A number of studies have begun to validate behavioural or economic models, using proposed neural correlates of their model parameters/value distortions in the human brain (Hsu, Krajbich, Zhao, & Camerer, 2009; Tobler, Christopoulos, O'Doherty, Dolan, & Schultz, 2008). Indeed modern economic theory relies on a series of testable assumptions about the mechanisms underlying choices.

The first two axioms of completeness and transitivity are the simplest. The first proposes that people have both either a preference for or against an option compared with another or are always indifferent [A>B, A<B or A=B]. The second assumes that preferences are stable in that an option cannot at once be better than one alternative

and worse than a third which is less preferred than the second [A>B,B>C necessitates A>C].

The Continuity axiom proposes a continuous utility function, meaning it should always be possible to create a mixture of outcomes that are functionally equivalent to another alternative.

The last axiom is the axiom of independence. It assumes that the preference between two options is unalterable by context, i.e. not impacted by what other alternatives are presented nor by how those two options are framed. This however, has been questioned most frequently and very famously by psychologists demonstrating differing use of probability according to frames of loss aversion (Kahneman & Tversky, 2000).

Generally, economists have themselves recognized that humans are not perfectly following economic axioms. Such violations of expected behaviours, often referred to as fallacies, have been used to give insight into the neural mechanisms underlying choices (De Martino et al., 2006).

However, most economic and psychological theories are blind to considerations regarding how adaptive different mechanisms would be or their origins and equivalences in other species. Specifically, so far most models of choice have proposed one serial system that is involved in valuation, comparison and selection of an option (Padoa-Schioppa, 2011; Rangel, Camerer, & Montague, 2008; Rangel & Hare, 2010), since they are trying to solve the selection problem or any set of arbitrary options. There are no a priori distinctions between decisions based on different specific features, between the choice of an environment or the choice of a more circumscribed and specific offer or between different kinds of rewarding objects. There might easily be ecologically meaningful considerations and categorical differences between different kinds of value computations that violate the economic axioms, but which allow adaptive behaviours.

## Goods-based decision making

Padoa-Schioppa has argued that orbitofrontal cortex (OFC) neurons represent abstract economic value, the universal quantity Pascal already referred to as expected utility/value. He believes it represents foremost the chosen option, and makes decisions between different offers in “goods” or value space. He supports this claim mostly by his electrophysiological recordings of neurons from the central region of orbitofrontal cortex in the lateral bank of the medial orbital sulcus during a simple binary decision making task (Padoa-Schioppa & Assad, 2006).

In this task the animals had to decide between different amounts of two distinct types of juice rewards such as peppermint or grapefruit flavour. The identity of the reward each option, on the left and right side of the screen, were indexed by the colour of square targets. The number of targets comprising each option, on the other hand, corresponded to the amount of reward that would be received. Animals made decisions by saccading to either one of the two targets after a delay period. They were sensitive both to amount and type of reward, each animal had a distinct point of equivalence for a given pair of options. In other words they appear to trade-off the amount and identities of offers and at the point of equivalence being offered an extra amount of a less preferred juice can compensate for its less preferred status in comparison to a smaller amount of a preferred reward. This can be considered as the animals’ conversion rate for computing a universal currency between the qualitatively distinct offers.

The highest proportion of neurons appeared to respond foremost to the chosen offer value independently of the type of reward chosen, in essence encoding a universal currency, whereas other neurons encoded the value of only one specific offer. In essence, Padoa-Schioppa argues that the co-occurrence of offer and chosen value coding shows this region's involvement in the decision process, since those two pieces of information need to converge to make decisions in his tasks.

Subsequently Padoa-Schioppa further showed that neural responses were invariant to the overall context and whether a given type of reward was the preferred one in a given condition. The same offer value related firing occurred for the identical abstract values independent of the type of alternative offer available for a given decision (Padoa-Schioppa & Assad, 2008). However, neurons in OFC show strong range adaptation (Padoa-Schioppa, 2009) coding value relative to the values of available alternatives, a necessary feature for effective decision-making across broad ranges of values. In an independent line of argument he proposed that variability in firing of OFC neurons causes choice stochasticity (Padoa-Schioppa, 2013). He shows increased activity in neurons that represent chosen value if an offer of the same objective value is chosen under the context of a more compared to less desirable alternative. Thus, the only difference between difficult and easy decisions in his recordings was the value of the unchosen offer. He argues that the subjective value of the same offer needs to be higher to generate decisions of an offer with an increasingly desirable alternative. However, his results could easily be explained by range adaptation, since the value range is smaller in difficult decisions, as well as effects of overall value on the activity of "chosen value" neurons possibly mediated by non-selective neurons within the network (Hunt et al., 2012). His counterargument to such an interpretation would be that offer neurons were not effected by difficulty, although they do seem to be sensitive to range adaptation.

Thus, it seems that OFC neurons are sensitive to current subjective desirability which is both dependent on the range of values of possible alternatives and when an objectively unchanged offer is preferred over an increasingly desirable alternative. On the other hand, by only changing the type of alternative offers in the comparison, OFC encoding does not appear to change, although there are neurons dedicated to each individual taste and taste value/quantity.

More appealingly, an additional finding in the same recording experiment was that "juice-type" specific or taste neurons appeared to be more active, if that juice was subsequently chosen even before decision onset. This suggests the possibility of neurons in this region carrying information about the animals online/current preferences or desirability for specific types of rewards. This interpretation is further supported by lesion studies of the OFC, in the context of reward devaluation (Izquierdo, 2004; Peter H Rudebeck, Saunders, Prescott, Chau, & Murray, 2013). In these studies monkeys learn a series of visual discriminations some of which are reinforced by one reward type and some by another reward type. One of the rewards is then devalued by allowing the animals free access to the same reward and then animals are given choices between pairs of stimuli each associated with just one of the two rewards. Normally animals chose the stimulus associated with the reward that has not been devalued but this is no longer the case after an OFC lesion.

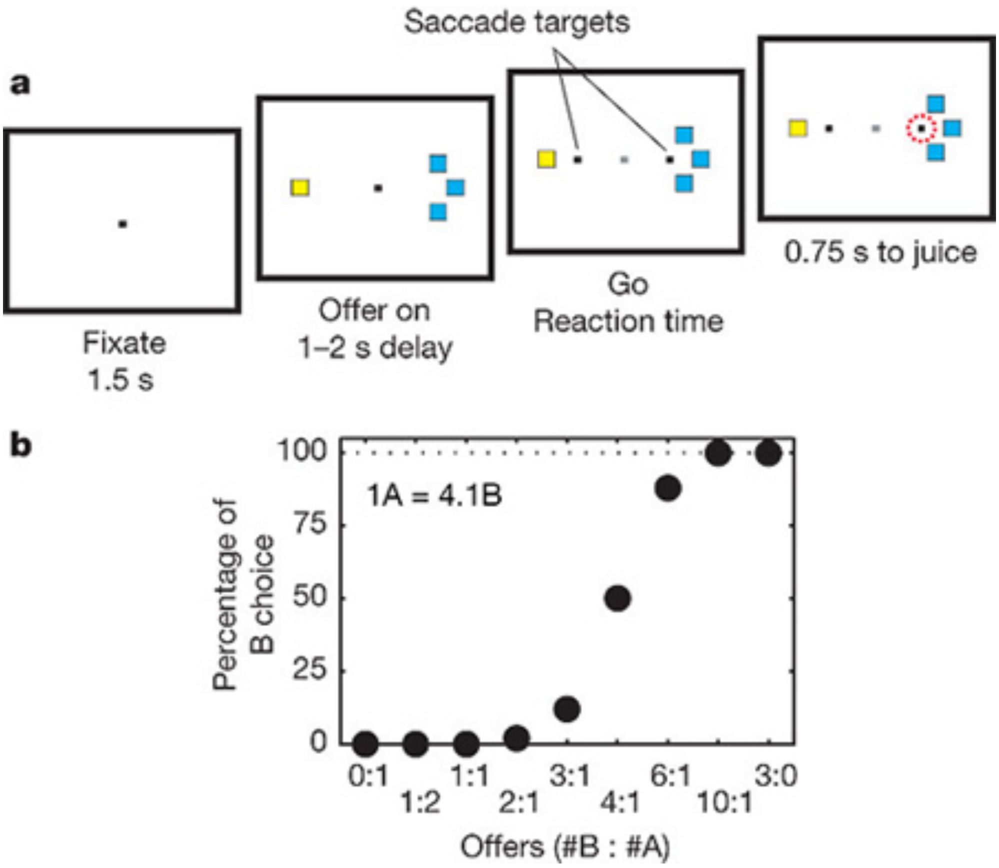

***Figure 1:*** *a) Timecourse of Padoa-Schioppa's neuroeconomics goods trade-off task: Two offers are presented to the macaque on the left and right side of the screen. The animal has to saccade to its preferred offer to select it. The type of reward delivered after a choice is presented in form of the colour of the rectangle on the corresponding side. The quantity of reward is equivalent to the number of rectangles. b) This plot shows the percentage of B choices as a function of changing quantities of B. Although offer A is generally preferred to B, all other things being equal, as the quantity of B increases, it is selected more frequently. If a sigmoid is fit to the data, it is possible to find the point of equivalence between the two offers, here 1A=4.1B (Figure taken from and part of* (Padoa-Schioppa & Assad, 2006)*).*

To summarize, there are two major parts to the goods-based account of decision-making. Firstly, that the co-occurrence of reward type specific and abstract chosen value neurons in a universal currency make the decisions that the task requires (Cai & Padoa-Schioppa, 2012; Padoa-Schioppa & Assad, 2006, 2008; Padoa-Schioppa, 2009, 2013). Secondly, that all value-guided decisions can be reduced to the type of decision studied in this investigation and could be accomplished by the types of neurons found in the OFC.

It is unlikely that the narrow part of central OFC that is recorded from is capable of all decisions in isolation. There is some evidence that the vmPFC, adjacent and heavily connected to the central OFC of Padoa-Schioppa's recordings, holds signals in a simple binary decision making task consistent with a value comparison process (Boorman et al., 2009) together making the orbitofrontal network a very good candidate for value-based comparisons and option selection. This signal furthermore has a temporal profile consistent with biophysical attractor model dynamics if decisions were made between different competing option specific neuronal pools (Hunt et al., 2012). More recently (McNamee, Rangel, & O'Doherty, 2013) showed that both category specific and non-specific information are encoded in the vmPFC, which is highly connected and adjacent to the central OFC, further supporting the claim of category specific information in the OFC system.

The vmPFC has recently been found to encode the value of novel goods, i.e. combination of flavours thus far not encountered, potentially allowing it to implement decisions even between imagined and not experienced options (Barron, Dolan, & Behrens, 2013).

In contrast, after carrying out similar experiments in the ACC, Padoa-Schioppa concluded that the ACC plays no role in decision making, the goods-to-action transformation as the decision is effected, nor any other online role in decision-making itself (Cai & Padoa-Schioppa, 2012, 2014). Cai and Padoa-Schioppa did, however, find neurons that appeared to encode the conjunction of saccade direction and offer value in the dACC during the post-offer or decision period and others that encoded the movement direction itself.

Arguably this sparse coding of both movement direction and reward might in principle be sufficient for effective decisions to be made using the action value information. Action value specific neurons can compete and drive decisions in action space, even if such neurons are not ubiquitous in a specific area. Particularly, if an area were responsive to many different types of tasks, actions and contexts, it would be surprising if the majority of neurons encoded one specific action. There are other problems with Padoa-Schioppa's reasoning. For example, in his task each juice type corresponds to one particular colour. Thus, it is not surprising that lateral PFC neurons encode juice/colour specific value early during decisions as well (Cai & Padoa-Schioppa, 2014), since lateral PFC has regularly been observed to be involved in arbitrary rule association and rule selection. It poses a larger problem to his argument, as lPFC and central OFC essentially encode the same decision parameters, thus meeting the only criterion he has set in order for a region to be regarded as a substrate for decision-making. However, the far larger problem with his theory is that he generalizes from the selection between qualitatively different kinds of concrete rewards such as apples and oranges, to all other types of decision. In other words, just because a region does not make decisions in a framework of apples against oranges, it does not mean that it does not participate any decision, or evaluation at all. Since we already know that macaques cannot use information regarding which primary reinforcer is associated with which particular abstract stimulus after OFC lesions (Izquierdo, 2004; Peter H Rudebeck, Saunders, et al., 2013), it is not too surprising that it should encode reward type specific value information.

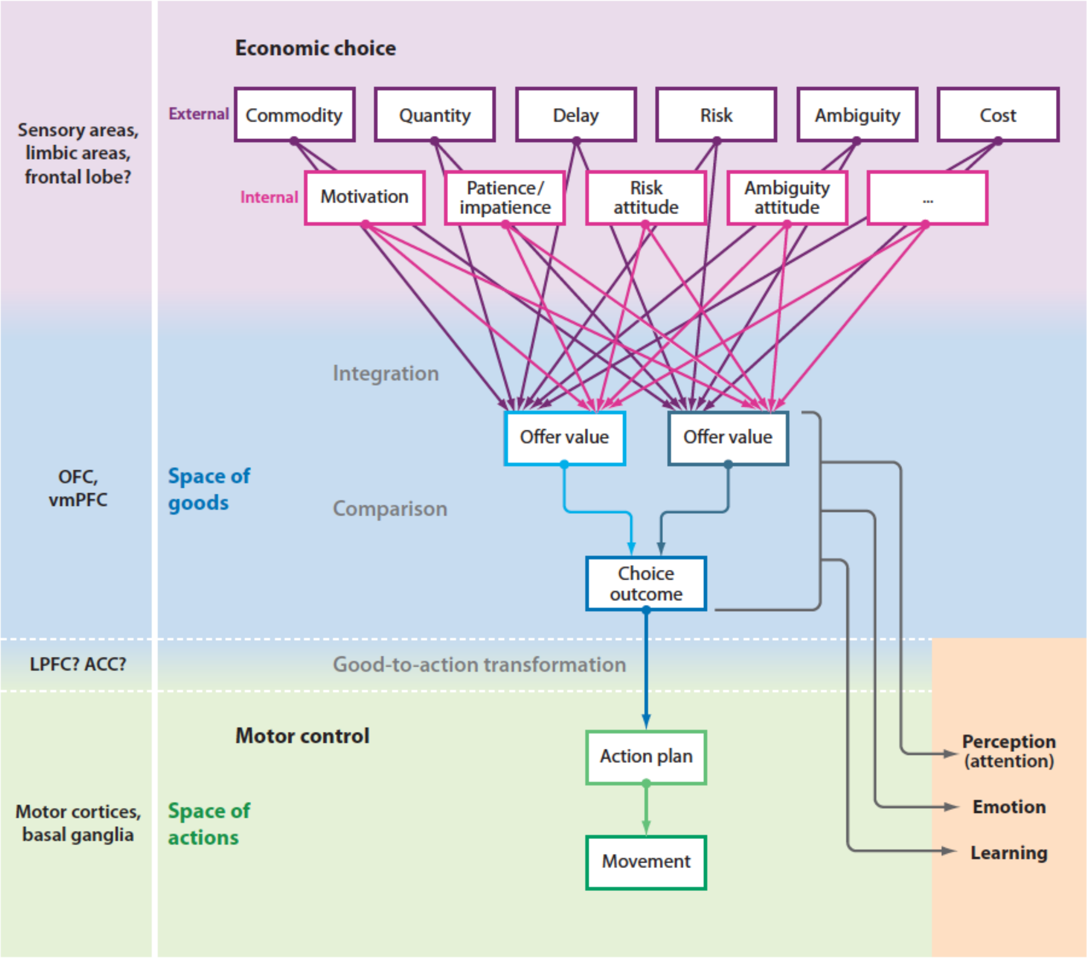

***Figure 2:*** *An example model from Padoa-Schioppa taken from* (Padoa-Schioppa & Cai, 2011) *offering a goods-based account of decision-making. Here all decisions are made in "Goods Space" and selection is the simple matter of comparing between all available offer values as soon as they are computed.*

## Action-based decision making

An alternative to the goods-based neuroeconomic account of decision-making is an action-based one. According to this theory, selection of a particular option occurs in relation to the action necessary to acquire the outcome rather than exclusively in a goods space that defines the outcomes themselves. There are very prominent proponents of action-based selection processes, especially in the context of saccade-based decision tasks (For a review (Gold & Shadlen, 2007)). They argue that even during perceptual discriminations of direction in random-dot-motion stimuli, evidence accumulation and choice happens in the system used to drive the action itself, in their case in form of a saccade. If animals need to make a saccade toward the direction of the coherent motion then lateral interparietal neurons (LIP) (Shadlen & Newsome, 1996) will accumulate evidence for motion direction for saccades into their receptive field. Additionally LIP can be considered a canonical accumulator system, since it is also sensitive to value other behaviourally relevant information. LIP neurons activity is biased by value (Dorris & Glimcher, 2004; M L Platt & Glimcher, 1999) and by prior information regarding which choice is more likely to be correct (Sugrue, Corrado, & Newsome, 2004). Furthermore, microstimulation of LIP neurons biases decisions in a manner consistent with evidence accumulation, as it effects proportion of responses and reaction times in favour of one decision, but only when evidence had to be accumulated to drive a discrimination decision (Hanks, Ditterich, & Shadlen, 2006). It is possible that other value-based decisions are made in connexion with other action systems, for examples those in other parts of the parietal cortex such as the medial intraparietal (MIP) area that more closely associated with limb movements. If decisions were to be made exclusively in action space, it would be necessary to employ more abstract representations of actions when an effector has not been specified, or alternatively assign a response modality temporally.

Wunderlich and colleagues designed a human fMRI experiment in which participants had to choose between two different kinds of action, an eye or hand movement, in a direction specified by a red dot appearing on the left or right side of the screen (Wunderlich, Rangel, & O’Doherty, 2009). They learned about the probability of reward for each type of action, eye and hand, independent of movement direction. In this task, the vmPFC was more active as a function of the chosen value, but dorsal anterior cingulate cortex was active as a function of the difference in value between the unchosen and chosen option. They took this as suggesting dACC is involved in the decision process, since activity scaling with the relative unchosen value difference, i.e. more activity with small differences between the options, could be to the hallmark of a comparison process; when the decision is difficult the decision process is more extended in time and the activity that can be recorded from ACC is greater. Due to fMRI's poor temporal resolution the other prediction from this suggestion, a slower evidence accumulation process on difficult trials, could not be validated in this experiment.

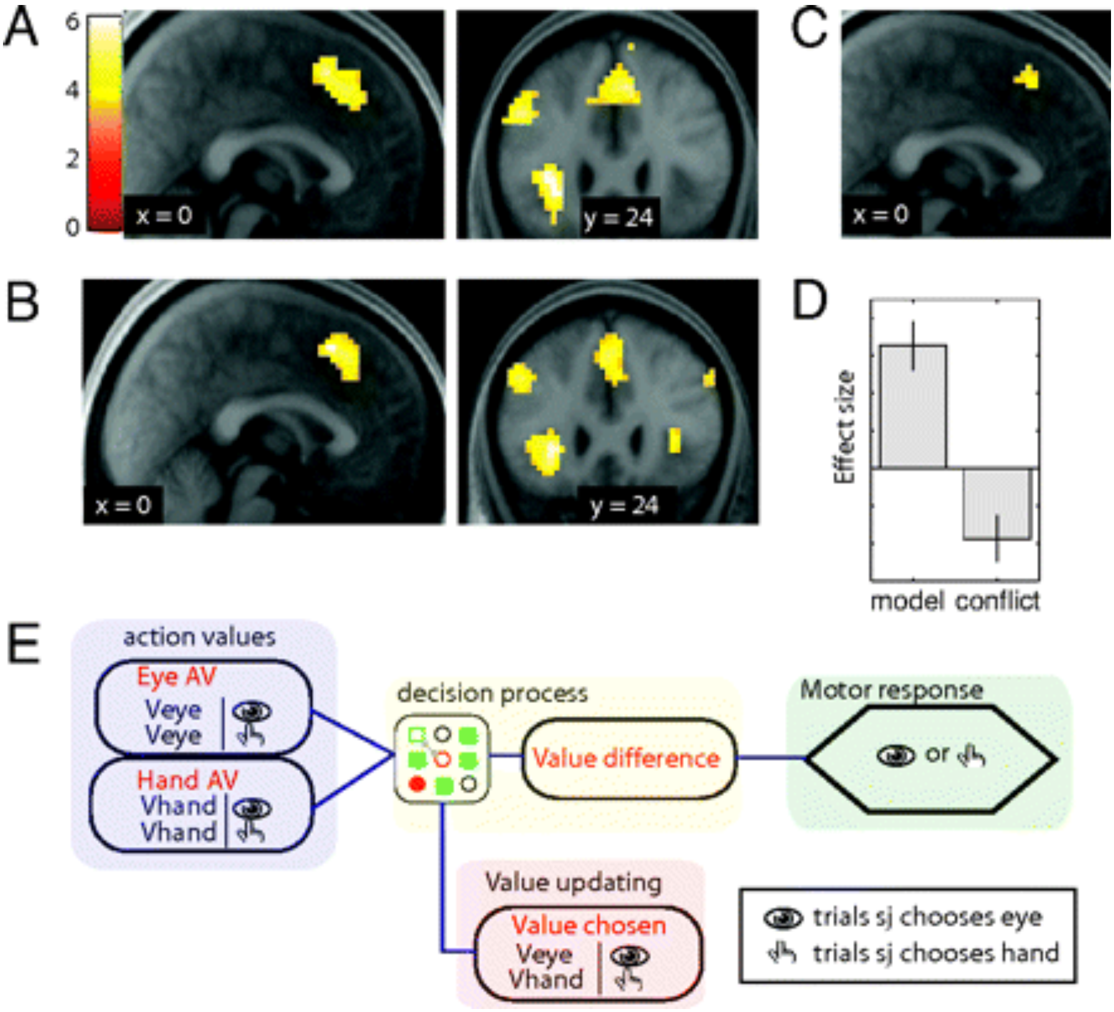

***Figure 3:*** *The results from Wunderlich taken directly from* (Wunderlich et al., 2009)*. Participants had to learn whether to make a saccade or a button press response into a direction specified by the location of a red circle on the screen. A) Activations related to the difference between the unchosen and chosen action values. B) Shows a regression with the output of their diffusion decision model, revealing the same areas. C) Shows this*

*activity exists to some extend even when no obvious error occurred. D) compares the effect sizes of the model output with conflict. A decision process clearly explains the neural data better than simple conflict. E) A schematic of how decisions could be made in action-space between eye and hand movements.*

With an almost identical design to (Padoa-Schioppa & Assad, 2006) Hare and colleagues (T. A. Hare, Schultz, Camerer, O'Doherty, & Rangel, 2011) again suggested an action-based selection process. They were inspired by competitive comparator mechanisms called diffusion decision models (Ratcliff & McKoon, 2008) of which a biophysically plausible version is the model put forward by Wang (Wang, 2002) and used for value-guided decisions in humans (Hunt et al., 2012). They used this model to generate their own predictions of how the activity should evolve if a region was participating in of the process of the value comparison. Again, activity in the dACC was a focus in this study; activity in dACC increased when the options' values were closer together and more difficult to choose between. By contrast activity in the vmPFC correlated with the sum of values of the choices that were presented. Investigating the neural dynamics during the task further, they showed dACC changing its functional connectivity with the left or right motor cortex, depending on which hand movement was ultimately chosen and a generally increased connectivity with vmPFC during the choice. This further suggests that dACC's involvement in the comparison process, especially if different effectors are involved.

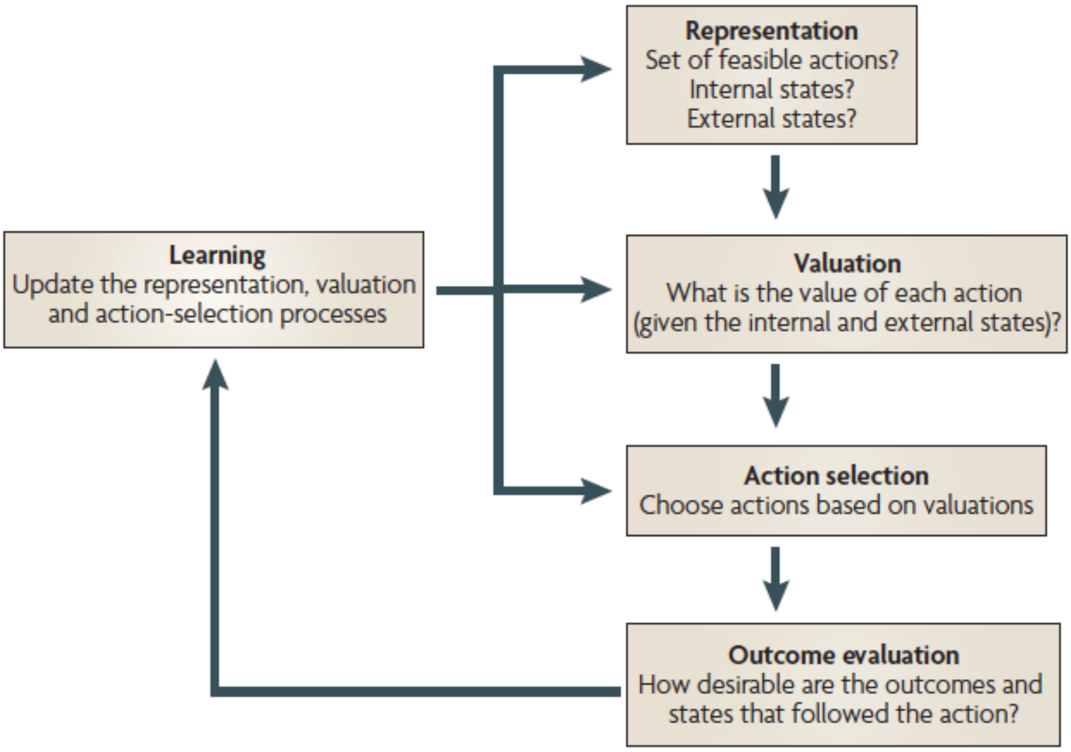

***Figure 4:*** *Selection model using actions, i.e. putting the comparison process on the level of actions. Those actions are each associated with value through a preceding valuation process that takes into account many factors such as prior experiences and internal states and external constraints. (Schematic taken from* (Rangel et al., 2008)*)*

To summarize, the action-based account of economic choice posits decisions are made by directly comparing evidence in favour of different courses of action in the framework of the specific actions required to obtain such goals. Even though this has been a very popular model for visuo-oculomotor decisions it is only recently that value-based choices have been considered and tested in this framework. However, one important restriction to this account is, how are decisions made before the action required to obtain any offer is known? According to the strictest interpretations of this approach, decisions are not possible without at least rudimentary or temporary actions assigned to the options being considered, i.e. decisions cannot happen in purely abstract or stimulus spaces.

## Actions or stimuli – what are we choosing?

To explicitly contrast both action and stimulus based decision theories, Wunderlich and colleagues presented participants with a decision between two different stimuli to choose between either before revealing the actions required for selecting either stimulus or concurrently with a stimulus-action assignment. Participants had to track the changing reward probabilities associated with each stimulus across trials. The information about stimulus-action association was only useful in as far as it told the subjects what to do in that specific trial to acquire the stimulus of their choice. This revealed that participants were able to make decisions purely in goods-space and that such comparisons appeared to engage the vmPFC, but not the dACC (Wunderlich, Rangel, & O'Doherty, 2010). However, in this paradigm actions were so tangential to the task that participants had to focus completely on stimulus identity. The absence of evidence for a neural mechanism encoding action values may have been due to the absence of any meaningful and consistent link between actions and values in this particular task.

On the other hand, if macaques need to associate a particular value to an action, as was the case in Wunderlich's earlier paradigm, and there was no possibility of learning an association between stimuli and rewards then lesions of ACC led to impairments (Peter H Rudebeck et al., 2008). When, by contrast, the macaques learned about stimulus values then no such impairments were found (Peter H Rudebeck et al., 2008). Instead OFC lesions impaired learning. One interpretation of such a pattern of results is that there are separate mechanisms for learning action-based and stimulus-based reward information. The same double-dissociation has been since confirmed with human patients with lesions in either ACC or OFC (Camille, Tsuchida, & Fellows, 2011). Neurons in the OFC of macaques appear, correspondingly, more sensitive during decisions between stimulus-based associations, compared to action-based associations, whereas in the ACC the opposite pattern was found (Luk & Wallis, 2013).

Action and stimulus-based associations are qualitatively distinct in at least one respect. Whereas there is usually no inherent relationship between the value of an item and any cost that it might impose on an agent, for example in terms of the effort that is needed to choose it, costs such as effort are intrinsic to action-based choices. Thus, it is not very surprising that effort and action-related information could change the neural mechanisms of decision making (Peter H Rudebeck, Walton, Smyth, Bannerman, & Rushworth, 2006). If rats had to choose between different effort costs, lesions of ACC impaired their ability to pursue the more effortful goal, whereas OFC-like lesions impaired delay-based evaluations, clearly dissociating the requirements for comparison of delay and effort information. Subsequently, it has been shown that effort costs activate the ACC but not vmPFC (Croxson, Walton, O'Reilly, Behrens, & Rushworth, 2009) and decisions made on the basis of effort/reward trade-offs engage the ACC in ways not observed in the vmPFC, whereas the vmPFC encoded delay based costs (Prévost, Pessiglione, Météreau, Cléry-Melin, & Dreher, 2010). Moreover single neurons in ACC have now been shown to code the integrated utility of cost benefit decisions (Steven W Kennerley, Behrens, & Wallis, 2011).

To conclude, even the studies adopting a more or less neuroeconomic perspective suggest that decisions are made using different frameworks. For example, value can be assigned either to actions or stimuli and comparison is then likely to be made in relation to actions or stimuli. However, debates are still fiercely fought over the nature of a single general unitary decision framework employing a single and universal decision making mechanism.

However, the alternative to such an approach would be to assume the existence of multiple concurrent decision making mechanisms, each using a distinct brain structure to converge on choice signals that have the capacity to drive voluntary behaviours in situations where there are multiple choices. In other words, the brain concurrently solves the problem of what to choose using a variety of possible approaches as well as selecting between different ways of choosing on a more network-based level. For such a view to result in any significant gain in our understanding of general decision mechanisms, it is critical to use a principled approach to consider what might be the unique reference frame for any particular brain region when it contributes to decision making.

Another distinct framework, following naturally from foraging-style perspectives to decision-making, is that the key decision to make is between the value of an option that is encountered and the average value of the environment; the choice is whether to engage with an option that is encountered or whether the environment is sufficiently rich that it is worth refraining from engaging with the encountered option and continuing to search for better alternatives elsewhere. In other words, a key issue under such a framework is to work out what are the rules that guide decision making within an environment between options and between environments themselves? Is there a specific neural mechanism underlying these types of decisions, optimized to solving this unique and very important ecological challenge?

Before elaborating on the implications of an ecological modelling for studying decision making mechanisms, it is important to outline the major types of ecological models first.

## Ecological Approach to Decisions

Firstly, it is important to establish the broader framework of ecological modelling approaches. Optimal foraging theory tries to establish in every instance, what is crucial to maximize for a biological agent in an ecological environment. From there more specific models are generated to explain the particular behaviours at hand. In essence it is an evolutionary perspective suggesting that behavioural strategies evolved because they aided an animal's survival. Whichever constraints are particularly relevant for survival are under the strongest selection pressure and over time optimized the most rigorously. To understand a particular adaptation/behaviour, one must identify the most important constraints and goals and use optimal foraging theorems to generate a set of equations that ought to describe ecological behaviour if it is indeed adapted to the particular problem at hand.

Another special aspect of ecologically and biologically informed models is that they make distinctions between common vs. different mechanisms and common vs. different functions. In other words, a very important ecological problem such as basic foraging behaviour could have several different mechanisms as a solution, whereas a powerful mechanism evolved to solve such a problem could also solve different computationally similar problems.

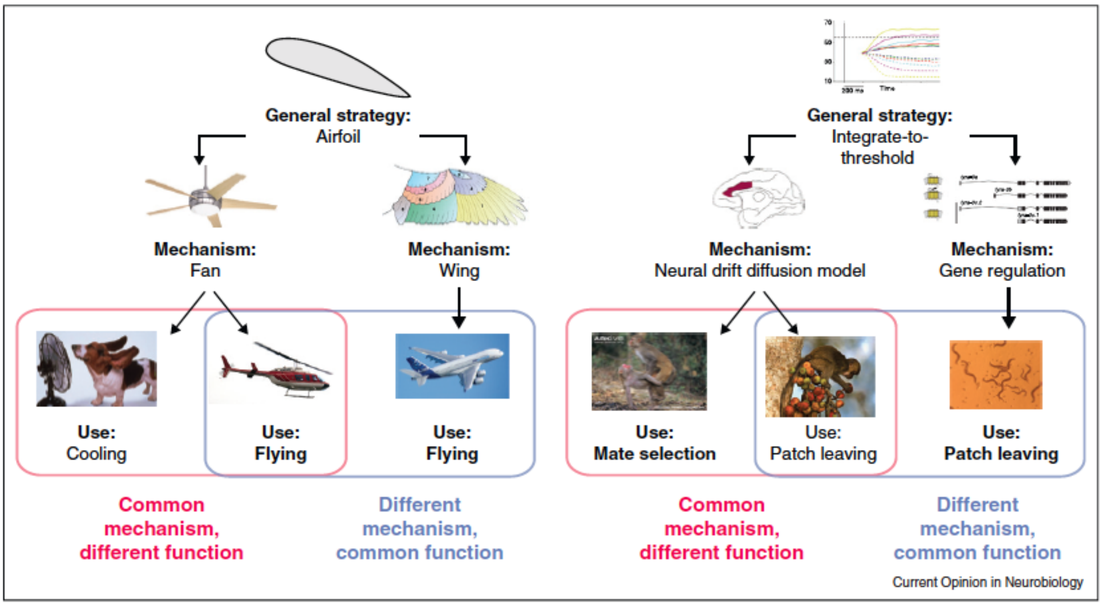

***Figure 5:*** *An example of how a general approach or concept, here called strategy, can be implemented by different mechanisms. These mechanisms share functions such as flying or patch leaving as well as one mechanism possibly having more than one use such as Neural drift diffusion as a mechanism for accumulation of evidence allowing both patch leaving and mate selection. Figure directly taken from* (Pearson & Platt, 2013)

For understanding human cognition and decision making, ecological models of animal behaviour offer powerful means for describing behavioural adaptations in ways particularly relevant for solving naturalistic problems of behavioural optimisation and ultimately survival. Since humans, like other animals, evolved to solve such problems, it stands to reason that such an approach could help uncover the principal components of human cognitive functions particularly when it comes to making adaptive choices.

More specifically, it is quite plausible that unique neural mechanisms exist for different kinds of behaviours, such as foraging versus precise comparison between a limited number of specific objects and their selection. Indeed some have suggested that precise deliberation for comparison between concrete options only makes sense for animals with long distance vision, specific associative learning, capacity for credit assignment and exploitable knowledge about causal structures together allowing sophisticated goal directed behaviour. According to this perspective it comes as no surprise that the vmPFC has been proposed as a region that compares the values of concrete options. It is particularly pronounced in non-human primates in comparison to rodents (Carmichael & Price, 1995; Wise, 2008). Furthermore, in interaction with the central and lateral OFC, vmPFC could allow selection between stimuli on the basis of their specific value that have been learned through means of credit assignment (Mark E Walton, Behrens, Buckley, Rudebeck, & Rushworth, 2010) and in relation to their motivational significance (for example as studied in reward devaluation tasks – (Izquierdo, 2004; Peter H Rudebeck, Saunders, et al., 2013).

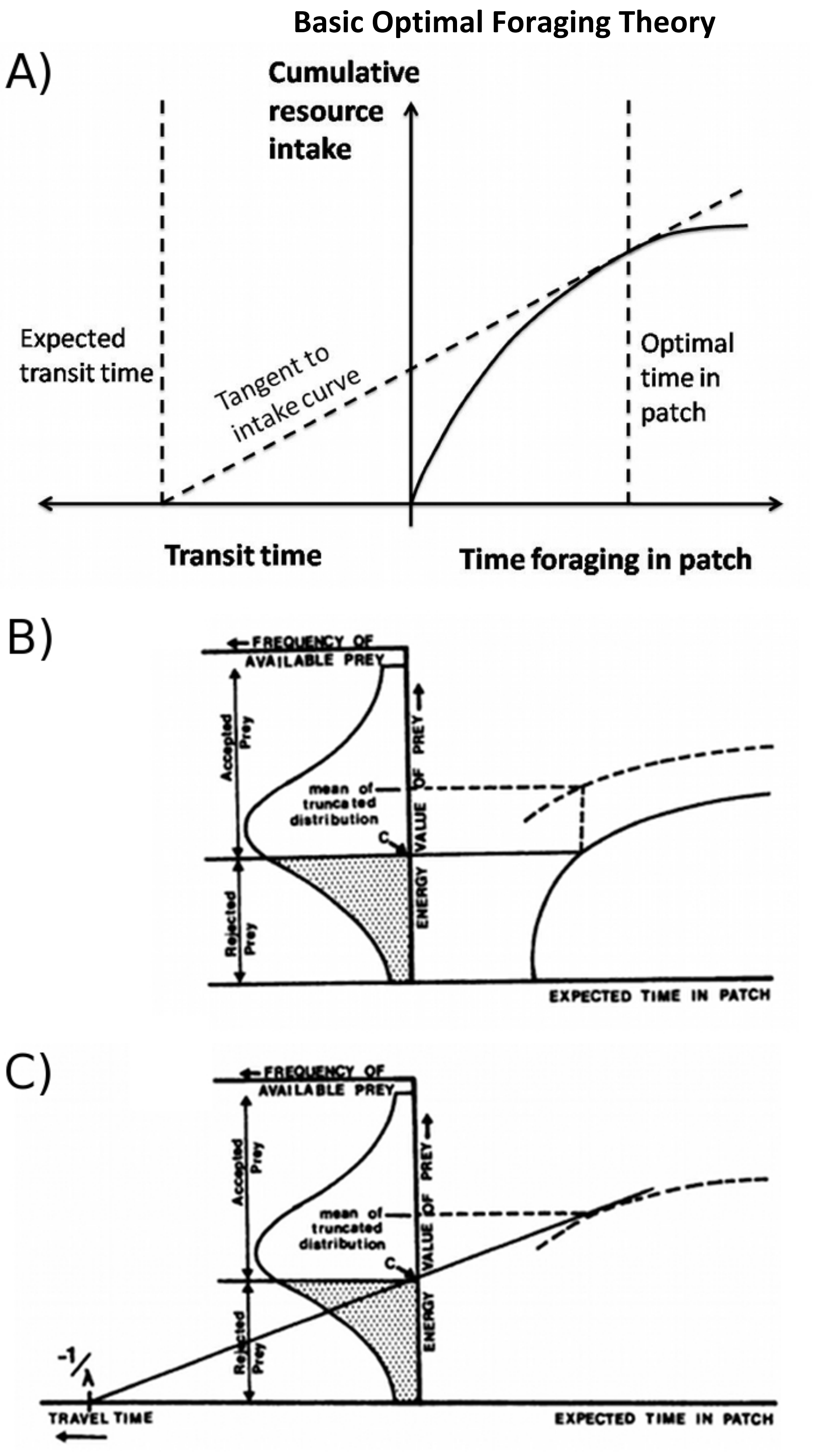

***Figure 6:** (A) A depiction of the relationship between the most basic components of a patch leaving decision according to the marginal value theorem (MVT). It shows how expected transit or travelling time and the harvesting rates effect the optimal leaving time in a specific patch to maximize the average resource intake. (B) A more complex*

*model of optimal prey selection, if an agent can only engage with one particular prey item at a time. The Single Prey loader model (taken from Stephen and Krebs 1986* (Stephens & Krebs, 1986)*) suggests a way to determine whether to include a particular reward value or prey in the current prey selection. Whereas the model in (A) has a continuous resource intake, the prey model as a distribution of possible prey items being encountered. The animal has furthermore the choice to continue searching within the habitat, rejecting the current prey if it falls below a certain threshold c. This of course increases the expected time in a particular patch, but also increases the average expected resource gain within that patch. The solid line shows the relationship between increasing levels of c and increased time in a patch, whereas the dotted line indexes the resulting average expected gain from the now truncated distribution of reward outcomes. As the chosen threshold together with the prey distribution completely determines expected patch time, and with it expected average gains, it is the only free parameter to determine optimal preference. (C) Just shows how different travelling times effect the optimal threshold c and how, again like in (A), the "tangent solution" can be used to find the optimal leaving time and with it c value for a given travel time. Overall average reward rates can be determined by taking the average expected gains and dividing it with overall time.*

The most famous example of optimal foraging models is the marginal value theorem. At its heart it is an elegantly simple solution to the complex problem of when to best stop "harvesting" a depleting resource in favour of a more plentiful potential alternative. To solve this problem the theorem in its simplest form relies on the use of the average reward or intake rate, which has to be optimized over the longer term future. The intake rate of course is nothing more than the cumulative resource intake divided by time. However, the use of longer term intake rate estimates allows such models to integrate a variety of constraints such as travelling or transit times as well as other forms of delays through their effect on the average rate. The model easily predicts the patch leaving time for a given patch in a graphical manner if one draws the tangent to the intake curve (starting from the point corresponding to the travel time), which represents the highest average rate and therefore optimal leaving time for a specific patch (see Figure 6A). More complex models can also implement rules for the acceptance/rejection of particular prey items within a specific environment by using a simple threshold rule (Figure 6B,C). In Chapter 6 of my thesis I will be using a threshold model inspired by this approach toward future option selection.

The so called zero-one rule in foraging states, like the Hard-Max rule in Economics, that animals should only and always select food items above their desirability threshold and reject the ones below it. Debates about the validity of this are equally fierce as those that concern the nature of stochasticity in human decision making, but in both cases useful predictions can be made with these models even if the rules are not applied in an all or none fashion.

Experimentally, many of those theories were tested in birds such as great tits in elegantly simple experiments. One of the most basic setups uses a conveyor belt with different worms, large and small, being presented to the bird with variable frequencies and variable handling times/difficulty by wrapping those worms with more or less sticky tape. Using such a setup the birds were shown to be sensitive to the frequency of good and bad worms, i.e. environmental richness, as well as to the increased time it took the animals to actually acquire the food item, i.e. handling time (Stephens & Krebs, 1986).

As in the above-mentioned models and experiment, many ecological foraging-style decisions occur in the form of sequential encounters, during which the decision is about accepting or rejecting a concrete offer, which is compared against a sense of what the current or future environments have to offer. This asymmetry of options, rather than computationally equal concrete offers, which is common in neuroeconomics paradigms, is one of the obviously distinct features of foraging distinction. However, it is crucial to highlight that rejection of a concrete offer is not just a selection of a valueless nothing, but a strongly value-based comparison between two different deliberate behaviours of either engaging with the option at hand or following the possibility of higher longer term values in the environment. However, environmental value is sometimes underdefined and certainly requires different information than the comparison of a selection of concrete potential outcomes, which the OFC/vmPFC appears to be exceptionally well equipped to handle.

In another series of experiments animals such as birds have also been shown to be very sensitive to their current energy budget when making decisions between more profitable but also more risky options and less profitable but also less risky alternatives. In risk sensitive foraging theory animals risk taking behaviour is indeed considered to be a direct function of the animal’s current needs and context i.e. internal and external risk pressure (Caraco, 1981; McNamara & Houston, 1992). This is a very potent example in which basic parameters of economic theory such as probability and reward size should be considered very flexibly, depending on ecological constraints important for survival. It suggests a very dynamic modulation of a behavioural parameter, namely risk aversion/seeking, often considered to be fixed according to neuro-economic models.

One question that has not been directly addressed using a foraging approach is the use of future or prospective behaviour to determine the value of future foraging states. Especially during sequential behaviours it is crucial to take into account not just the next outcomes of foraging, but also the overall average value of a state that an agent might enter given a certain behavioural strategy. Foraging theory offers the notion of a fixed optimal threshold for determining such future behaviours, allowing the calculation of utility of sequential foraging. However, an important question for human foraging is, how deliberate this computation is and whether it can be modulated by information such as varying the number of future opportunities, restricting possibilities of using future behaviours to increase the value of foraging. Neurally, this is particularly interesting, because a true simulation of one's own future decisions, given certain outcomes, is a complex process, which might rely on additional neural circuitry. Such circuitry could in turn be used to simulate other people's decisions and underlie theory of mind functions. Independently, prospective foraging behaviours could be a strong ecological argument for the development of certain computational mechanisms that take into account additional information about future environments and simulated behaviours in it. By computing the consistency between such simulations and later choices one could also shed more light on the origins of choice inconsistency in more natural settings.

A distinction in perspective between economic and psychological theories of decision-making and a biological ecological approach is that decisions in ecological terms are a lot broader. In other words, adaptive responses in ecological models can be implemented through selection processes and genetic mechanisms, longer term developmental pathways as specific responses of the biological agent itself, whereas the psychological concept of a decision is more restricted and internal. Indeed ecological modelling does not assume the processes of adaptation to be intrinsic to a specific agent itself, whereas this is a necessary component of a psychological definition of most decisions. This does however, not imply that an ecological approach towards decision making is inapplicable to internal decision making mechanisms confined to the human brain. The human brain could have evolved to actively compute internal variables that allow a specific agent to generate adaptive and flexible decisions, whereas other organisms might need to adapt through selection over generations. In other words, humans might be able to implement higher order adaptability through computational processes that are internal to their brains using a unique mechanism to solve the same ecological problems that other organisms also face.

It is therefore not surprising that one candidate region for the implementation of foraging behaviours is ACC, particular its dorsal part, since it belongs to the evolutionarily oldest and most preserved parts of the medial prefrontal cortex.

Thus far I have discussed the idea that 1) foraging-style decision problems are ubiquitous for humans and other animals alike, requiring a unique comparison process to optimize the ability to select optimally in this set of problems. 2) It is therefore not surprising that an area that is relatively preserved across species such as the ACC, might be involved in such comparisons, forming the most basic mechanism for foraging decisions in the cortex. 3) However, I have also highlighted the fact that many very simple organisms can follow behaviours prescribed by optimal foraging theorems and are thus able to forage effectively and adaptively. This is however, not at odds with ideas of ACC being important for foraging functions in mammals, because there is both the possibility for multiple systems that can generate foraging behaviours within an organism, but also different mechanisms solving common problems in different species. What is crucial in this perspective is not that all species use the same mechanism such as a specific cortical brain area to solve a specific problem, but that they all have found adaptive solutions to a common problem exploiting optimally efficient mechanisms for their specific environments and constraints. The primate ability to perform very complex, but effortful foraging and information seeking behaviours could have led to the evolution of a mechanism in form of the ACC that allows longer-term value driven often sequential actions, taking into account their effort costs and the current environment.

Foraging tool-using omnivores like humans are particularly invested in making longer term value decisions accurately. They frequently need to decide whether to engage with a particular offer or forage elsewhere instead. Sometimes this requires developing very elaborate and effortful reward seeking behaviours with very long term reward horizons. They furthermore, need to often consider their environment and longer term consequences of their behaviours, to dynamically modify the way they evaluate specific offers.

Of course it is possible to refer back to formal formulations of expected value-based decision-making problems. However, the crucial question is whether it is particularly helpful for understanding the neural mechanisms underlying the ability of biological agents to perform adaptively in a natural environment. Alternatively, it can be argued that neural circuits develop to generate adaptive behaviours against environmental challenges and that since neural systems are about solving those kinds of challenges, they probably are structured in a way as to solve them optimally, even if this means violating economic principles. Therefore, when trying to understand prefrontal cortex function in decision making, being able to draw from potentially neurophysiologically meaningful, but certainly ecologically essential distinctions, could have many advantages. It not only argues for fundamental frameworks for foraging-style behaviours, but also suggests what kind of quantities might be computed, represented and used to drive behavioural adaptations.

One might ask whether foraging is even a behaviour that humans pursue, or whether they rather use different, unique mechanisms to guide their behaviours. It is certainly true that many of the decisions and behaviours described in this thesis could be discussed without using foraging terminology. However, humans do engage with search behaviours, evaluate whether to engage with an offer or reject it. They are further, very sensitive to environmental value information when making those considerations in ways very similar to foraging animals (see Chapter 4). The experimental chapters of my thesis further show that there seem to be neural systems dedicated to solving such kinds of comparisons.

Another interesting point about functional differences in foraging behaviour is that it is often under/undefined in terms of outcomes, objects, offers, but specified in terms of behaviours, actions and costs. In other words, what is selected is an unspecified reward environment with the prospect of future longer term beneficial outcomes

So far, a selection of brain regions, most prominently the dACC and the vmPFC have emerged as having particular relevance for simple comparison processes, but a broader reference to the activity that they have been reported to have in decision making is still lacking.

In the following paragraphs I will focus more on each brain area's activity and anatomical connectivity, separating them into the orbital prefrontal cortex (**OFC**), the vmPFC , and the **ACC** including the dACC as well as the lateral PFC areas of dorsolateral (**dlPFC)**, frontal pole (**FP**) and inferior frontal gyrus (**IFG**). In the light of the outlined ecological perspective, I will briefly interpret some of the major observations on the activity they exhibit in order to make an argument for some of their functions.

## Categorizing and Measuring Frontal Lobe Function

Lesions of the prefrontal cortex were among the first to spark an interest in the relationship between the human brain and behaviour, due to their drastic effects on patients' emotions, traits and behaviour (see the case of Phineas Gage, 1823-1860).

## Neuropsychological methods

Since the earliest days of neuropsychology/physiology there has been immense progress in dissociating the specific functions of the different regions of the prefrontal cortex in supporting aspects of decision making. Most of this advancement can be attributed to a wealth of new neuroscientific methods unavailable to the earliest pioneers of neuroscientific research.

Firstly, it is now possible to map behavioural changes measured by far more selective cognitive tests onto specific sub-regions using structural brain scans through lesion overlap analysis. Acute and chronic animal lesion studies allow for even more selective and controlled interference of only one specific brain area or in some cases only the interactions between areas (see pathway lesions and cross lesions), further testing causal relevance of very specific systems for different functions.

## Electrophysiological methods

Furthermore, we are not limited to just aligning different dysfunctions to distinct brain areas. We can also measure the functioning brain and relate its activity to very specific aspects of the external or internal environment, the planning or execution of behaviours or more intermediate processes such as the selection/comparison process itself.

Nowadays it is possible to measure neural firing patterns in very localized parts of the brain directly either by using single electrodes or more exhaustively through whole multi-electrode arrays. However, most neural recording studies are limited to animals such as macaques, rats or mice, since human recording studies are only possible in specific patient populations such as people suffering from chronic epilepsy. Furthermore, electrode placement in humans is dictated not by scientific considerations, but by medical relevance, making it harder to have consistent samples. Electrode recordings are further limited by the manner in which neurons studied are selected. Using a task or reward sensitivity criterion during electrode insertion biases results in favour of highly active (Shoham, O'Connor, & Segev, 2006), often larger, neuron populations leading to an under-representation of other neuronal populations within the network.

Still, knowing the correlation between neuronal firing and a specific feature of the environment or intermediate cognitive processes, we can induce firing in a subset of neurons through electrical stimulation and observe the effect on behaviour. These tests of causal relevance of particular neural codes are incredibly useful but only possible when a spatial organization of the coding allows for specific stimulation of some neuron types preferentially. However, new methods such as optogenetic (E.g. DA neurons firing, (Tsai et al., 2009)) and magnetogenetic stimulation are promising increasingly targeted stimulation and therefore unprecedented experimental control for probing neural networks.

Additionally, there are many more methods enabling the testing of theories about neurotransmitter functions in a highly temporally resolved manner using e.g. fast scan cyclic voltammetry (Gan, Walton, & Phillips, 2010; Phillips, Robinson, Stuber, Carelli, & Wightman, 2003), or tonicly using dialysis. Alternatively, optogenetic stimulation can be targeted to neurotransmitter specific neurons to test the causal effect of such release.

## Human functional imaging

With humans the possibilities are sparser. We rely heavily on indirect inferences about neural activity from measurements of Blood-Oxygenation-Level-Dependent (BOLD) signals in MRI scanners. Although spatially constrained to the activity across many neurons within usually 3x3x3mm large voxels and temporally smeared to a resolution of a few seconds, it allows us to measure the neural correlates of complex human cognitive processes throughout the whole brain. The BOLD signal is an combination of many different aspects of neuronal dynamics within a voxel. Logothetis and colleagues (N K Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001) showed BOLD signals to most strongly co-vary with local field potentials (LFP's), but also neuronal spiking i.e. action potentials in visual cortex. Since pre-synaptic activity underlies most of the LFP, it is reasonable to assume that the BOLD co-varies with both inputs to the region as well as internal computations within in it. However, since such studies are limited in comparing one recorded brain region using electrodes to the corresponding MRI signal, the concrete relative contribution of post-synaptic spiking against long-range inputs from other areas and inter-neuronal computations, is unknown for most cortical regions, except early visual areas. Thus, we can assume that the BOLD signal reflects the computational properties of a region through correlations with its input, output or intermediate parameters computed within, with varying sensitivities.

We can further use the BOLD signal to investigate functional networks e.g. by inferring a changing functional coupling between those areas from changing correlations between brain areas as a function of a psychological parameter. This method is often referred to as psychophysiological interaction (PPI) analysis (Friston et al., 1997) as it looks at changing interactions between different physiological signals as a function of a psychological manipulation.

For a better temporal resolution and the possibility to investigate oscillatory properties in humans, electro-encephalography (EEG) and more recently intercranial EEG (iEEG) as well as Magneto-encephalography (MEG) are available. However, the increased temporal resolution comes at the price of bad spatial localization in MEG and even more so in EEG, or specific limitations to patient populations and lack of whole brain coverage in iEEG.

Instead of the more selective stimulation possible with animals, in humans we can stimulate a relatively large area of brain tissue, probably a centimetres or so in diameter, using a technique called Transcranial Magnetic Stimulation (TMS) for a couple of milliseconds. With some TMS protocols and with transcranial direct current stimulation (tdcs) it is possible to change the overall properties of a particular area (Dayan, Censor, Buch, Sandrini, & Cohen, 2013).

Importantly, all those functional methods are embedded in a steadily growing anatomical understanding of the cytoarchitecture and connectivity patterns of the cortex and its boundaries as well as gradients. For more than a century, staining and tracing techniques have been refined and used on a variety of species to map their cortical and subcortical organization. Furthermore, such anatomical profiles can be used to find cross species correspondences, crucial for using animal models for understanding human brain functions. Due to the unavailability of tracer injections in humans, diffusion tensor imaging, functional coupling analysis and post mortem anatomical studies are used to find equivalents between human and animal neuroanatomy. In this thesis I will use anatomical labels and divisions to define different brain areas and their function. However, for simplicity, I will consider sub-regions that might differ in some respects but which are co-active in many functional imaging studies.

## Model-based approaches to data analysis

Without sufficiently sophisticated analysis approaches complex data can be very difficult to interpret. Fortunately we can use powerful computational models to generate specific intermediate parameters such as internal task state estimates allowing the dissociation of interrelated mechanisms in complex data sets. Such models allow us to further test different theories about the mechanisms underlying a particular behaviour. They are therefore particularly valuable when trying to understand which brain areas represent different elements of a valuation/value estimation, evaluation/value comparison and contextual influences on such processes. The best-known example of a simple modelling approach is reinforcement learning. Using a simple prediction error updating rule it allows the dissociating between error, expectation and outcome signals. Additionally, individual differences between estimated learning rates can further reveal neural differences that underlie changed evidence integration.

Clearly when working with humans it is not possible to use the most precise methods for recording or manipulating neural activity. However, apart from the obvious value of validating theories about specific neural mechanisms in the human brain derived from animal studies, there are other advantages to the human as an experimental system. The behaviour can be more complex and flexible when it comes to multi-facetted decision problems. This is particularly important when looking at dynamic influences of context on difficult evaluations and more abstract deliberation, since humans can accomplish such tasks with comparatively little training.

## How to find the neural mechanisms of decision-making

To conclude, there is an ever-increasing repertoire of methods and analysis approaches allowing us to relate neural processes to cognitive and behavioural phenomena in the frontal lobes. We can observe the neural dynamics coinciding with a particular decision process, but also directly test the relevance of such encoding by actively changing the firing rates of specific neuronal populations. Functional separation of the distinct roles of parts of the frontal lobes using neural activity relies heavily on the appropriate use of computational models. Computational models have to be completely specified if they are to be fit to behavioural data, and thus implement a process through which particular information is processed e.g. a stimulus is valued, comparisons are implemented and resulting choices are generated. Therefore the decisions about the computational model are crucial, since they inform the interpretation of the function of all different parts of the system. Since there are often several ways to model the same behaviours using different mechanisms, such as modifying option values when they are assessed or effecting behaviour through changes in the decision process itself, neural measures can be incredibly useful for differentiating between possible neural mechanisms. Further, only by using experimental manipulations that generate the crucial differences between different candidate models, can we distinguish between mechanisms that in many other scenarios generate indistinguishable predictions. Good models generate hypothesis that can be tested in many different experimental modalities seamlessly (e.g. Alexander and Brown 2011, Nature Neuroscience).

The most specific models of decision making generate categorical predictions of the choices to be made, parameters to encode, but also the temporal dynamics assumed and neuron types engaged in such processes. However, before being able to generate biophysically plausible neural networks of the evaluative process, as has been done for the case of simple comparisons, and the vmPFC (Hunt, 2012), it is important to get a sense of the framework in which computations are generally made in an area. For this, using a method such as fMRI, which gives a very integrated measure of neuronal firing and processing (N K Logothetis et al., 2001; Nikos K Logothetis, 2008), is very useful. It allows us to generate and compare candidate models of the mechanisms actually underlying an area's function.

Causal manipulations are often lacking or inconclusive in investigations of value-guided learning and decision-making by the frontal lobes. This may be because traditional electrical microstimulation experiments exploit the topographic organization of a brain area in order to change preferences. Whereas, perceptual choice paradigms such as random dot motion tasks that require motion congruent saccades can often be related in straightforward ways to the spatial organization known to exist in parietal and extrastriate cortical areas such as LIP and MT, we know very little about potential spatial organization within medial prefrontal areas. Even more neuron-type specific stimulation methods can be inconclusive, if we do not know which population to target and what preference it should affect. On the other hand neuronal interference or activation is particularly important when it is necessary to determine if neural activity is causally important for a decision rather than simply decision adjacent or post decision-related.

More fundamentally, it is not always clear what types of comparison processes are implemented in different prefrontal brain areas, since there are many reasons an area might be sensitive to reward information or to features of the current choice during the decision period. Equally, comparisons could be implemented in simple choice problems in many ways, often generating equivalent behaviours. What is needed, are dissociable predictions of neural activity for different evaluative mechanisms and subsequently a coherent framework describing and predicting the major functional properties of a particular network. The neuroethological approach (Adams, Watson, Pearson, & Platt, 2012) adopted in this thesis is aimed at relating more natural decision processes as well as ecologically relevant types of evaluation problems to the function of distinct brain areas. The optimal foraging literature is both useful for informing my understanding of foraging behaviours and decisions on both an algorithmic and functional level. On the algorithmic level, foraging theory has useful models of animal behaviour with the models attempting to optimize on overall fitness. This also gives a clear rationale for the evolution of such cortical mechanisms, since foraging, as a strongly ecologically important type of behaviour, ought to have neural mechanisms dedicated to solving such problems optimally. However, how such algorithms are actually implemented by neural mechanisms is largely unknown. Therefore, in this thesis I attempt to both test the utility of this ecologically inspired approach to studying behaviours and brain function and I attempt to shed light on the precise details of neural mechanisms for value comparison.

Before beginning with the actual experimental chapters, I will discuss the current understanding of functions of parts of the frontal lobes, with emphasis on their roles that are related to decision-making and reward learning. I have attempted to divide the medial and orbital frontal cortex into a limited number of areas and in the following sections I compare some of what is known about their functions with each other and with other areas on the lateral surface of the frontal lobe and in the frontal pole.

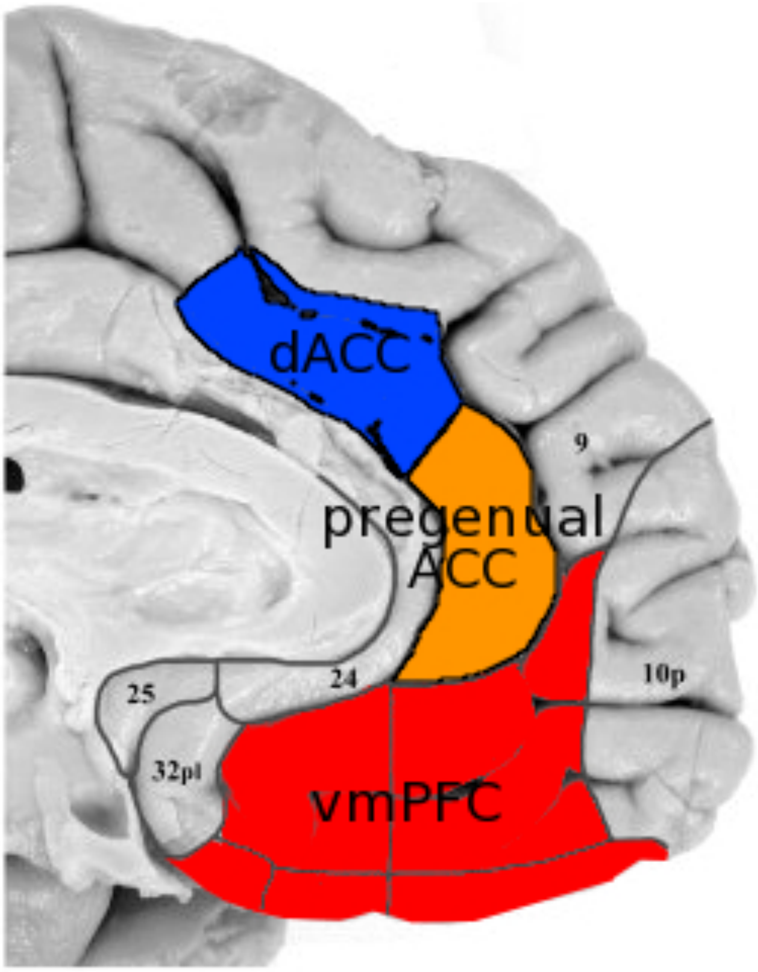

***Figure 7:*** *Three distinct regions of interest on the medial surface adapted from* (Ongür & Price, 2000). *Many different subdivisions of the medial prefrontal areas exist, but this rough division aligns quite well with functional imaging studies.*

## Anterior Cingulate Cortex

Despite few other brain regions spurring more interest, the exact role of anterior cingulate cortex has eluded researchers for many decades. One of the problems in identifying this medial prefrontal brain region's primary function lies in its engagement in a variety of heavily investigated cognitive concepts. From social cognition (P H Rudebeck, Buckley, Walton, & Rushworth, 2006), conflict monitoring (Matthew M Botvinick, Braver, Barch, Carter, & Cohen, 2001; Matthew M Botvinick, Cohen, & Carter, 2004), cognitive control (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), action planning (Picard & Strick, 1991), sequential behaviors (Procyk, Tanaka, & Joseph, 2000) and reward-based learning (Behrens et al., 2007; Holroyd & Coles, 2008) to decision making (T. A. Hare et al., 2011; M. F. Rushworth, Kolling, Sallet, & Mars, 2012), barely a higher cognitive function exists that has not in some way been associated with ACC activity. Coupled with the fact that ACC has connections to many associative parts of the neo-cortex, no single obvious unifying function has been proven, although some controversial unifying theories exist such as Conflict theory (Matthew M Botvinick et al., 2001, 2004) or Action-Outcome associator (Alexander & Brown, 2011).

In this thesis, I will focus on the dorsal ACC, sometimes referred to as the anterior mid cingulate cortex (aMCC) or rostral cingulate zone anterior part (RCZa) and the pregenual/perigenual anterior cingulate cortex anterior to it. I will however, also discuss studies involving the more ventral parts of the ACC both in the sulcus and gyrus (sACC and gACC).

## Conflict theory

One of the most well known observations about ACC activity in humans is during tasks involving conflict (Matthew M Botvinick et al., 2001, 2004). Most commonly studies try to induce conflict by using one of three different experimental paradigms. Firstly, the Simon task has been used to induce conflict between the side of space in which an action will be directed and the side of space in which a stimulus is shown. Secondly, Flanker tasks induce conflict between multiple directional cues only some of which are relevant for performance. Often such tasks are also analyzed in terms of sequential dependencies between trials such as the Gratton effect, which is the reduced incongruency-related slowing when one incongruent trial precedes another. Thirdly, simple variations of the Stroop task, which induces a conflict between the colour and the meaning of words, have been used. All those tasks have consistently been shown to invoke ACC activity in situations of higher conflict.

However, subsequently many studies were not able to confirm more specific predictions of conflict theory during decisions. First of all, no “conflict” neurons, i.e. single neurons that represent the amount of conflict in a particular task have convincingly been shown in electrophysiological recording studies. In a saccade counter-manding task, no neurons in the ACC representing conflict were observed (Ito, Stuphorn, Brown, & Schall, 2003). Instead many were responsive to error and reward omission. Sheth et. al., (Sheth et al., 2012) did suggest conflict related activity in their neural recordings in humans. However, they only observed increased activity in situations of possible alternative actions compared to fewer alternatives, which is consistent with any of the other accounts of dACC function. Furthermore, Fellows and Farah (Fellows & Farah, 2005) observed to no change in conflict related behavioural change due to ACC lesions, nor impaired post-error slowing. Patients were also able to change their speed-accuracy trade-off when it was cued explicitly. This means lesions of the ACC do not produce simple problems during conflicting situations. Indeed monkeys with lesions in ACC have no obvious conflicted-related impairments (M F S Rushworth, Hadland, Gaffan, & Passingham, 2003) most, human patients appear to have deficits in some behavioural adjustments and using sequential relationships such as measured by the Gratton effect (di Pellegrino, Ciaramelli, & Làdavas, 2007; Sheth et al., 2012). Although, broadly speaking consistent with a role of the ACC in conflict monitoring and subsequently behavioural updating, the findings are also consistent with many other models of ACC function. Furthermore, it is puzzling that no more direct conflict related effects can be observed, if conflict monitoring is really its principle function. Furthermore, many imaging studies found their ACC activity could be better explained by other constructs (Alexander & Brown, 2011; Basten, Biele, Heekeren, & Fiebach, 2010; Boorman et al., 2009; Mark E Walton, Devlin, & Rushworth, 2004; Wunderlich et al., 2009).

More recent modifications of the conflict theory, have suggested it is responsible for evaluating the expected value of cognitive control (Shenhav, Botvinick, & Cohen, 2013) rather than just fulfilling a function in conflict monitoring or adjustment itself. However, by introducing the concept of expected value comparison into theories of conflict, arguments about the ACC's role are becoming increasingly convoluted and un-parsimonious, since all activity can now be explained by reflecting conflict, i.e. the proximity of two competing values with each other, or the expected value of resolving this conflict, which could be anything from the total value of all options, the best option's value or even the opposite contrast to conflict, i.e. of the distance between all values, further weakening any possibility of making falsifiable predictions using this model. Other re-formulations of conflict theory as difficulty suffer from the same problem of difficulty being very ill defined. Decisions with very similar values are at the same time very difficult, as finding the correct option requires better discrimination and very easy, as the gains for picking the better option are marginal. This is of course different from perceptual paradigms in which only the correct response is rewarded and difficulty is less ambiguous.

## Behavioural flexibility, switching and learning

We now know that ACC is frequently active when behavioural flexibility or change is necessary, often through switching away from previously executed actions (K Shima & Tanji, 1998) or task rules (Crone, Wendelken, Donohue, & Bunge, 2006; Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Liston, Matalon, Hare, Davidson, &

Casey, 2006), particularly if such actions are voluntary or self-generated (Lau, Rogers, Ramnani, & Passingham, 2004; Mark E Walton et al., 2004) and often reward-guided (K Shima & Tanji, 1998). Although behavioural change is frequently related to some degree of conflict and error monitoring, they are by no means interchangeable concepts. We know that outcomes that inform a participant about the need to change future responses, but without them making an actual error, do indeed activate the ACC (Mark E Walton et al., 2004). However, it is important to note that some studies on task switching (M F S Rushworth, Hadland, Paus, & Sipila, 2002) and response competition (M Ullsperger & von Cramon, 2001) see activity more centred around the SMA and pre-SMA for response competition. One distinguishing feature that could lead to activity in the dACC could be an increased reward relevance and sequential dependencies in a task such as action learning, whereas the pre-SMA/SMA might be more concerned with the selection of concrete “action sets” (M F S Rushworth, Walton, Kennerley, & Bannerman, 2004).

The ACC has also been identified as being sensitive to Surprise for rewards in single cell recordings in macaques (Hayden, Heilbronner, Pearson, & Platt, 2011), and violation of error expectation i.e. accuracy surprise rather than errors themselves (Brown & Braver, 2005; Jessup, Busemeyer, & Brown, 2010), although it could be more specifically updating of future expectations (O’Reilly et al., 2013) and learning. Mechanistically, it might serve as modulator of learning according to the current environment (Behrens et al., 2007) rather than actively implementing simple reinforcement learning itself.

Lesioning ACC can impair monkeys’ ability to use past reward experience to maintain reward-guided behaviours (S W Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006); Although unimpaired in using the last trial’s outcome for making choices, more remote outcomes had significantly less impact on monkeys’ behaviour. In other words, they were unable to use the longer term reward history to adapt their behaviours so as to obtain rewards more optimally across several trials. dACC has neurons encoding reward with different time constants, which would allow for longer term and dynamic use of reward information (Bernacchia, Seo, Lee, & Wang, 2011). A neural network with such neurons could easily use those different time constants for reward employing simple rules of synaptic weights and mutual inhibition to compute past reward trends and future reward projections. Thus, it is biophysically plausible that dACC could be used to make estimates about how a reward environment is going to develop in the future and whether it is going to be more valuable to remain in such an environment or not. It has been suggested that ACC is important for average reward rate computation (C Amiez, Joseph, & Procyk, 2006) and also other environmental information such as uncertainty or volatility (Behrens et al., 2007).

Confirming the relevance of ACC activity for reward guided behavioural changes as suggested from learning and simpler switching studies, an impressive amount of research has looked at the role of ACC in performance monitoring and behavioural adjustment (reviewed in (Ridderinkhof et al., 2004; Markus Ullsperger, Danielmeier, & Jocham, 2014)).

## Performance monitoring and behavioural adjustments

Arguably the most important piece of information when monitoring the performance of oneself or others is error detection. Neurophysiologically, the most extensively studied neural correlate of error processing are error related negativity (ERN) (W. Gehring, Coles, Meyer, & Donchin, 1990; W. J. Gehring, Goss, Coles, Meyer, & Donchin, 1993) and feedback related negativity (FRN) (Miltner, Braun, & Coles, 1997) which can be measured using EEG and which have been associated with dACC. A similar error related potential has been established in electrophysiological recordings made in macaques in ACC (area 24) (Gemba, Sasaki, & Brooks, 1986). Error and reward related activity in macaque dACC have been dissociated from conflict-related activity (Ito et al., 2003) using a saccade countermanding task.

ERN's can be measured when observing one's own mistakes but also when others make an error (van Schie, Mars, Coles, & Bekkering, 2004), but they appear to be conditional on the requirement to adjust one's own future behaviour (de Bruijn, Schubotz, & Ullsperger, 2007). FRN's appear most extensively when volitional motor actions have to be executed subsequently (Walsh & Anderson, 2012). They might in part reflect quantitative reward prediction errors in the ACC that are signed (Behrens, Hunt, Woolrich, & Rushworth, 2008; Fischer & Ullsperger, 2013; Rutledge, Dean, Caplin, & Glimcher, 2010; Talmi, Fuentemilla, Litvak, Duzel, & Dolan, 2012) and unsigned (Hayden, Heilbronner, et al., 2011; Klavir, Genud-Gabai, & Paz, 2013). Apart from errors, novelty appears to also invoke ERN like activity in EEG (Wessel, Danielmeier, Morton, & Ullsperger, 2012), suggesting the key to both ERN and FRN might lie in behavioural consequences that might require adjustments in the way responses are made either by changing future choices directly or response parameters such as speed-accuracy trade-off. Correspondingly, not just errors per se but error relevance for future decisions modulate ACC activity (Markus Ullsperger & von Cramon, 2004). Interestingly, the neurotransmitter dopamine has also been implicated in signalling reward prediction errors (Schultz, Dayan, & Montague, 1997), novelty (Dulawa, Grandy, Low, Paulus, & Geyer, 1999; Wittmann, Daw, Seymour, & Dolan, 2008) and the need or ability for behavioural change (Berridge & Robinson, 1998), suggesting a possible link between the functions of ACC and striatal and prefrontal dopamine (Holroyd & Coles, 2002).

The concept of performance monitoring especially for the purpose of behavioural adjustments as a core function of ACC, has received much support (Markus Ullsperger et al., 2014) and allows the integration of many of the above mentioned findings on ACC activity during situations that require behavioural adaptation, induce response conflict or lead to switching as well as reward learning (Holroyd & Coles, 2002). It goes beyond a simple error detection function, as behavioural adjustments occur even without the occurrence errors (Mark E Walton et al., 2004) and increases in accuracy after lowered reward expectation are both also associated with ACC activity in monkeys (K Shima & Tanji, 1998) and humans (Williams, Bush, Rauch, Cosgrove, & Eskandar, 2004). In fact, the apparent error sensitivity of the ACC might be in large part be due to the usefulness of errors for behavioural adaptations, i.e. the frequent need to effect behavioural change after mistakes have been made and the high information content of mistakes in most instances (Mark E Walton et al., 2004). Indeed, if errors appear more to be more frequently associated with dACC than do correct responses, it should be noted that dACC signals have also been shown to increase with rare correct responses (Jessup et al., 2010) and the feedback obtained after rare correct responses are made is arguably particularly informative. Importantly, ideas of monitoring and behavioural adaptation can be reconciled with evidence for value related representations in ACC as well as electrophysiological and lesion evidence better than can accounts of ACC that simply emphasize conflict. Such accounts are reconciled by suggesting that the ACC is not just tracking value but that it is also monitoring the need for behavioural change on the basis of value estimates.

Behavioural adjustments can be broadly categorized into at least three different timescales. Firstly, adjustments can be made within the event in which a need for adjustment is detected. This process is sometimes called reactive control (Braver, 2012; Hikosaka & Isoda, 2010; Markus Ullsperger & King, 2010). In fact, when participants correct an error by making the right response immediately afterward, larger ERN's (W. J. Gehring et al., 1993) as well as increased ACC BOLD signals (Fiehler, Ullsperger, & von Cramon, 2004) can be observed. Secondly, error information as well as feedback processing in general can be used to inform next trial behaviour, directly or indirectly. This is often referred to as proactive control (Braver, 2012; Hikosaka & Isoda, 2010; Markus Ullsperger & King, 2010) and is most prominently discussed in the form of post error slowing and accuracy increases as well as sequential congruency effects such as the Gratton effect mentioned earlier. Indeed the ACC has been associated with both post error slowing (Chevrier & Schachar, 2010) and post error improvements (M. X. Cohen & van Gaal, 2013) as well as the Gratton effect itself (Kerns et al., 2004). Thirdly, behavioural change can also happen on a longer time scale through learning based on reward history that leads to improvements in the value of behaviour over the longer term (Behrens et al., 2007; Steven W Kennerley, Walton, Behrens, Buckley, & Rushworth, 2006). This suggests that behavioural adjustment on many different timescales all involve the ACC and that this often, but not exclusively, is the consequence of error monitoring.

Performance monitoring and foraging functions might at first seem to be quite distinct processes. The former is about the optimization and adjustment of currently attended action plans, whereas the latter is about value-guided changes of such plans in favour of other alternatives. However, both approaches assume that factors such as uncertainty, surprise, novelty and reward information should all be tracked to serve ACC function. Furthermore, other environmental and meta-representations such as volatility should also be considered to enable better foraging or monitoring. Also, foraging evaluations could be seen as a particular form of performance monitoring, in which a current engagement is compared against the potential alternative values and long term outcomes in the environment and behavioural adjustments are initiated when superior alternatives are detected. In other words, monitoring then would be about the validity of what the agent is doing, rather than improving upon its accuracy. Events such as errors could in fact induce both processes in parallel with the current context and other control processes determining which is ultimately executed.

On the other hand performance monitoring and the resulting behavioural adjustments could also be seen as a special case of foraging behaviours, since they are a change of the current response strategy and more volitional response selection. Thus, it is possible that performance monitoring and adjustments are based on a neural mechanism which initially was concerned with evaluating current and alternative response strategies taking into account the current environment and potential longer term values i.e. a form of foraging comparison mechanism. In fact, Gehring (W. J. Gehring & Willoughby, 2002) found that the early ERN can reflect the outcome, i.e. whether participants gained or lost money, rather than the correctness of choices. However, note that in this simple task there was nothing to learn about either available option and only the chosen option's outcome determined overall task performance. Thus, rather than the unchosen option value, the last trial's outcome modulated ACC activity as a function of gains on previous trials. This suggests an environmental contextualization of reward processing or a comparison of current against past or longer term rewards. Such a proposed foraging mechanism could in other tasks further adjust and optimize upon current task performance by effectively considering environmental reward trends. Behaviourally, this would often mean "switching" or changing into a different, often more deliberate response strategy and monitoring of performance and reward outcomes more carefully, to evaluate potential behavioural changes that need to be implemented to enhance performance and achieve longer term gains. Consistent with such a proposal, ERN like signals can be observed during outcomes with explicit signals for behavioural adjustment or exploration (Sallet, Camille, & Procyk, 2013).

Thus far I have focused on immediate ERN signals and the aspect of foraging, which is concerned with the initial decision and initiation of foraging choices and suggested a potential role of ACC in this process. However, this does not preclude the ACC from being involved in another component of successful foraging behaviours related to performance monitoring and adjustments, i.e. the maintenance or execution of deliberate search behaviours after initiation. Generally, longer term foraging goals are preceded by the successful execution of a sequence of novel or learned behaviours and require precise monitoring of one's own performance and possible outcomes. In fact, many foraging context are marked by a need to persevere through a relative lack of immediate rewards to achieve a longer term gain of increased reward rates, as do post error effects. Additionally, both often require a sequence of deliberate and novel or improved sequences of behaviours.

As further support for such an interpretation, ACC has been repeatedly identified as being crucial for the implementation of such sequential and persistent actions especially in the context of costs that need to be overcome. However, although performance monitoring and adjustment theories can explain effort related and sequential action effects in the ACC, they are not seen as a direct consequence of its function in monitoring or adjustments. On the other hand, such effects do fit very nicely to the two necessary stages of successful foraging behaviours, i.e. its initiation and subsequent implementation and maintenance.

In the following section I will discuss the relationship between ACC and action invigoration and how it fits in the foraging perspective further.

## Actions, Effort, Sequences and Persistence: For the greater good!

The anterior cingulate cortex is one of the cortical areas concerned with reward and value that is most closely related to the motor system. Stimulation in macaques can lead to movements (Luppino, Matelli, Camarda, Gallese, & Rizzolatti, 1991) and impairments after lesion can be modality specific (Turken & Swick, 1999). Furthermore, parts of the ACC have extensive connexions with motor areas (Picard & Strick, 1991).

Matsumoto and colleagues (K. Matsumoto, Suzuki, & Tanaka, 2003) showed that neurons in the ACC encoded reward expectation in conjunction with whether the reward was paired with a Go or No-Go response. Some ACC neurons were only reward sensitive in Go or No-Go trials. This was in contrast to dlPFC, in and around the posterior half of the sulcus principalis, which during the same time coded the conjunction of visual stimulus and reward expectation, visual stimulus-motor association and reward expectation but not the interaction of reward and action itself. When macaques learned action reward associations ACC neurons also encoded the prediction errors for such associations, suggesting neurons there can link reward value to specific actions (M. Matsumoto, Matsumoto, Abe, & Tanaka, 2007). In humans, the possibility of consistent reward-type action mapping increases ACC activity (M P Noonan, Mars, & Rushworth, 2011). Such mapping helped participants when first acquiring stimulus response associations.

In a parallel line of research, the ACC has been investigated in relation to its role in sequential behaviours. Procyk (Procyk et al., 2000) Showed that ACC neurons are sensitive to whether a behaviour is indeed sequential or non-sequential. There is an increased response in the ACC when going through a multi-step sequence of otherwise equivalent responses as monkeys proceed through this sequence toward the reward (Shidara & Richmond, 2002).

In summary, there is evidence that ACC is active during sequential behaviours such that it is increasingly active when an animal persists through a sequence of voluntary actions in order to acquire a distant reward (Shidara and Richmond, 2002) and also that it encodes conjunctions of Go/No-Go responses and reward expectations (K. Matsumoto et al., 2003). However, others have failed to find an interaction between Reward expectation and Saccade direction, but rather an independent coding of both (Hayden & Platt, 2010). Cai and Padoa-Schioppa (Cai & Padoa-Schioppa, 2012) reported a limited number of cells encoding action-reward conjunctions as evidence against action-based selection in ACC.

Most relevant for the issues of decision making however, Kennerley and colleagues (Steven W Kennerley et al., 2011) showed that ACC had greater access to more decision relevant parameters in simple binary cost-benefit trade-off decisions than did other brain areas. They trained macaques to select between two picture stimuli that either varied with each other in terms of their effort, probability of reward or reward magnitude. The other two dimensions of utility were present in every trial but they fixed while the third dimension was varied. The difference between the options' values within the relevant domain (effort, reward magnitude, reward probability) was always the same (thereby controlling for difficulty), but the overall value of the trial was varied (e.g. monkeys might choose between effort levels of 1 and 3 presses or between 5 and 7 presses). ACC, dlPFC and OFC neurons were recorded throughout the task. Only ACC neurons appeared to "multiplex" the relevant reward information, i.e. the same neurons encoded all three utility dimensions when they were relevant to choice (costs were coded with a  opposite change in activity to that seen when benefits were manipulated). Some of the ACC neurons that represented utility of a choice positively also represented the prediction error positively. When the focus was just on how neurons encoded one aspect of value, the probability that a choice would lead to reward, then even more neurons encoded reward probability in the same manner as they encoded prediction errors at the time that feedback on the choice was given. Thus, ACC neurons encode all components necessary for simple cost-benefit decisions in an integrated fashion and could easily drive both cost-benefit comparisons as well as support learning about future value using the information from feedback through prediction error learning.

In humans, dACC neurons are particularly sensitive to reductions in rewards particularly when this coincides with a need to change behaviour (Williams et al., 2004). A similar observation was made before in macaque neuron recording studies (K Shima & Tanji, 1998). This corresponds well with most theories of updating and behavioural flexibility and the ACC, however in humans, increased activity of dACC neurons in such a context predicts higher accuracy, suggesting this activity is not merely a reflection of lowered reward expectations and switching. Furthermore, when dACC was ablated, patients performed particularly poorly in those lower reward trials. This rather suggests a function in overcoming situations of lowered rewards without a drop in motivation or accuracy. Furthermore, stimulation of neurons in the ACC of humans have been reported to induce a sense of a “will to persevere” in patients (Parvizi, Rangarajan, Shirer, Desai, & Greicius, 2013). They reported an increased motivational drive and anticipated new challenges, which could be interpreted as an exploratory drive or a motivation to maintain a new action sequence over time.

To conclude, two seemingly contradicting literatures have emerged, linking the ACC both to learning, contextualization by the environment and changes in behaviour as well as to staying on course, following a sequences of actions or choices or investing continuous effort into an activity. However, it might be possible to reconcile those findings. Conceptually, both value guided behavioural changes and effortful persevering sequential behaviours can be motivated through the expectation of potentially higher longer term rewards. In other words, it is possible that ACC holds information about the current reward environment to allow for contextually appropriate longer term value driven behaviours, which could mean changing the current course of action, exploring or sometimes going through one or a sequence of costly actions to optimize the average reward rate in the longer term. It could be that those deliberate and often costly actions that need to be maintained over time are what ACC implements, regardless of whether they imply change of behaviours because of a changing reward environment or maintaining responses toward the same goal over an extended period.

## Value driven exploration and novel/voluntary/effortful actions

So far I have described a relationship between the ACC and learning as well as behavioural flexibility, switching and performance monitoring. Although behavioural changes can be the consequence of changes in reward contingencies, they can also precede them. An agent makes exploratory responses when choosing to engage with novel voluntary behaviours rather than exploiting known past reward contingencies. Daw and colleagues (Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006) made participants chose between four different drifting one-armed bandits (probabilistic gambles with different probabilities). After participants selected a bandit, they would only receive feedback on their chosen option and gradually with every stay decision grow more uncertain about the value of the alternative bandits. As soon as they decided to switch however, they were essentially exploring the other bandits' reward probabilities. Such exploratory decisions engaged the frontal pole, intraparietal sulcus (IPS), as well as the ACC.

In a more open exploration paradigm Yoshida and Ishii (Yoshida & Ishii, 2006) showed strong ACC activity for goal directed exploration in a chess-like grid environment. It further was most active when participants were likely to back track, i.e. had to act against simpler reward approach mechanisms, to receive their desired goals. In rats spatial decision lead to interactions between hippocampus and ACC in the theta band (Remondes & Wilson, 2013). Such an interaction could support an ACC role also in spatial decisions and foraging behaviour.

In another paradigm designed to investigate exploratory behaviour in macaques (Quilodran, Rothé, & Procyk, 2008) a distinctive pattern of exploratory-related activity could be observed. In the task the animals had to explore four stimuli in order to find the rewarded one. Once they found the rewarded option they could exploit this knowledge four times by repeating their choice. Subsequently, the rewarded item was changed and they had to enter another exploratory period. Subsequently, Amiez (Céline Amiez, Sallet, Procyk, & Petrides, 2012) used fMRI to report similar differences in activity on exploratory versus exploitative trials in human dACC.

An additional key component for foraging style evaluation, average reward rate, seems to be computed in the ACC in macaques. Macaques were tested on average reward value encoding by choosing between two options with different averages through different probabilities of the same large and small rewards. ACC appeared to be crucial for finding the better stimulus, i.e. the one with the higher average reward (C Amiez et al., 2006), since muscimol deactivation impaired this ability.

One study directly testing whether foraging-type evaluation is conducted in ACC was reported by Hayden and colleagues (Hayden, Pearson, & Platt, 2011). They recorded from single neurons in dorsal ACC in macaques’ while the macaques were doing a variant of a classic foraging task mentioned earlier. They trained them to recognize that bars of different height on a computer screen indicated the “handling time” that taking such a choice would take. Furthermore, the animals received continually updated information about the handling time because the bar heights decreased throughout the period that the animal engaged with a choice (engagement with the choice was achieved by the animal fixating it as drifted down the length of the computer monitor on which it was displayed. Animals had to choose between fixating either a blue or grey bar. If they fixated the blue bar for a certain time then they received a reward. The size of the reward decreased with each successive selection of the blue target. If, however, they decided to fixate on the grey target the animals effectively moved to a new foraging patch and the level of reward available for taking further blue target choices was renewed to its original level. As further were made in this new patch, however, the reward level began to decrease again. This mimics a traditional patch-leaving problem in which animals have to decide between continuing to harvest a depleting patch, (the blue bar choice in the current experiment), or leave the patch for an un-harvested alternative through the intermediate step of the grey bar which itself did not lead to reward. The height of the grey bar was varied to represent different travelling times before the next patch would be reached (as noted above the travel time between patches as well as the reward rate experienced in a patch should affect the animal's choice of a leaving time. By posing this foraging decision as a series of discrete events they can nicely investigate neural activity triggered by such decisions and find activity predictive of categorical decisions.

Another pressing issue that remains to be resolved is that many electrophysiological studies do not consider the ACC to have an active role in reward-guided evaluation of choices at least in their tasks (Blanchard & Hayden, 2014; Cai & Padoa-Schioppa, 2012). Although it is undisputable that some reward guided decisions can still be made without ACC and that its importance depends on the particular decision problem at hand, there are reasons for thinking that it might contribute to reward-guided decision making. Particularly, when taking a foraging perspective, the ACC appears to represent a lot of the information necessary for making and implementing foraging style decisions. It knows about time, is sensitive to the sequential nature of tasks, represents the costs of actions and is particularly important for learning about and selecting specific, novel, and voluntary actions and for the motivation to maintain such responses. Furthermore, it is responsive to a wealth of reward information, it can also use fictive outcomes for changing actions (Hayden, Pearson, & Platt, 2009) and environmental changes. It has all the crucial information to compute reward trajectories, signals when an environment is changing more rapidly or when there is a need for learning new contingencies. It represents all costs and benefits (Steven W Kennerley et al., 2011) in a motivational frame.

In summary, ACC carries a lot of the information necessary for foraging style evaluations and subsequent behaviours. 1) It is active during spatial and non-spatial exploratory decisions as well as during feedback processing. Both forms of exploration are crucial for adaptive patch foraging. 2) It appears to have many roles in contextually sensitive reward learning such as volatility modulation and computes averages reward rates, a crucial parameter in foraging theory. It could compute environmental trajectories to enable prospective foraging decisions. 3) It strongly cares about sequential behaviours and persevering through them; This is particularly important for foraging which relies on a series of deliberate responses to acquire new, better longer term goals, rather than potentially regressing to inferior concrete offers. 4) It knows about costs and benefits of voluntary effortful actions and can learn about specific action values. Generating such new sets of behaviours is needed while trying to acquire novel rewards, shaping new strategies toward a longer term reward through combining different actions with each other. 5) For this it uses performance feedback to modify behaviours for future situations.

Thus, ACC is not an area that is only concerned with monitoring conflict, nor trying simple switches. It represents a wealth of reward and cost information about the environment and potential sequential behaviours, particularly novel actions, in order to motivate choices with longer term value. By contrast giving subjects the sort of decisions that are typically studied by neuroeconomists might not be the best way to study ACC activity.

## Orbitofrontal Cortex

The earliest lesion research found changes in patients' behaviours, including inappropriate social behaviour, and a lack of inhibition of behaviours inappropriate to a particular context (Damásio, 1994). More specific theories of OFC function have since emerged as a result of using more rigorous psychological tests and more specific lesion populations. Bechara and Colleagues (Bechara, Damasio, Damasio, & Lee, 1999) showed that in simple reward learning tasks, patients without the OFC had profound deficits in learning and repeatedly selecting the more profitable of two options in a set of alternatives. In their most famous paradigm, the Iowa gambling task, participants had to learn to select among four decks of cards, each with distinct outcome distributions. The subjects' aim was to select the two "good", i.e. more profitable decks and avoid the two "bad" ones using outcome feedback. Patients with OFC lesions kept picking the "bad" decks, i.e. they could not utilize the outcome feedback appropriate for selecting the options associated with better outcomes. This impairment coincided with a lack of galvanic skin responses related to reward expectation, and an abolishment of a differentiation of deck value through this measurement prior to behavioural preference. This pattern of behaviour contrasts with that seen after amygdala lesions, which, unlike OFC lesions, also interfered with galvanic skin response at the time of outcome delivery. The patient studies led to the origin of the somatic marker hypothesis (Bechara, Damasio, Damasio, & Anderson, n.d.; Bechara, Damasio, Tranel, & Damasio, 1997, 2005), which states that the simulation of somatic states or emotional responses are exploited to allow adaptive decisions to be made and that this process is in part driven by the OFC.

Many of the earlier lesion studies neglected to distinguish clearly between the ventromedial part of prefrontal cortex, (here vmPFC) from the adjacent but more lateral OFC. Lesions of medial or lateral OFC cause dissociable impairments (M P Noonan et al., 2010; Mark E Walton et al., 2010) and both systems have different activity profiles (Boorman et al., 2009; M P Noonan et al., 2011) as well as unique anatomical connectivity (Carmichael & Price, 1995, 1996; Ongür & Price, 2000). Thus, I will in the following section distinguish between studies that looked at either the central and or lateral OFC as opposed to the medial OFC or vmPFC.

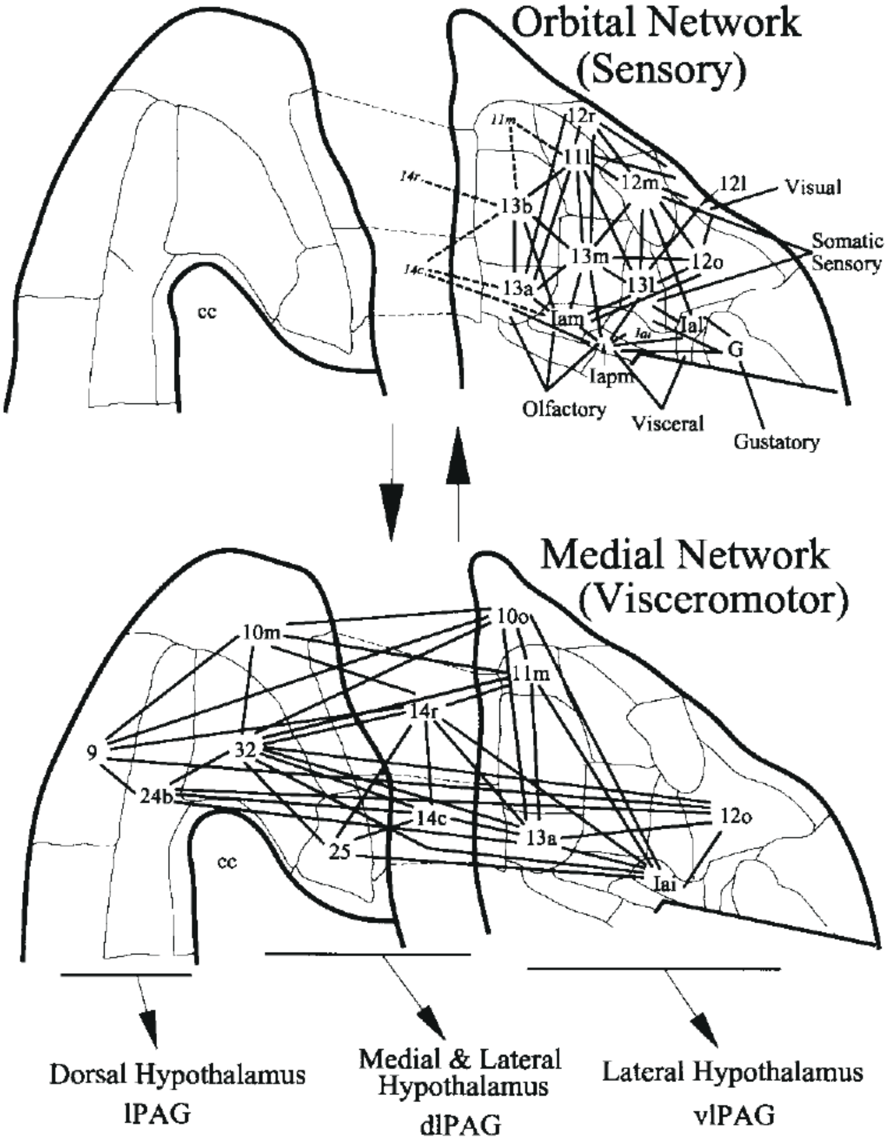

***Figure 8 :*** *An illustration of the strongest close range connections within the orbital and medial frontal cortex (directly taken from (taken from* (Carmichael & Price, 1995, 1996; Ongür & Price, 2000)*). The area labelled orbital network by Carmichael and Price corresponds to the region that is referred to as the central and lateral OFC in this chapter while the area labelled medial network or visceromotor network by Carmichael and Price corresponds to the region that is referred to as medial OFC and vmPFC in this chapter. There is a clear separation into two networks. There are dense connections within each network but spare connections between networks. The more lateral network strongly connects areas 13, 12 and 11 with each other and with sensory and visceral cortical regions and perirhinal cortex. The more medial network has stronger connections to*

*different parts of the hypothalamus and medial temporal areas such as the amygdala and hippocampus and with sme parts of the striatum*).

## Reward, Learning and Choices

In parallel to human patient studies, Murray and colleagues investigated the role of OFC in value contextualization and modification. Instead of looking at emotion or simple gambling learning, they investigated how the desirability of food items was modified by recent experience and internal states, i.e. satiety. Using selective satiety protocols and excitotoxic lesions, they observed monkeys were unable to update the values assigned to abstract stimuli associated with specific food items, following OFC lesions. In other words, in contrast to control animals, they could not synthesize an online value of the stimuli when making decisions, taking into account that their internal state had changed, making particular foods less desirable. Instead the stimuli themselves were perceived as desirable and selected repeatedly.

Subsequently, Rudebeck and Murray showed more specifically that lesions of the more lateral parts of OFC, areas 11 /13, but not the medial area 14, which is part of the vmPFC, led to impairments in appropriate food preference after selective satiation (Peter H Rudebeck & Murray, 2011). On the other hand lesions of area 14/vmPFC led to failures of ignoring the previously rewarded option in favour of the now rewarded alternative, in a simple reversal task. They also showed macaques to be relatively unimpaired in emotional regulation in contrast to more complex value updating after OFC lesions (Peter H Rudebeck, Saunders, et al., 2013), speaking against the strongest somatic state theories, and in favour of more complex value integration and updating.

Indeed there are many experiments relating the OFC with reward value updating in animals such as monkeys (M P Noonan et al., 2010; Mark E Walton et al., 2010), rats (Geoffrey Schoenbaum & Roesch, 2005; Takahashi et al., 2009) as well as learning in humans (Gläscher, Hampton, & O'Doherty, 2009; M P Noonan et al., 2011; O'Doherty, Critchley, Deichmann, & Dolan, 2003). VmPFC lesions more specifically appear to induce problems in the consistent preference or selection of the best option in learning paradigms (Fellows & Farah, 2007; M P Noonan et al., 2010; Mark E Walton et al., 2010), changes in other measures of choice consistency (Camille et al., 2011), and in other measures of accuracy (Henri-Bhargava, Simioni, & Fellows, 2012). Compared to the Iowa gambling tasks even simple reversal learning paradigms have the advantage that they allow tests of continuous updating and reversed preferences under competing option values. Furthermore, more specific computational models inspired by economic approaches to choice allow the dissociation of many different dimensions of value, such as probabilities or reward magnitudes. Such models also offer very useful tools for the integration of such information for the generation of single utility values in a universal currency, which can be directly compared to allow for adaptive decisions. VmPFC lesioned patients are particularly impaired when decisions require the integration of two option values across several dimensions (Fellows, 2006). This is coupled with a changed attentional strategy of valuing each individual option in isolation, whereas controls followed a strategy of comparing individual attributes with each other independently. We also know that the vmPFC is more engaged in decisions that require integration between the different value dimensions of an offer (Hunt et al., 2012), compared to "no-brainer trials" in which all dimensions are better for one option compared to another. Thus, vmPFC might be particularly important for multi-attribute comparisons between different dimensions of value.

As mentioned earlier, neuroeconomic approaches have shown that stimulus values are represented in the whole OFC, with reward type specific representations most closely linked to central OFC (Padoa-Schioppa & Assad, 2006, 2008; Padoa-Schioppa, 2013) while brain activity correlated with total, chosen or value difference is more prominent in the vmPFC (Boorman et al., 2009; De Martino, Fleming, Garrett, & Dolan, 2013; FitzGerald, Seymour, & Dolan, 2009; Philiastides, Biele, & Heekeren, 2010; Wunderlich, Dayan, & Dolan, 2012). Together with Hunt and colleagues (Hunt et al., 2012), I have put forward a mechanistic explanation for why both total and value difference representations appear in the same brain area on the basis of, a biophysical model. Our model of value comparison was very much inspired by biophysical models of perceptual choices in LIP (Wang, 2002), except that input to the different pools of neurons was not motion coherence but option value. The different neural pools representing the available options competed indirectly through inhibitory interneuronal populations, receiving continuous input as a function of the value of the option they represented. This first led to an overall increase in neural activity in the network as a whole as a function of the combined value, and later the difference between value of the ultimately chosen and unchosen alternative. A prediction directly confirmed for the vmPFC using MEG. Subsequently, with Chau and colleagues (Chau, Kolling, Hunt, Walton, & Rushworth, 2014) I have shown that the very same model can also explain counterintuitive sub-optimalities in peoples' decisions when there are multiple alternatives and distracting choices. The model also explained changes in the size of the vmPFC value difference signal in the presence of distracters. Subsequently Strait and colleagues (Strait, Blanchard, & Hayden, 2014) recorded single vmPFC neurons during very similar binary decisions like the ones in Hunts' original study and found evidence for a comparison process relying on mutually inhibition.

Very few investigators have directly compared ACC and OFC neuronal firing during decisions. However, Rudebeck and colleagues recorded from ACC and central OFC during a sequential decision making paradigm in macaques (Peter H Rudebeck, Mitz, Chacko, & Murray, 2013). They considered ACC to not be directly decision relevant because more neurons encoded the value of two sequentially presented stimuli they refer to as the first stimulus (S1) and second stimulus (S2) independently in OFC. They assume a region involved in the decision making process needs to have a relatively large number of neurons coding for both offer values independently to allow comparison and choice. However, ACC encoded all necessary parameters for the making of a decision, movement direction, S1, S2 and the difference between S1 and S2. A sparse coding of all those parameters would be sufficient for decision making especially as the relative preference for S1 encoding in ACC suggests a strong contextualization and sequential dependency that might underlie the area's ability to make choices in a switch-stay decision framework. It would be interesting to test more rigorously whether all important decision parameters in such a relative framework between staying or switching from one option to the other are represented sufficiently to drive decisions. Furthermore, it remains to be seen whether neuronal variability in ACC neurons is predictive of choices in such a sequential paradigm. Either way, central OFC did represent both offer values S1 and S2 independently and it did so prior to the ACC. This is consistent with many theories of central/lateral OFC function, since it does not in itself show this part of the OFC is involved in making decisions, just that it knows about option values relatively early. VmPFC recordings might be more choice predictive, although some have found sequential choices to engage the vmPFC less (Hunt, Woolrich, Rushworth, & Behrens, 2013). It is possible that in sequential paradigms the ACC is more choice relevant. Interestingly, behaviour in the task was unaffected by amygdala lesions and so were the ACC value signals (Peter H Rudebeck, Mitz, et al., 2013). Central OFC value signals, although still present, were reduced and had an increased latency, even though there were no changes in the animals' accuracies. This means either that the task was too easy for accuracy effects on decreased value coding in the OFC or it was not directly choice relevant.

However, we know from other studies that the amygdala has a particular link to the lateral OFC when it comes to learning and value updating (G Schoenbaum, Chiba, & Gallagher, 1998), rather than in decision making per se. So far I have presented a lot of evidence in favour of an OFC system being important for making decisions, particularly by creating online goal-directed value expectations (e.g. specific satiety), enabling the integration of multidimensional aspects of value into a single utility and supporting better learning, especially during difficult reversals. I will now try to address the question of what the precise role of the OFC system, particularly the lateral and central OFC might be in reward learning, but also what specific functions the more medial OFC, i.e. vmPFC has in value comparison i.e. decision making.

## Concrete comparison, credit assignment and current preferences

The OFC is not necessarily involved in all forms of decision making or preference. There are many cases in which vmPFC more specifically might be less important for making appropriate choices (Camille et al., 2011; N. Kolling, Behrens, Mars, & Rushworth, 2012; Nils Kolling, Wittmann, & Rushworth, 2014; Peter H Rudebeck et al., 2008). However, many human imaging experiments have shown activity in the vmPFC that is related to the decision making process and individual variability in preferences (Plassmann, O'Doherty, & Rangel, 2007) (similarly in macaques (Padoa-Schioppa, 2013)) and even confidence in one's own choices ((De Martino et al., 2013) DeMartino 2013 Nature Neuroscience, also shown in rats (Kepecs, Uchida, Zariwala, & Mainen, 2008) Kepecs 2008 Nature). As mentioned above, although lesions of the vmPFC in animals and humans does not completely destroy an agents' ability to make decisions, they do become worse at selection, particularly when a concrete options has multiple dimensions of value associated with it.  Furthermore, vmPFC might be particularly engaged if current preferences are reflected in behaviour, since vmPFC activation was observed in free bid trials related to the "Willingness to pay" (Plassmann et al., 2007), but not to passive viewing, although others can see some signals even under Pavlovian conditions (Lebreton, Jorge, Michel, Thirion, & Pessiglione, 2009). Subsequently Rangel and colleagues have shown in a series of studies that the vmPFC is particularly active as a function of the relative desirability of concrete rewards such as food items. It is possible there are functional gradients during evaluation in the vmPFC, representing different types of rewards at slightly different locations (McNamee et al., 2013).

Furthermore, if vmPFC represents hypothetical value it does so less than real rewards (Kang, Rangel, Camus, & Camerer, 2011). Interestingly, it does however appear to represent other peoples preferences if you choose for them (Nicolle et al., 2012) as well as more social altruistic value (T. A. Hare, Camerer, Knoepfle, & Rangel, 2010), indicating that although value might be concrete or stimulus/goal specific in the vmPFC, this can encapsulate many types of preferences in humans.

In rats, which do not have dissociable vmPFC and lateral OFC, just a possible "proto-OFC", OFC has been shown to be necessary for model-based or inferred value based decisions, but not for decisions simply based on past experiences, learning or working-memory (Jones et al., 2012). Barron and colleagues (Barron et al., 2013) have shown that the lateral OFC of human participants presents the value of food combinations that have not been experienced yet, further supporting claims of its necessity when combining different specific value information, especially when rewards are qualitatively distinct. Thus, it appears that lateral OFC can be crucial for making decisions, if those require the synthesis/inference of novel values from past experiences. Although lateral OFC and medial OFC/vmPFC are distinct anatomically, they are also highly interconnected (see Figure 8). As mentioned before, Lateral OFC carries a lot of value information especially in reward type discrimination tasks. It is also important for devaluation of specific food items and more complex reward updating is impaired if competing outcomes and stimuli are present.

Such reward-type specific representations would be immensely useful for generating a more specific model about the relationship between value and the world as well as dynamic updating of such values, according to internal and external context such as satiety or reliability. In other words, by helping to assign credit to specific outcomes, the central OFC might serve to disambiguate the relationships between stimuli and events with rewards (Mark E Walton et al., 2010). These specific links in turn could then be used in the vmPFC to guide accurate selection of the correct goal or object on the basis of specific expectation of the rewards to come (M P Noonan et al., 2010; M. P. Noonan, Kolling, Walton, & Rushworth, 2012). This would explain both the presence of a wealth of value information in both vmPFC and lateral OFC, their differences in encoding rewards and consequence of their lesioning. Behavioural impairments should be particularly strong if either decision or learning is particularly difficult and rely on disambiguation between many different choice dimensions or possible stimulus-reward association. The idea regarding concrete decisions reflecting currently held subjective preference between specific options relying particularly on the OFC circuit is more speculative. However, we do know that OFC seems to be more important for stimulus-based learning, compared to action-based learning (Camille et al., 2011; Peter H Rudebeck et al., 2008). Most studies engage the vmPFC by asking participants to choose according to preference and many individual differences in vmPFC signal are reflected also in current preferences. Not making deliberate choices between specific offers, reduces vmPFC activity during decisions. Interestingly, if concrete offers are compared and selected requiring the overcoming of a delay which can be seen as a feature of the reward, vmPFC seems to be involved. However, if additional cost constraints are introduced that are independent of the properties of the concrete offers such as effort, vmPFC appears to be disengaged.

Ecologically, the OFC might indeed be a system evolved for deliberation (Steiner & Redish, 2014) between specific potential goals or concrete options optimizing such decisions, rather than being necessary for making all decisions in the first place. In favour of this is the fact that the primate OFC is a relatively recent evolutionary development, possibly particularly relevant for a species with long distance vision and the possibility/luxury of selecting between more than one concrete offer at the same time. Indeed, such concrete offers could even be imagined or inferred, furthermore generalizing its usefulness for a species with a strong model of the world. Primate OFC has direct connections to higher order object-based representation in the perirhinal cortex. However, the usefulness of such a circuit might be reduced in situations in which such decisions are impossible or implausible. In a natural environment often a decision ought not to be made between a set of concrete options present or imagined. Instead, it could be made as a rejection or acceptance of an offer or course of action, in the presence of a richer sense of the reward environment and risks as well as other costs. But how is this process of deciding how to decide controlled if neither is sufficient for adaptive behaviour?

More specifically, there is considerable amount of evidence implicating both OFC as well as ACC in representing the value of different past and potential options. Although both seem to be decision relevant, they might be more or less so, depending on the particular decision problem, since they evaluate different possible courses of action, in their unique reference frames. However, one very important question still remains. How is it determined, if both systems are capable of making decisions, , whether circumstances demand that one area or the other should become decision dominant. To come closer to answering this question, it is important to briefly discuss the role of more lateral prefrontal areas in decision making and selection.

## Lateral PFC: Dynamic control, strategic preference and alternative courses of action

Cognitive psychologist always had a keen interest in the lateral prefrontal cortex, due to its large role in many cognitive processes of interest to psychologists. It appears to have many functions in attention (Corbetta & Shulman, 2002), working memory ((Passingham & Sakai, 2004) but also (Levy & Goldman-Rakic, 1999)), behavioural inhibition (Swann et al., 2009) and abstract reasoning (Donoso, Collins, & Koechlin, 2014). Since my thesis is focused on value guided choice and foraging behaviours, the discussion of the lateral prefrontal cortex will be more limited. This does not mean that it lateral PFC is unimportant in the comparison or selection process, but that my projects, like much research on value-guided choices, is less informative about their influence. There are, however, two exceptions I will discuss specifically. Firstly, the lateral frontal pole at the apex of lateral frontal cortex, appears to have important functions in modification of value guided behavioural strategies (see Chapter 5), holding alternative hypothesis in mind (Donoso et al., 2014). Secondly, the frontal operculum, part of the IFG, appears to influence behaviour not just through reactive or proactive inhibition as commonly thought previously, but it may also modify the impact of other regions on decision making (Chapter 5). Additionally, although I will not discuss dorsolateral prefrontal cortex (dlPFC) in depth, I will briefly mention some studies investigating how it might be involved in supporting more strategic choice behaviours. My hope is that the eventual neural data from my last behavioural experiment will shed some more light on the dlPFC’s involvement in decision-making and foraging style behaviours.

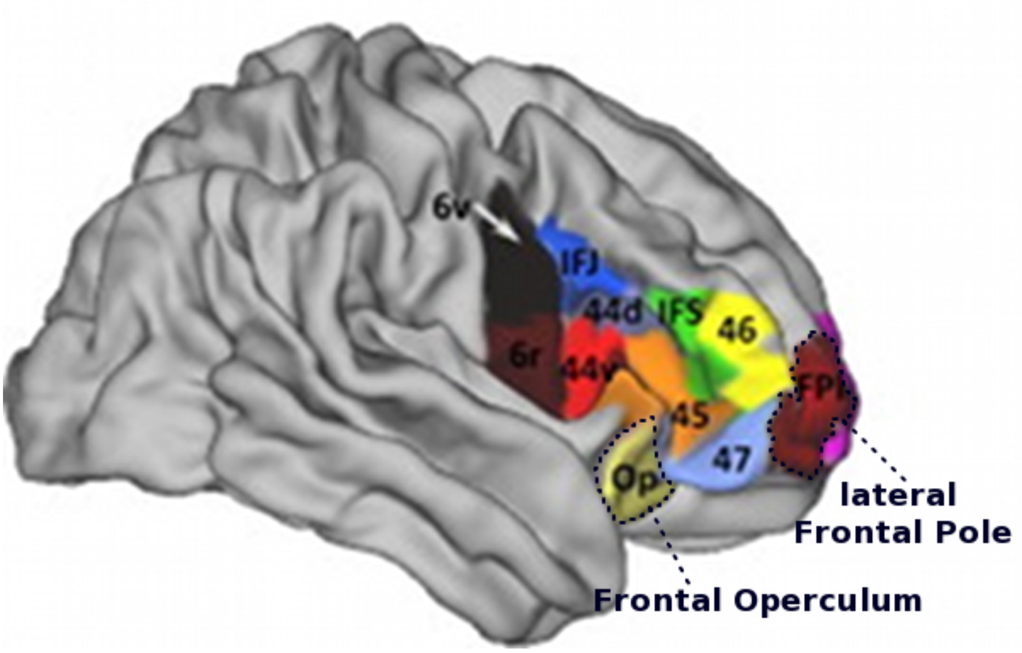

***Figure 9:*** *A recent Parcellation of the ventrolateral prefrontal cortex and Frontal Pole (Figure taken from* (Neubert, Mars, Thomas, Sallet, & Rushworth, 2014)*). The frontal operculum, part of the inferior frontal gyrus and the lateral frontal pole are both highlighted for clarity.*

When it comes to making choices, we know that simple rule representation exist in dorso- and ventro-lateral prefrontal cortex as well as in OFC, since single neurons are sensitive to non-match or match rules in macaques (J D Wallis, Anderson, & Miller, 2001). Many other studies have shown different ways in which dorso- and ventro-lateral cortex encodes current rule information in macaques (e.g. (Bussey, Wise, & Murray, 2002; Toni, Ramnani, Josephs, Ashburner, & Passingham, 2001)). Even more specifically, it has been shown how neural assemblies in the principle sulcus (dlPFC) can transition from input representations into response states (Go/No-Go) in goal specific manner using pattern classification (Stokes et al., 2013).

In a neuro-economic decision paradigm, very similar to the experiments mentioned before, Padoa-Schioppa recorded from principle sulcus neurons on the ventral and dorsal bank. However, whereas, in his other experiments macaques made decisions between two stimuli associated with different types of juice and could immediately move their eyes to the desired options, this time, the direction of eye movement needed for choosing either options was revealed only after the options had been shown. Thus a choice made between goods, had to be converted somehow to a spatial distinct response or action.

However, it was already known that the vlPFC is involved in action selection that is conditional on the use of an arbitrary and learned rule (Bussey et al., 2002; Toni et al., 2001). It is possible that the neurons studied by Cai and Padoa-Schioppa are simply carrying out this function.

We also know that neurons in the lateral prefrontal cortex in the posterior PS in front of the arcuate sulcus (dlPFC) can encode reward information (Barraclough, Conroy, & Lee, 2004) in a complex mixed strategy game. Furthermore, it encodes all decision relevant parameters in a simple neuro-economic tasks (Cai & Padoa-Schioppa, 2014) similar to his juice task discussed earlier. Neurons in the Principle sulcus can also encode different categories of behavioural sequences before their initiation in macaques (Keisetsu Shima, Isoda, Mushiake, & Tanji, 2007).

Additionally, neurons in the principle sulcus of macaques can distinguish between win-shift or win-stay strategies (Genovesio, Tsujimoto, & Wise, 2008) and do so less in incorrect strategy trials but always at outcome. They could thus support a simple behavioural strategy during such tasks.

A pressing question is then what is the difference between lateral PFC and medial PFC when it comes to reward-guided decision making, if both circuits are able to represent value and goal information?

Buckley and Colleagues (Buckley et al., 2009) are one of the few to directly compare lesions of parts of the lateral and medial PFC with each other regarding choice behaviour using a Wisconsin card sorting task (WCST) version on macaques. In this task the animals had to find out the rule by which a sample stimulus matching one of four presented options. Only the correct rule leads to reward at the end of the trial. They found that, ACCs lesions affected the ability of macaques to maintain rewarded rule responses even after several prior correct responses, whereas central OFC impaired rule based learning the most, especially after the first reward, as if they had problems learning the rule-reward association in the first place. Principle Sulcus/dlPFC lesions on the other hand made animals worse after interruptions, but did not have the same impact on reward learning. vlPFC lesioned animals had problems implementing the WCST rule post-lesion even without rule changes, preventing them from performing the task in the first place. Superior dlPFC lesioned animals showed the least impairments in the task. Thus, dlPFC appears to not impair reward learning itself, but supports interference free or learning over delays, often important to exploit reward information properly. For example, it appears to be important for calculating simple average reward rates, when both delay and reward size are manipulated at the same time (Simmons, Minamimoto, Murray, & Richmond, 2010), although this is a relatively simple piece of information for choice. Since ACC is also important for the creation of average reward estimates (C Amiez et al., 2006), a dlPFC-ACC circuit might be necessary to extract longer term value estimates encapsulated in average reward rates in behavioural models.

Another important fact about the dlPFC in regard to value-guided choices is its strong anatomical connections to the dACC (M. Beckmann, Johansen-Berg, & Rushworth, 2009; Lu, Preston, & Strick, 1994). Several people have argued that this circuit implements different forms of cognitive control for performance adjustments (Ridderinkhof et al.,

2004) or conflict resolution(J. D. Cohen, Botvinick, & Carter, 2000). It is plausible that the strategic information present in the dlPFC can modulate value representations and comparisons particularly in the dACC, adding several more abstract, strategic and rule based elements to the behavioural flexibility and foraging behaviour implemented by the dACC circuit. In many ways my last experimental chapter is meant to address this question, by asking whether truly prospective foraging requires dlPFC-centric computations to support dACC in its evaluation process.

Some have argued that its concurrent representation of reward and attentional/working memory information and relevance for rule learning and exploitation of arbitrary rule information, means it has a role in both letting reward affect other cognitive processes such as WM and allowing arbitrary associations between the world and rewarding outcomes (Jonathan D Wallis & Kennerley, 2010). Given the wealth of studies showing strategic information being encoded in the lPFC this might be a too narrow a role for the lPFC for decision making.

Apart from potentially helping with more complex average reward rate computations, we also know that dlPFC has a lot of information about current goals and strategies that can be used in order to pursue them. Thus, the dlPFC might be particularly important in decisions in which deliberate strategies need to be pursued, such as is the case in most model-based paradigms (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Gläscher, Daw, Dayan, & O'Doherty, 2010) or arbitrary symbols need to be transformed into a behavioural plans or cognitive strategies. Additionally, when rewards guide the use or maintenance of particular information about the outside world, dlPFC might help with such abilities.

## Frontal Pole

The function of the most anterior part of the prefrontal cortex is still highly disputed. Very few animal studies have looked at its function. One of the few experiments to look at it suggests it has a function in using and adjusting self-generated decisions strategies, but not explicitly cued ones (Tsujimoto, Genovesio, & Wise, 2012).

It has even been argued (Neubert, Mars, Thomas, Sallet, & Rushworth, 2014) there while there may be correspondences between the human medial frontal pole and the macaque frontal pole there might be no equivalent to the human lateral frontal pole in macaques. In terms of connectivity, the closest equivalent in macaques of the human lateral frontal pole might be dlPFC. Like human lateral frontal pole the macaque dlPFC representssome of the same pieces of information such as counterfactuals (Abe & Lee, 2011; M. F. Rushworth et al., 2012).

A couple of theories have emerged regarding frontal pole function from human fMRI experiments. For example Koechlin argued that it lies at the apex of a hierarchy of cognitive function, allowing the selection between cognitive sets (Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999). When people had to suddenly follow a temporary exogenous task instruction, and pick up on their own internal plan subsequently, lFP was also involved (Koechlin, Corrado, Pietrini, & Grafman, 2000).

Koechlin and colleagues (Donoso et al., 2014) further recently suggested lFP codes for the reliability of alternative courses of actions. This is very close to Boorman's and colleagues (Boorman, Behrens, & Rushworth, 2011) findings that the relative value/evidence for next best alternative choice is represented in the lFP. Interestingly, here the third best suppressed the lFP, whereas in Koechlin's study it activated it. This could be because sometimes participants had to pick the second best strategy in Boorman's study, but could never follow the worst strategy, whereas in Koechlin's study each strategy could be tested.

During a complex value guided spatial navigation task, Yoshida and Ishii (Yoshida & Ishii, 2006) observed lFP activity when there were many competing hypothesis, i.e. great uncertainty about what path would ultimately lead to the goal, and one particular one had to be selected as working hypothesis for navigation.

Overall, it could therefore serve as a representation of alternative courses of action or the testing i.e. confirmation or disconfirmation of other, alternative, hypotheses.

## Introduction summary

The last decades have greatly advanced our understanding of the function of the prefrontal cortex in decision-making and behavioural flexibility. However, much of this progress has been based on approaches to decision making that are grounded in neuroeconomics and cognitive psychology. A new approach applied in this thesis is neuroethological and combines models of ecological theory such as optimal foraging theory with their economic and cognitive counterparts to test their usefulness for understanding the solution of selection problems in the human brain. Some of the implications of foraging theories are exceedingly simple, such as the asymmetry in decisions between foraging and engaging with a concrete pray, food item or patch. Others are more complex, such as contextual modulations of decision processes according to ecological environmental constraints. This thesis is however not an attempt to just simplifying foraging theory and place it in the brain, but rather an attempt to use some of the concepts offered by these models in order to see if they shed light on the neural mechanisms of decision making. Other distinctions or concepts regarding foraging theory are inapplicable for understanding the neural mechanisms underlying decision making in humans. Equally, many conventional components from neuroeconomics and cognitive neuroscience are used to understand what drives human value guided behaviours.

By combining different perspectives for understanding of the generation of flexible behaviour and choice I hope to have used a novel approach to some important behavioural problems. For example how do we decide to engage or forage/search? Or do we take variable risks? How do we behave with foresight and with self-terminated strategies?

Secondly, I hope to contribute to a novel understanding about the brain, by asking what some of the evolved principal mechanisms of behavioural adaptation are. One major goal is to look at the possibility of there being multiple neural mechanisms for decision making or modules for choice. If such multiple decision modules exist, then the question is what is special about each. Here I argue that one possibility is that different areas are able to represent decisions in different reference frames. However, what information do they represent and for what adaptive purpose? Additionally, how do they flexibly interact to change decision making perspectives dynamically? Ultimately, integration between different perspective to decisions could converge dynamically through a flexible weighting of the influence of each evaluative system in their influence on behaviour.

More specifically, I will address three distinct issues of human behaviour, cognition and the neural mechanisms underlying it:

1) How are alternatives compared with specific offers? How do people make the judgement to pursue value driven foraging or decide to select between a limited set of offers instead? In essence, what are the neural mechanisms of human foraging?

2) How can people take risks flexibly when deciding between concrete options? Does the competition between different decision-relevant areas changes as preferences are updated?

3) Can humans use self-generated prospective information to guide their foraging behaviours? Can they take into account their own potential future behaviour when doing so. Furthermore, to what degree do they persevere through a sequence of choices after making the initial decision to be prospective?

# Chapter 2: Human Foraging Behaviour

## Abstract

*Adaptive foraging decisions are crucial for survival in the wild as well as more contemporary environments. Foraging decisions are essentially about whether to engage with currently encountered concrete options, or to instead continue searching elsewhere for superior alternatives. Whereas for most animals those decisions are mostly based on sequential reject/accept decisions between a concrete food/prey item and the environmental richness surrounding it, for humans such choices could be more abstract. We have shown that in a human equivalent to traditional foraging choices, participants took into account all necessary pieces of information to adaptively either engage with a current offer or reject it for the chance of better alternatives. As predicted from optimal foraging theory, they were sensitive to the costs as well as the average value associated with options constituting a foraging environment, as well as the current offer. Furthermore, they weighted the components that constituted the offer according to their relative desirability with the better component's value impacting choices more than the worse one during foraging decisions. Additionally, subject behaviour confirmed that subjects had learned and could use the values associated to each of the twelve stimuli used in this experiment.*

# Introduction

Recent insights into the neural mechanisms of decision-making have come from investigations in behavioral economics. Participants typically decide between limited numbers of options differing in probability, risk, and amount of reward (Michael L Platt & Huettel, 2008). Despite their success in explaining the choices animals make (Charnov, 1976; Stephens & Krebs, 1986) the optimal foraging models of ecology have had little impact on cognitive neuroscience (Hayden, Pearson, et al., 2011) or economics (Freidin & Kacelnik, 2011). The key foraging choice is usually not a binary one between currently available options, instead it is whether or not to engage with options as they are encountered (Charnov, 1976; Freidin & Kacelnik, 2011; Stephens & Krebs, 1986). It depends not just on 1) the value of the option encountered (*encounter value*) but also on estimates of 2) the environment's average value (*search value*) and 3) the cost of leaving to forage for alternatives (*search cost*) (Charnov, 1976; Hayden, Pearson, et al., 2011; Stephens & Krebs, 1986). I used functional magnetic resonance imaging (fMRI) to examine the neural mechanisms mediating foraging (next chapter). Prior to the fMRI analysis, in this chapter, I examined subjects' behaviour and the relationships between factors that would go on to be used as regressors in subsequent fMRI analyses, to establish their validity for describing this type of decisions.

Human participants made foraging-style choices (*forages*) to either *engage* with current options of known value or *search* among a set of potential alternatives also of known value. All the stimuli were drawn with replacement from a set of 12 with values, learned in a previous session (for training task see Figure 1 and methods). Pre- and post-scanning checks and analyses of choices during scanning confirmed value retention (Figure 4). Two visual stimuli indicated reward magnitudes potentially available if the subject engaged (their weighted combination constituted the *encounter value*; Equations.2-4). Rewards were points that translated into money on experiment completion. Six additional boxed stimuli indicated the values of the potential alternatives (*search value*). Choosing to search entailed a risk of paying a *search cost* (high, mid or low) in loss of points indicated by box color. If the subjects engaged they went on to make a comparative *decision* between the two components that constituted the encounter option, after being informed about their associated reward probabilities (Figure 2a). The introduction of probability information ensured decisions could only be made at this point and that forages and decisions were separated in time. When participants chose to search, new options drawn at random from the boxed alternatives were encountered. Participants searched as often as they wished but risked the same costs each time.

# Methods

## Subjects

We scanned 20 (12 female) healthy right-handed participants (age 22-32 years). We had to exclude two participants in all analyses, because they did not search more than 7 times during the experiment, preventing choice-related analysis. All participants gave informed consent in accordance with our national health services (NHS) research ethical approval (07/Q1603/11).

## Behavioral training

An initial training procedure was used to teach the subjects about the values of the stimuli they were to choose between in the scanning session. Participants were trained and tested on 12 stimulus-reward associations prior to the scanning session and retested afterward. The training was presented in the form of binary choices between the 12 stimuli drawn randomly (Figure 1) until they reached two criteria: (i) 95% correct performance over at least the last 40 trials; (ii) correct choice on the last occasion every stimulus was presented. As an additional test participants were asked to order all 12 stimuli according to their value. All participants were able to order the 12 stimuli correctly before MRI scanning. After scanning they were asked to order them again and at most two ordering errors for adjacent positions occurred. We took this to be sufficient evidence for the participants' knowledge of the value of all 12 shapes. For each subject tested the 12 stimuli were drawn from a larger set of 20 and the reward assignments were approximately counterbalanced across subjects.

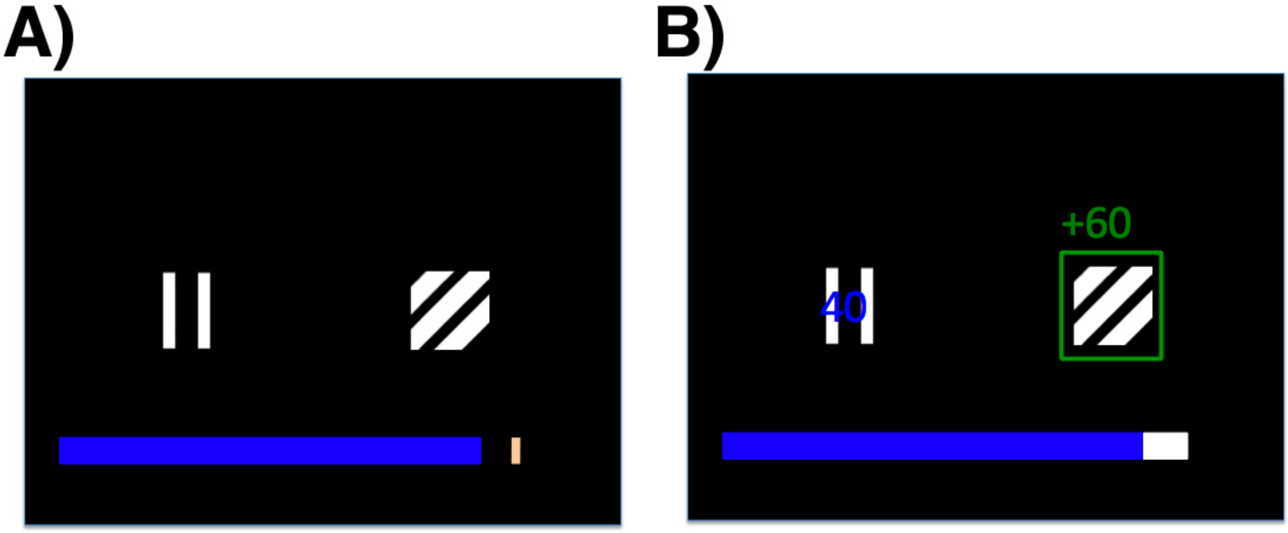

***Figure 1:*** *The binary choice paradigm used to train participant on the reward values of all 12 shapes. The participants chose between the left and the right object (A) and then received feedback (B) on the value of both stimuli. Training ended when participants reached criterion. Participants did not receive the money they earned during training. During training, however, the presented blue money bar served to illustrate the point system that was to be used in the main part of the experiment.*

### Experimental task

Participants had to repeatedly make two types of choice with the goal of maximizing the number of points gained throughout the task (overall 135 trials). The current amount of points they had accumulated was indicated by the position of a blue bar at the bottom of the screen that was present throughout the task. Every time the bar filled to a golden target, subjects received an additional UK £2. Reward magnitudes were all presented in the form of 12 pre-trained abstract stimuli worth between 20 and 130 points in 10-point increments (rather than direct numeric values), to ensure that the participants were engaged in value judgments.

First, subjects made a foraging choice whether to *engage* with a currently offered set of two options represented by two abstract stimuli to the left and right centre of the screen. The weighted value of these two options constituted the *engage value* (equations 2-4). Alternatively, they could choose to *search* for alternative options. A colored box framed a set of six alternatives in the upper part of the screen. The average value of these stimuli constituted the *search value*. Participants then encountered two new stimuli drawn at random from the set of six alternatives. Subjects were told that the stimuli would be drawn at random and that they would not be able to influence which of the six alternative stimuli were drawn. The old offer was put in place of the drawn alternatives in the alternatives box. Decisions to search entailed a risk of point loss (which constituted the *search cost*) with a likelihood of 70%. The number of points that might be lost was set to three different levels and the level on each trial was indicated to the subject by the color of the box surrounding the search alternatives (potential point loss: red = 20; white = 10; green = 5). Participants received feedback regarding point loss each time they made a search choice. Participants were allowed to search as often as they wished and indeed could potentially do so indefinitely. However, they only proceeded to the next trial, and encountered a new randomly drawn offer and set of alternatives, after making an engage choice. Variation in the search value, encounter value, and search cost is illustrated in Figure 3.

After engaging with an offer, participants had to make a decision, a binary evaluative comparison, between the two parts of the offer (left and right option). However, additional information was revealed at this stage, about the probability (between 20 and 90%) of actually receiving the points associated with each stimulus, together with the reward magnitudes, before making a decision. The proportion of a vertical bar that was filled indicated the probability of winning for each option. After choosing either the left or right option, participants were shown their choice and received feedback on the outcome of both gambles concurrently and were able to see the points they had gained.

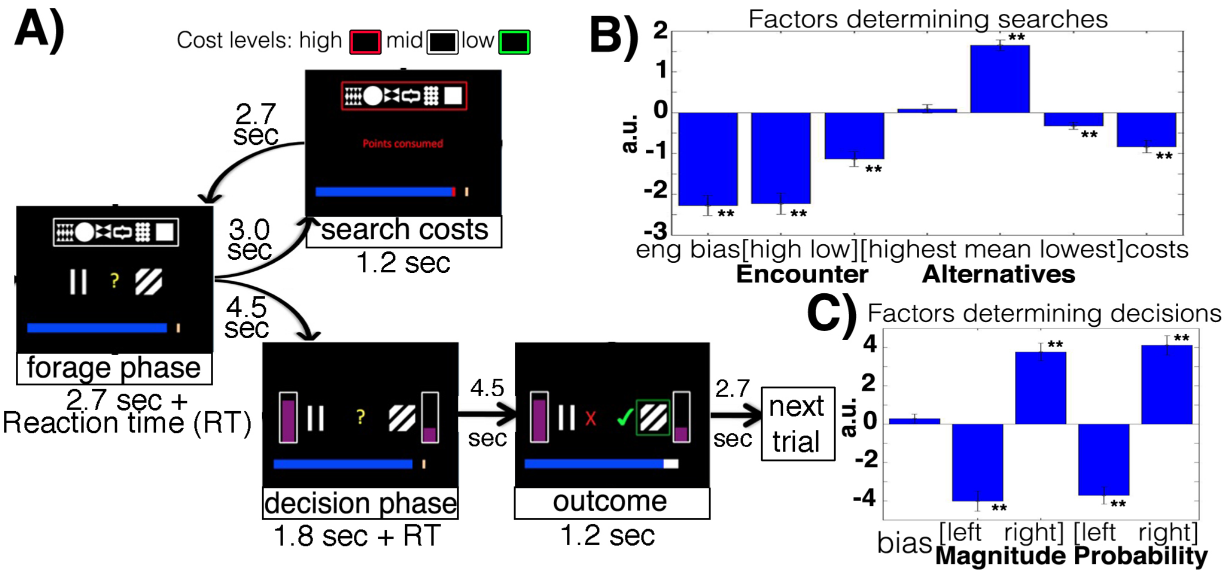

***Figure 2:*** *A) Trials started with two central stimuli (encounter value), six alternative stimuli (search value) in a box at the top (drawn from a set of 12 learned in a previous session) while box color indicated current potential search cost. The horizontal bar indicated previously collected points. The first choice was a forage – to engage with the encounter value or search for an alternative. Searching led back to the initial screen with a new encounter value drawn from the previous set of alternatives. Engaging led to the second type of choice – the decision – between the two component stimuli that constituted the encounter value. The pseudo-randomly determined reward probabilities were now*

*revealed. After the decision feedback indicated reward delivery. Factors (β-weights from logistic regressions) influencing likelihoods of search during forages (B) and picking the right stimulus during decisions (C) (**p<0.01).*

Participants observed the offers and alternatives, on average, for 2.7 s before they could make a choice (range 2-4 s). The jitter was drawn from a Poisson distribution with $\lambda$ 1.5 s and combined with an additional 200-ms random jitter to increase design efficiency. Subsequently participants were cued to respond by the presentation of a question mark in the centre of the screen, after which they responded with an average reaction time (RT) of 2.05 s [0.25 s standard error (S.E.) of mean]. If they chose to search they had to wait for, on average, 3 s (range 2-6 s). They were shown the cost outcome they had incurred for 1.2s (range 1-2 s). Participants then saw the new offer and alternatives. If participants chose to engage they waited for an average of 4.5 s (range 3-8 s), after which they observed the two options, including the newly presented associated reward probability for, on average, 1.8 s (range 1-4s). They responded, on average, 0.97 s (0.31 s SE) after the response cue was presented. The choice made was then framed and displayed for another 4.5 s (range 3-8s), after which they were shown the feedback from both options for an average 1.2 s (range 1-2s), followed by an inter-trial interval (2.7 s, range 2-4.5 s).

## Behavioral and fMRI parameters

### *Search value*

The parameter estimates representing the search value were simple averages of the values of the six stimuli that constituted the six alternatives.

$$Value_{search} = 1/6 * \sum_{i=1}^{6}(Value_i) \qquad \text{Eq. S1}$$

### *Encounter value*

The encounter value was a weighted mean of expected values ($EV_1$ and $EV_2$) of the two stimuli offered at the centre of the screen, with the weights ($w_1$ and $w_2$) determined by the optimal choice probabilities, given the two reward magnitudes associated with the stimuli ($r1$ and $r2$ for the lower and higher magnitude option, respectively) and the two associated probabilities ($p1$ and $p2$ for the lower and higher magnitude option respectively) that were to be drawn randomly from 20 to 90%. This was meant to generate a weighting more closely related to the one participants used as  it takes into account the fact that participants will not take an arithmetic mean between one very high and one very low stimulus to determine the value of an encountered offer.  So, for example, if a subject were presented with two stimuli associated with reward magnitudes 100 and 50 (in other words the magnitude of the first is twice that of the second) in the first part of a trial, then the weighting would be determined by the likelihood, in the next part of the trial, of the second stimulus having a reward probability at least twice as high as the first.  It expresses the fact that subjects should be biased towards the higher magnitude stimulus, which is a component of the encounter value, in estimating the encounter option's value before the probability information is revealed.

$$P(r_2/r_1 < p_1/p_2) = w_1 \qquad \text{Eq. S2}$$

$$1 - w_1 = w_2 \qquad \text{Eq. S3}$$

$$Value_{Offer} = w_1 * EV_1 + w_2 * EV_2 \qquad \text{Eq. S4}$$

Forage value difference, or search evidence was the simple difference in value between the average of all ***n*** potential new encounter values associated with the possible pairs of stimuli that might be drawn from the set of alternatives and the current encounter value offered.

$$Evidence_{Search} = 1/n * \sum_{i=1}^{n}(Value_{NewOffer}) - Value_{Offer} \qquad \text{Eq. S5}$$

The value of the two options at the binary gamble stage was the result of multiplying probability with magnitude of each option.

$$Value_{option} = magnitude * probability \qquad \text{Eq. S6}$$

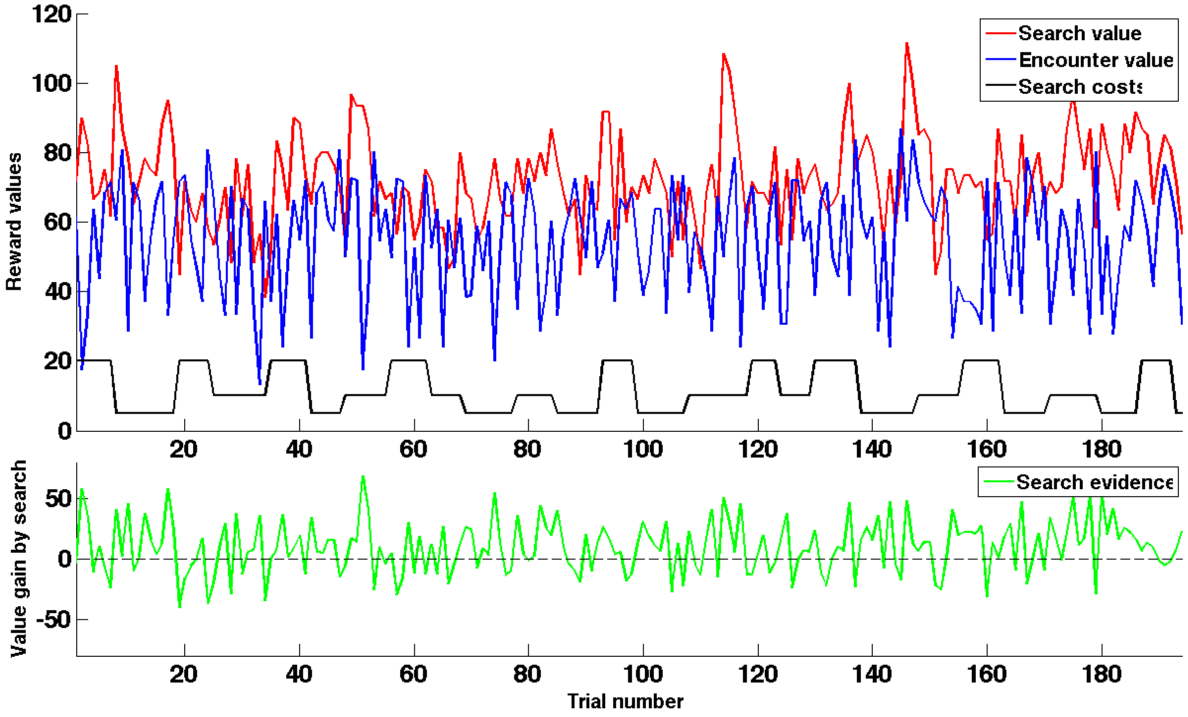

***Figure 3****: Variation in the parameters that determine foraging, search value, encounter*

*value, and search cost (top) and the evidence in favor of searching (bottom) that is, in turn, determined by these three parameters.*

# Results

Logistic regression identified factors weighing on forages and decisions. Engaging with the offers was promoted by increasing search costs and the higher and lower encounter values but retarded by all components of search values (Figure 2b, Figure 4, all p<0.01 in a one sample t-tests). More specifically, the average of the alternatives influenced foraging decisions, with each ranked stimulus value that was part of the alternative set having a significant impact on whether participants foraged or not (Figure 4 A). Participants were biased against search, requiring more value gain for searching than engaging, as seen by fitting a softmax on the relative foraging evidence minus the costs (Figure 5) allowing for a non-zero equivalence point (equivalence point significantly above zero [t(17)=3.17,p<0.01] in a one sample t-test) (differences in the constant from the regression reflects subjects' biases against searching; we call this parameter *forage readiness*). Decisions were influenced by reward probability and magnitude differences between options (Figure 1c, all p<0.01 in a one sample t-test). As a further illustration, in a regression with all twelve possible stimulus values, it can clearly be seen that the presence of the better stimuli makes forages more likely, due to the apparent negative slope of the beta weights.

Results of additional model-based analyses taking into account the possibility of repeated searches when computing the value of search are not reported here, as some participants very rarely searched more than once within a trial and the model-based analyses did not add any additional insight into subjects behaviours. However, a modified version of such a model is used in Chapter 6 to investigate prospective elements of the simpler foraging behaviours described here, in further detail.

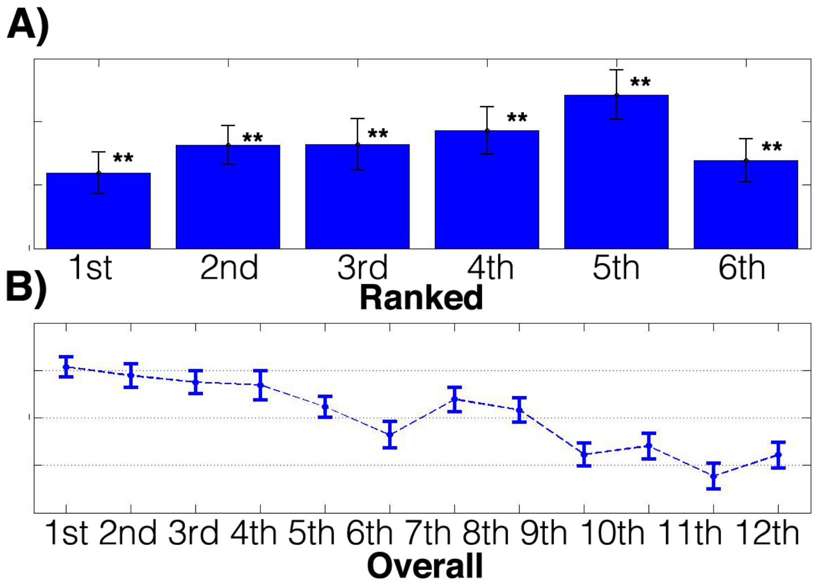

***Figure 4.** (**A**) The average β weights of a stepwise regression of the impact of the value of the stimuli that constituted the search value, ranked by value, on the likelihood of search decisions. (**$p<0.01$) (**B**) Logistic regression β weights of all 12 stimulus values on search decisions.*

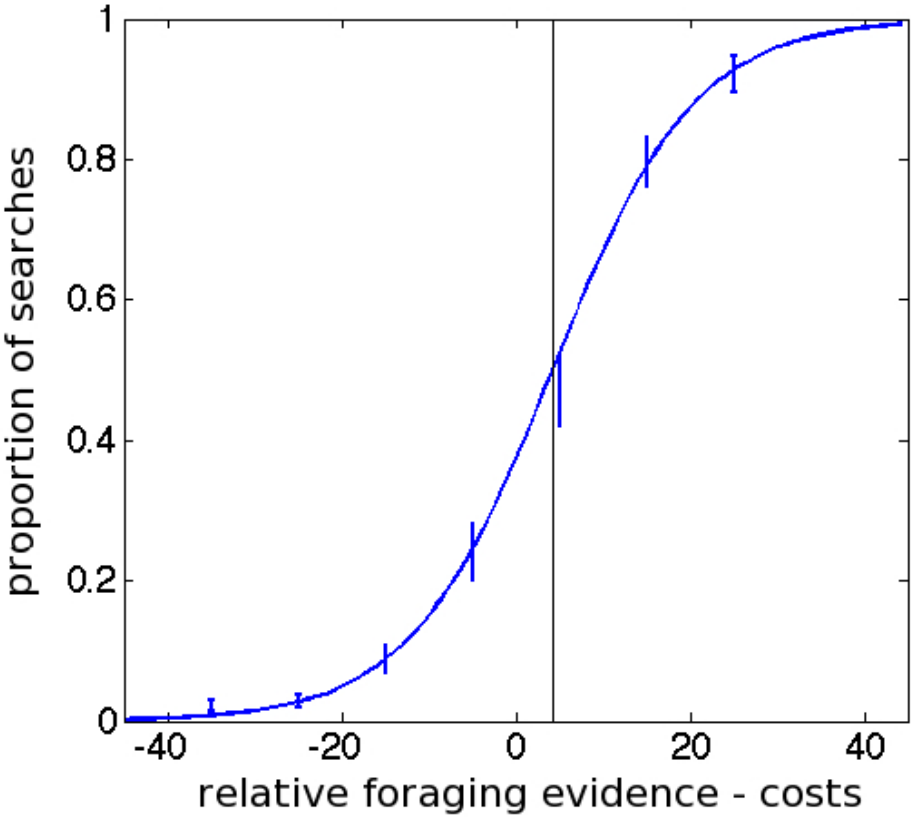

***Figure 5:*** *Softmax (blue) on proportion of search decisions in the foraging stage of the*

*experiment against relative foraging evidence minus the costs. This softmax allowed for a shifting equivalence point (black) and is plotted for all subjects together. Data is plotted as the mean with standard errors for the average between subjects for each bin. It illustrates that on average, participants were slightly biased against searching, but informed their decision strongly by the most relevant decision variable.*

## Discussion

Typically decision experiments within cognitive neuroscience have been about comparing two concrete offers when they possess one or more differing value dimensions, very much like the binary decision stage in this task. I, however, was interested in making participants make a different kind of decisions, more related to foraging-style understanding of common choice scenarios. Here, a sense of a current environment has to be compared to concretely encountered offers, also considering the costs associated with continued search.

In this chapter I was able to show that participants could make such kind of comparisons with prelearned values, also considering the costs of searching.

I used abstract stimuli as numbers could have led to more arithemetic strategies. Furthermore, by including the highest and lowest alternative into the regression model, I was able to exclude simpler heuristic strategies, compared to using the mean value of the environment itself. In the following chapter, I can therefore investigate the neural substrates of such an average value comparison with specific sets of offers.

# Chapter 3: Neural Mechanisms of Foraging

## Abstract

*Behavioural economic studies, involving limited numbers of choices, have provided key insights into neural decision-making mechanisms. By contrast, animals' foraging choices arise in the context of sequences of encounters with prey/food. On each encounter the animal chooses to engage or whether the environment is sufficiently rich that searching elsewhere is merited. The cost of foraging is also critical. We demonstrate humans can alternate between two modes of choice, comparative decision-making and foraging, dependent on distinct neural mechanisms in ventromedial prefrontal (vmPFC) and anterior cingulate cortex (ACC) employing distinct reference frames; in ACC choice variables are represented in invariant reference to foraging/searching for alternatives. While vmPFC encodes values of specific well-defined options, ACC encodes the average value of the foraging environment and cost of foraging.*

# Introduction

Through the advent of modern computing novel mathematical models of decision making have greatly helped our understanding of its neural underpinnings. From the beginning, cognitive psychology and economics inspired models have lent themselves particularly well to such modeling approaches, as they both already possessed some powerful descriptions of possible algorithms that could underlie human decision making (Kahneman & Tversky, 2000). Furthermore, those models offered possibilities of directly comparing neural activity recorded in different brain regions, with assumed intermediary processes that constitute a decision being made. This is particularly important for decision making, as it is by definition between perception and the actual action execution and not necessarily a direct consequence of either. When interpreting complex neural data, models are essential for constraining the interpretation of the observed activity.

However, as shown in the last chapter, human decisions can also be considered from a different perspective. Optimal foraging theory has also concerned itself with understanding decisions, albeit mostly decisions not made by humans, but by foraging animals. Nonetheless, it also offers unique perspectives and models to understand how adaptive decisions could be made.

In the past, the role of a two medial prefrontal brain regions in decision-making has been particularly contentious (See Introduction). Whereas some consider the OFC's lateral and medial surface to support decisions under any circumstances (Cai & Padoa-Schioppa, 2012, 2014; Padoa-Schioppa & Assad, 2006, 2008; Padoa-Schioppa, 2007, 2011, 2013), others have suggested the dACC to serve the role as the only prefrontal comparator region in the brain (T. A. Hare et al., 2011). Yet, other evidence is taken as suggesting that ACC might be particularly important for decision making in the action domain and learning about the environment, whereas OFC might contribute to stimulus based decisions and very precise credit assignment (M. F S Rushworth, Behrens, Rudebeck, & Walton, 2007; Matthew F S Rushworth, Noonan, Boorman, Walton, & Behrens, 2011).

In this experiment I set out to understand the neural mechanisms underlying decision making during foraging-style decisions. By looking at those decisions and comparing the activity with more commonly studied binary decisions, I wanted to dissociate different roles for the dACC and the vmPFC in generating adaptive decisions.

# Methods

## Data acquisition

A 3T scanner was used to acquire the T1-weighted high-resolution structural and the functional MRI data. For the functional MRI we used a Deichmann sequence (Deichmann, Gottfried, Hutton, & Turner, 2003) (EPI, repetition time (TR) = 3, voxel-resolution 3x3x3 mm$^3$, echo time (TE) = 30 ms, Flip angle = 87º. The slice angle was set to 15º with local z-shimming). This sequence reduces distortions in the ventromedial prefrontal cortex.

T1-weighted structural images were acquired for subject alignment using an MPRAGE sequence with the following parameters: Voxel resolution 1x1x1 mm$^3$ on a 176x192x192 grid, TE = 4.53 ms, inversion time (TI) = 900 ms, TR = 2200 ms.

## FMRI data analysis

The fMRI analysis was conducted with FMRIB's Software Library (FSL) (Smith et al., 2004). The data was denoised by using FSL Melodic independent component analysis after high-pass temporal filtering (3 dB cutoff of 100s), and Gaussian spatial smoothing (5 mm). Manually identified noise components were removed conservatively. Brain matter was separated from nonbrain by using a mesh deformation approach (Smith, 2002) and used to improve the registration of the functional data to the standard [Montreal Neurological Institute (MNI)] space by using affine registration with 7 degrees of freedom. Standard general linear models (GLMs) were fitted to the prewhitened data (M W Woolrich, Ripley, Brady, & Smith, 2001).

For group level analysis, images were registered to the high-resolution structural scan using 7 degrees of freedom and then to MNI space MNI152 template. We used FMRIB's local analysis of mixed effects (FLAME) (C. F. Beckmann, Jenkinson, & Smith, 2003) with outlier deweighting (Mark W Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). Reported cluster-corrected results came from these mixed effects analyses using Gaussian random-field theory and using the default $z > 2.3$ cluster-based thresholding and corrected cluster significance level of $p < 0.05$.

Regressors used in analyses of forages were time-locked to the onset of the forage stimuli (e.g. Figures 1,E and F, and 6A). The onsets of the encounter value stimuli, the search value determining alternative stimuli and the search cost determining box colors were simultaneous. Regressors used in the analyses of decisions were time-locked as follows. To analyze decisions four separate regressors were used to index the reward magnitudes and the reward probabilities associated with the option that was ultimately chosen and the option that was ultimately unchosen. The reward magnitude regressors were time-locked to the onset of the encounter value in the initial forage phase of the trial because it was at that point that the subjects first saw the stimuli which were associated with fixed reward magnitudes. The decisions could not be made at that point, however, because the reward probabilities associated with the stimuli were not revealed to the subjects. The reward probability information was only revealed at the point that the subject chose to engage and transitioned to the decision phase of the trial. The reward probability regressors were time-locked to the onset of the reward probability information at the start of the decision phase (Figures 1H and 6C). Thus, regressors associated with a decision were always time-locked to the moment the information became available.

The general linear model (GLM) used for the contrast of search value included the higher and lower encounter value, the highest and lowest alternative value and the cost (same parameters as in the behavioral GLM in the last chapter) during the forages.

The "foraging readiness" whole-brain contrast was generated by using this GLM and including the bias (shown in Fig. 1b of the last Chapter) together with the other normalized behavioral β weights in the higher order FEAT analysis (group analysis), and by looking at the correlation of such a bias with the main effects of the foraging phase.

The second GLM was constructed to directly compare contrasts of search evidence and foraging choice value difference with each other. This model included search evidence, foraging choice value difference as well as costs and reaction times. During the decision stage it included the value difference between both options, as well as the reaction time.

The time courses were derived from region-of-interest (ROI) analysis, calculating a mean time course within a ROI in each subject individually. The ROIs' coordinates were set in MNI space (spheres of a radius of three voxels, except the ventral striatum where the radius was two voxels) and then transformed to individual subject space by using the same linear registration as in the whole brain group analysis. Further details of vmPFC and ventral striatal ROIs are provided in Figure 5, and further details of BOLD effects in ventral striatum and ACC are shown in Figures 8 and 9. We then oversampled the time course by 10 and created epochs from the beginning of an event onward and applied a GLM to every pseudo-sampled time point separately. By averaging the β weights across participants we created the time courses shown (S.E. calculated between participants). For the slope estimate we oversampled by a factor of 20 to increase robustness of such estimates. Slopes were estimated using polynomial curve fitting on an interval of 6 s after event onset until each individual subject's signal peaked. The GLMs used in constructing the time series during foraging included the encounter value, the absolute weighting used to generate it (Eq. S2), the search value, and the search cost (the factors that are illustrated in Figures 1, E and F, and 6A). The subsequent decision-related time series were based on GLMs including chosen and unchosen reward magnitudes and probabilities and the search value from the preceding forage (the factors that are illustrated in Figures 1H and 6C).

I conducted a psychophysiological interaction (PPI) analysis (Friston et al., 1997) to investigate functional connectivity between ventral striatum and ACC. For this we generated a demeaned BOLD time-course regressor (from the ventral striatum) as well as an interaction term with the demeaned and convolved psychological regressor (search cost) separated according to choice, to generate an interaction term of the physiological and psychological regressor when search was chosen. We entered those regressors together with the other psychological regressors separated for choice into a FEAT whole-brain analysis.

## Results

As one of the main goals of our study was the comparison of foraging-style decisions and other types of decisions, we directly compared the main effect of both trial phases (Figure 1A). Furthermore, we looked at the difference in inverse value difference coding between the two types of decision, as this signal is commonly observed in the dACC (Figure 1B). More specifically we then tested what areas represent search value irrespective of choices made (Figure 1C), and what areas are more active as a function of relative foraging value difference, rather than the difference in value between the chosen and unchosen option (Figure 1D).

I then used an ROI based timecourse analysis as described in the methods to further interrogate the signal in the dACC (ROI coordinates shown in Figure 1C). Indeed, the timecourse illustrates that the signal appeared in both engage as well as foraging decisions, although the estimated slope was higher when people actually searched compared to engage trials ([t(17) = -2.54,p = 0.021]) (See Figure 2). Furthermore, costs only appeared to reduce ACC activity in search choice trials,.

Additionally, when searches were made, subjects with a higher beta-weight peak for the BOLD effect of the search value on dACC activity also exhibited a higher search value effect on their behaviour (Figure 1Gii), while a representation of the highest alternative value was predictive of participants making more search decisions after each other (Figure 1Gi). Looking at the dACC on trials when the engage option was chosen, it is possible to examine activity up until the point in time when subjects chose between the two components of the offer (Figure 1H). Doing so illustrates the commonly observed inverse value difference signal seen during such choices, but also a protracted search value signal that remained present until the binary choice had been made.

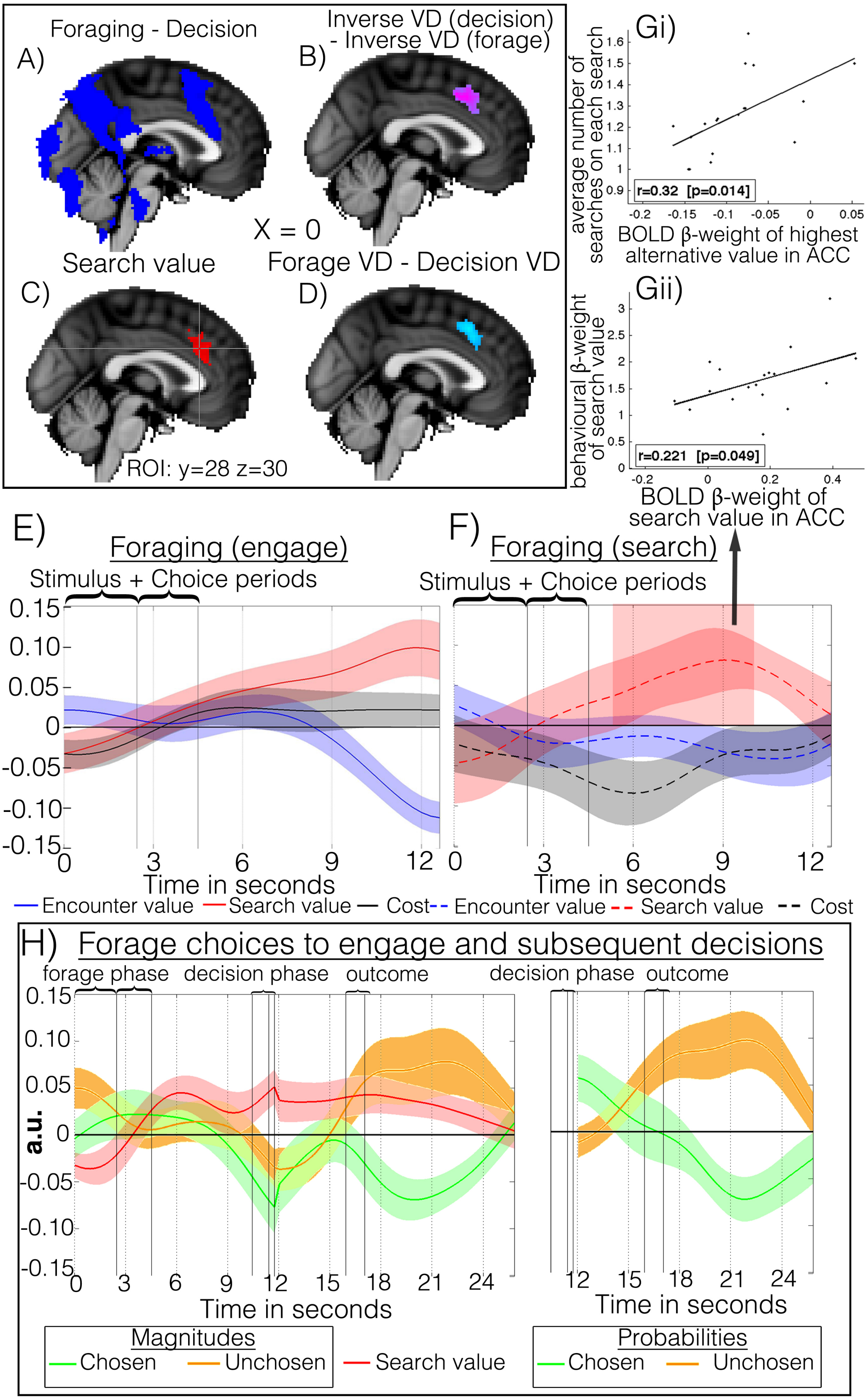

***Figure 1:*** *ACC activity was higher in forages than decisions (A), better related to the inverse value difference (VD) during decisions than foraging (B), reflected the main effect of search value during foraging (C), and better related to search VD than decision VD (D). ACC time courses during engage (E) and search (F). (G) Individual peak ACC BOLD β-*

*weights 5-10 s post-forage stimulus onset correlated with behavioral effects of the search value on search behavior (bottom) while ACC β-weights of best search value component predicted repeated searching (top). VmPFC exhibited no such correlations. (H) Time course for engage forages and the subsequent decision phase: The search value (red) signal continued into the decision phase. Reward magnitudes associated with chosen (green) and unchosen (orange) components of encounter value (left) were represented from their onset in the forage phase and into the decision phase. The reward probabilities of the chosen and unchosen options were only revealed after engaging and their BOLD effects therefore appear later (right).*

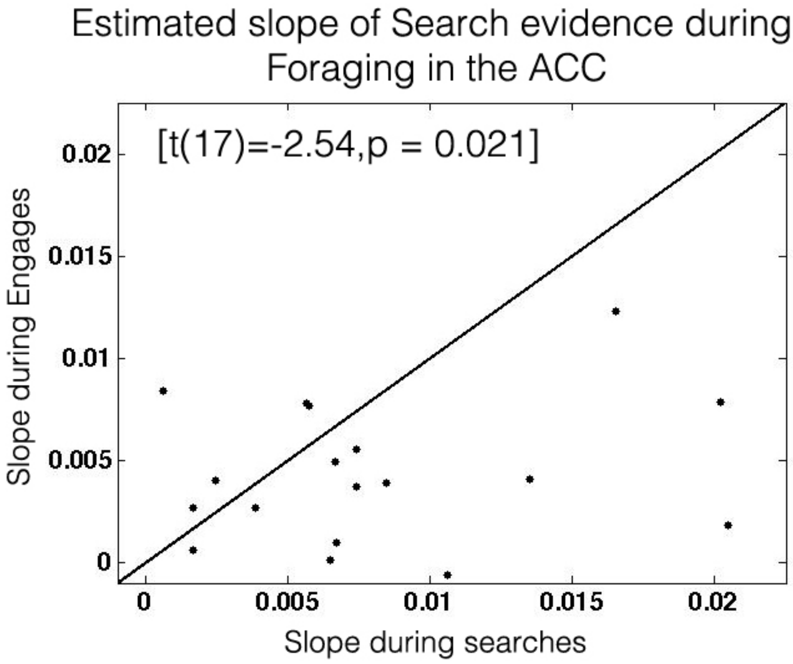

***Figure 2:*** *The estimated slope of the search evidence β weights in the ACC during foraging decisions, when search and when engagement is chosen. The slope difference is significant [t(17) = –2.54, p = 0.021].*

Although there were no strong overall effects of the highest or lowest alternative option values on dACC activity, the signal for both was correlated with their respective behavioural beta-weights predicting foraging choices (Figure 3).

Finally, after participants had chosen to search, an analysis of the timecourse suggests that the new encounter or offer value had a negative effect on ACC activity, whereas search value still activated it (Figure 4).

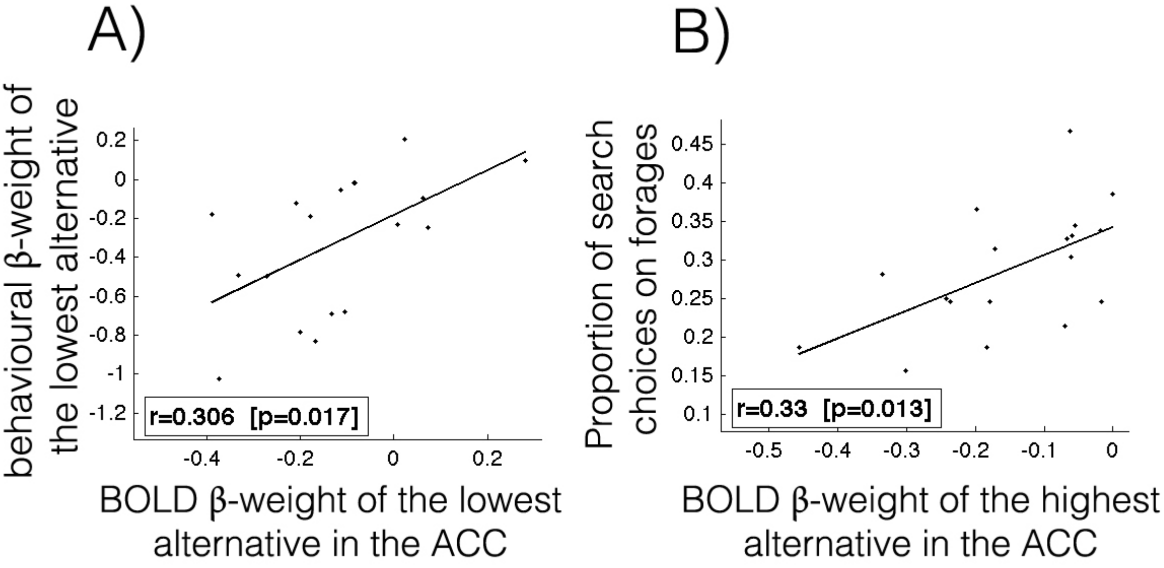

***Figure 3:*** *Correlations, across subjects, of individual subjects peak BOLD β-weights (standardized BOLD regression coefficients) in the ROI shown in Fig. 2, 5 to 10 s after event onset with behavioral effects (logistic regression coefficients for behavioral predictions from figure 1b, as well as proportion of choices as search). Peaks were measured from the individual subject time-courses averaged in Figure 1.*

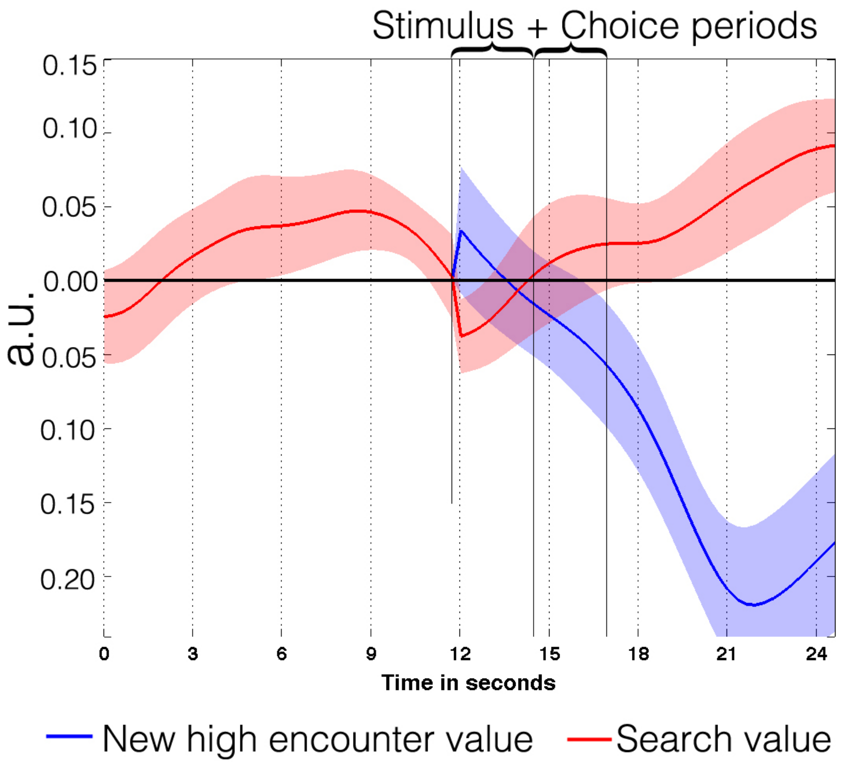

***Figure 4**: Time course of BOLD β weights in relation to feedback after foraging decisions in ACC: The search value (red) during search decisions and the new encounter value (blue trace) β weights are shown (compare Fig 8A showing comparable data from ventral striatum).*

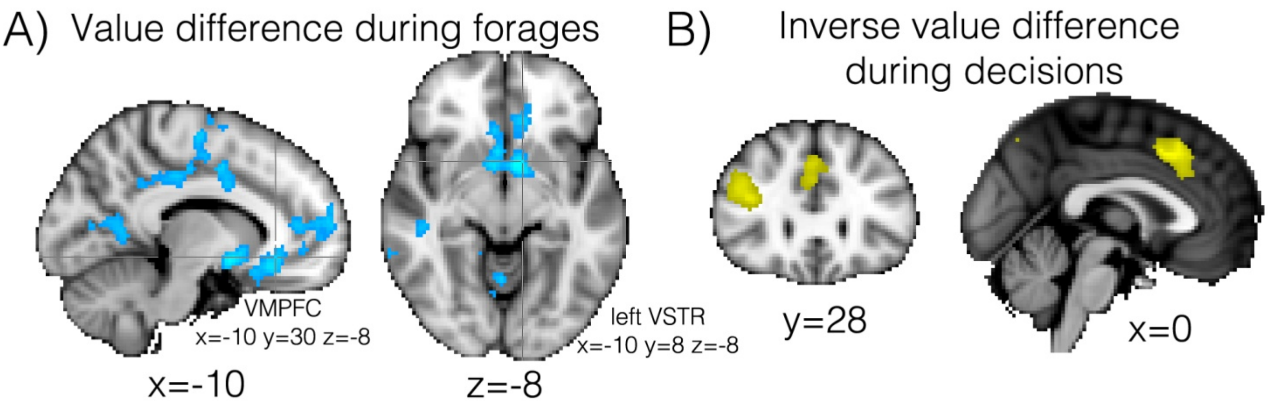

***Figure 5.** Whole brain level comparisons showing: (**A**) activity due to the choice value difference during foraging including the vmPFC (left) and ventral striatal ROI (right) coordinates; (**B**) activity due to the inverse of the value difference during decisions.*

By contrast the vmPFC was more active as a function of higher relative value for the chosen option (Figure 5A), whereas the inverse value difference appears in the ACC for the binary decisions (Figure 5B). Further investigation of the vmPFC signal (for coordinates see Figure 5 A), however, revealed that there was only a representation of the chosen value on the trials on which participants chose to engage, but not on trials when they foraged (Figure 6A). In other words, the average value of foraging never had a positive impact on vmPFC activity. This is consistent with the fact that the chosen-unchosen value difference contrast was associated with vmPFC activity while the contrast between forage and engage values was not associated with vmPFC activity (Figure 6B).

In addition, the amount participants relied on the higher or lower component of the encounter value behaviourally was also reflected in the vmPFC activity (Figure 7). At the decision stage, higher behavioural accuracy, as indexed by the inverse temperature, was positively related to the chosen reward magnitude signal in the vmPFC.

Finally, we also looked at the ventral striatum (coordinates shown in Figure 5A). The ventral stratium is perhaps most known for its prominent reward prediction error signals. Therefore, we looked at whether we could find a foraging prediction error signal on trials on which subjects opted to search. Analysis of the ventral striatal time course suggested that a new offer value increased striatal activity whereas a high expectation, in form of a high search value, decreased it. Not only that, but the peak activity in response to the new offer value, was negatively related to the proportion of occasions on which subjects actually opted to search. One interpretation of this finding is that the higher value signal at outcome was due to lower expectations about the value of foraging (Ci). Using the BOLD timecourse analysis approach, we could also look at effects of foraging cost on striatal activity. Subjects in whom ventral striatal activity was more closely related to costs when foraging, were also the subjects who were less inclined to forage (Cii), suggesting this activity might be related to overcoming cost related inertia. Additionally, the ventral striatum had an increased connectivity to the pre-genual ACC when there were higher costs and a subject actually searched (D), in exactly the same part that was more active in participants that were more inclined to forage, i.e. with increased foraging readiness as indexed by the constant of the behavioural GLM (Figure 1 in last chapter), suggesting a mechanism for overcoming such cost induced inertia.

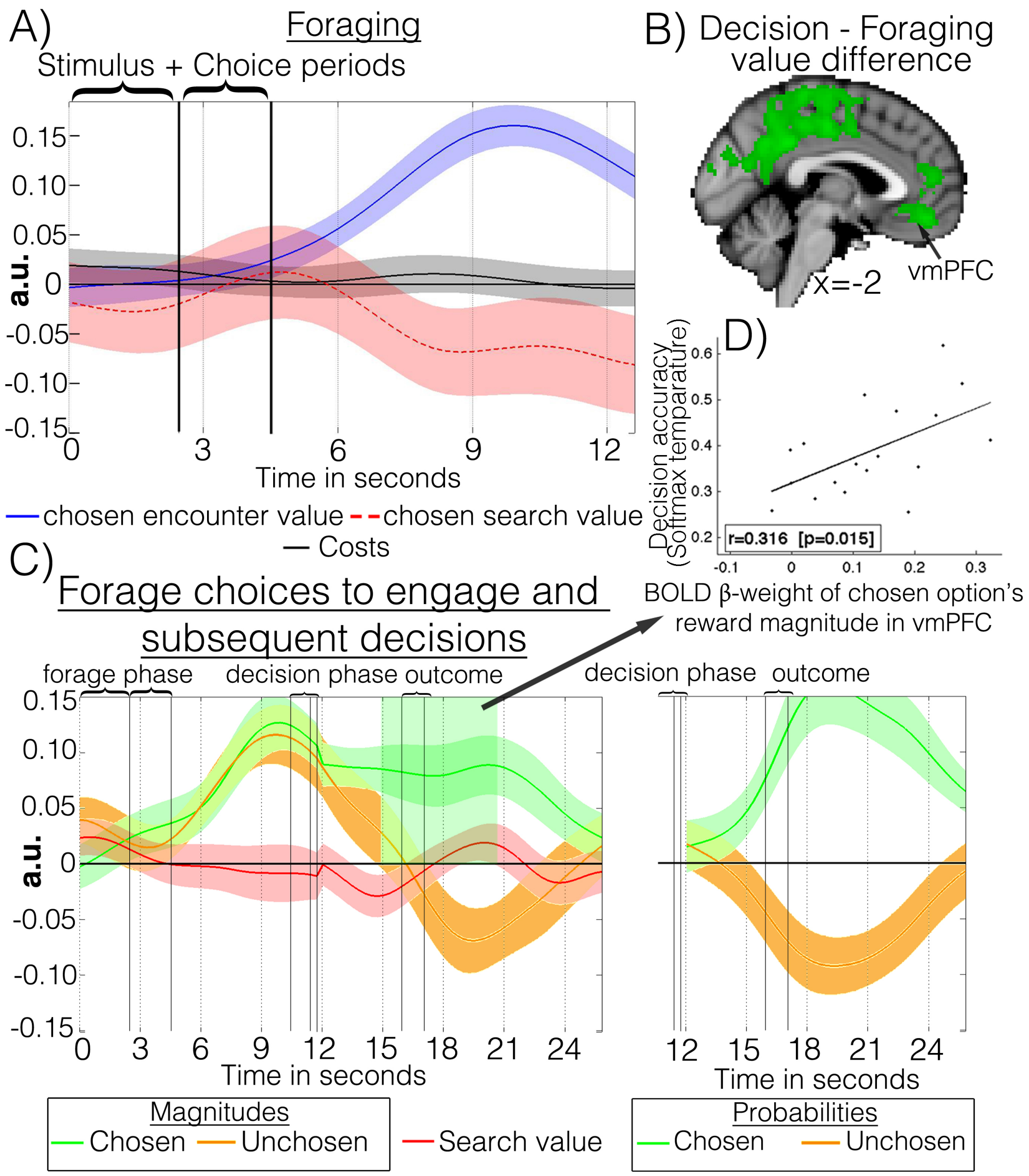

***Figure 6:*** *(A) VmPFC time courses during forages (conventions as Figure 1e). (B) Activity better related to decision VD than to forage VD. (C) VmPFC time course for engage forages and the subsequent decision phase (conventions as Figure 1h). D) Individual peak vmPFC*

*BOLD β-weights 5-10 s post-decision onset correlated with estimates of decision accuracy (softmax temperature).*

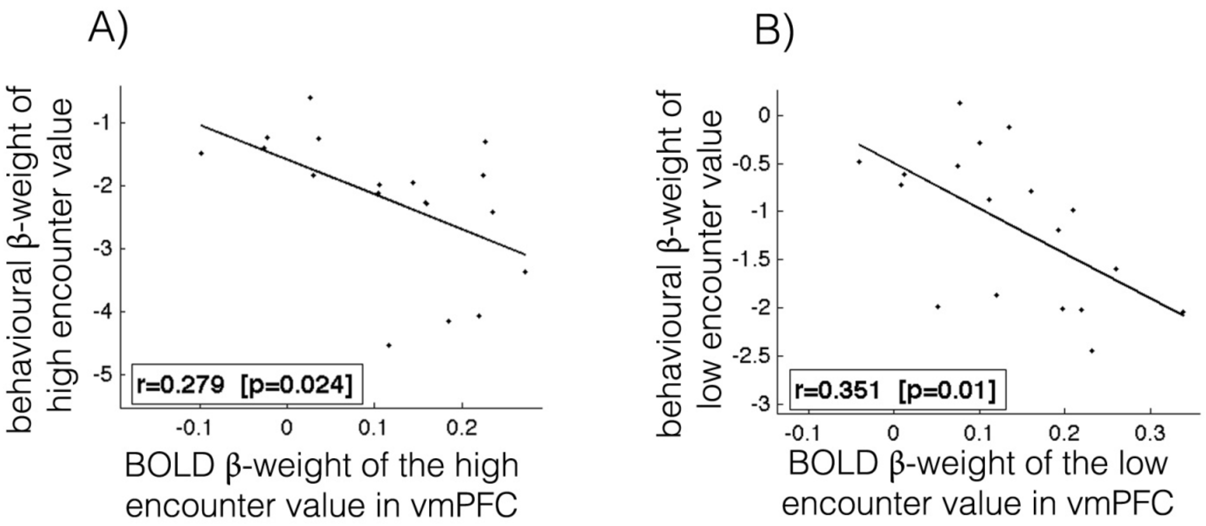

***Figure 7:** Correlations, across subjects, of individual subjects' peak BOLD β weights (standardized BOLD regression coefficients) 5 to 10 s after event onset with behavioral effects (logistic regression coefficients for behavioral predictions from Fig. 1B Chapter 2). Peaks were measured from the individual subject time-courses averaged in Figure 6 using the higher and lower encounter value as separate regressors.*

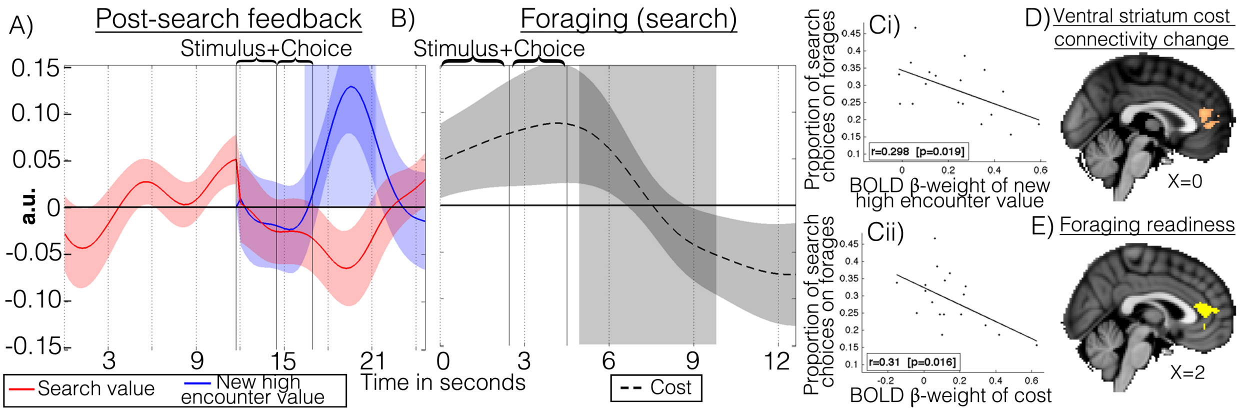

**Figure 8:** *Ventral striatal time courses after feedback following search forages (A). Effect of search costs when search is chosen (B). Individual peak BOLD β-weights for new encounter value (Ci) and peak BOLD β-weights for new search costs on searching (Cii) 5-10 s post-event onset both correlated with the proportion of forages on which participants searched. Increased coupling with left ventral striatum as a function of search cost during searches (D) and individual differences in foraging readiness (E) both revealed an ACC region anterior to, but overlapping with, that in Figure 1D.*

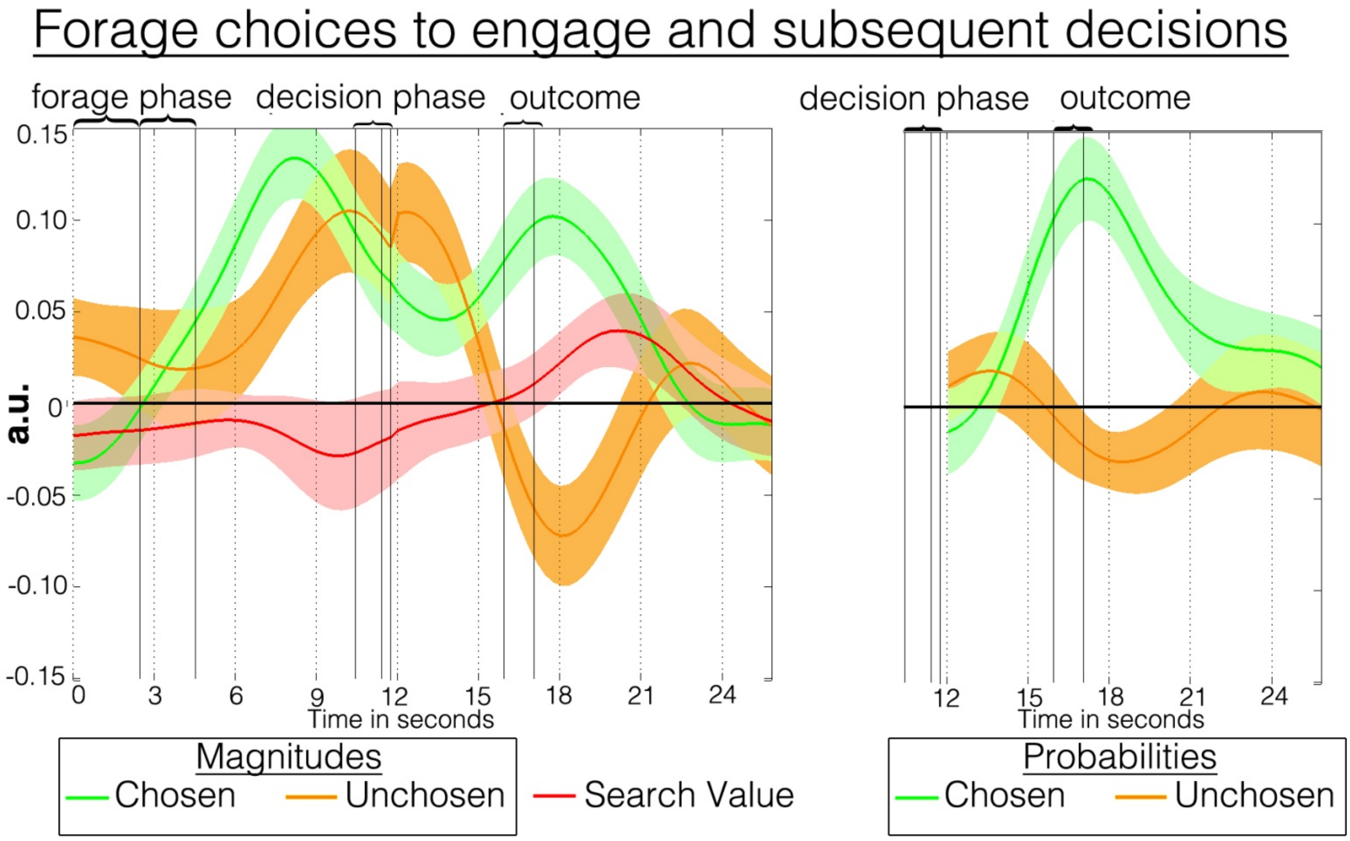

**Figure 9:** *Time course of the ventral striatal β weights relating the BOLD signal to chosen and and unchosen option values. Chosen and unchosen option values are decomposed into, first, reward magnitude components that were revealed at the start of the forage phase of the trial with the onset of the encounter stimuli and, second, reward probability components that were only revealed at the onset of the decision phase of the trial. The BOLD β weights for search value from the forage phase of the task is also shown*

*throughout both that task period and the subsequent decision phase. Data are shown for all trials on which subjects chose to engage in the forage phase. Overall a similar pattern to the vmPFC emerged, although some residual search value representation from the prior forage period that persisted into the decision period suggests possible effects of reward context on this structure.*

We also looked at the timecourse of the ventral striatum during and after engage choices. Overall the ventral striatum BOLD timecourse is very similar to the vmPFC BOLD time course.

| **Comparison** | **Anatomical Region** | **Hemisphere** | **Peak Coordinates (mm) (*x*, *y*, *z*)** | **Maximum *z* score** |
|---|---|---|---|---|
| **Foraging value difference – Decision value difference** | ACC | – | –2, 16, 42 | 4.41 |
| | Dorsolateral PFC | R | 40, 30, 20 | 4.09 |
| | Insula | | 34, 18, 4 | 4.45 |
| | IPS | | 16, –74, 48 | 3.47 |
| **Search value difference – Decision value difference** | ACC | – | 2, 12, 44 | 4.82 |
| | Dorsolateral PFC | R | 40, 30, 22 | 4.79 |
| | Insula | | 32, 18, –2 | 4.57 |
| | IPS | | 34, –58, 48 | 4.1 |
| | Visual Cortex | | –32, –88, –4 | 3.94 |
| | Visual Cortex | L | 20, –88, –12 | 3.27 |
| **Search value** | ACC | – | –4, 32, 20 | 3.32 |
| | IPS | R | 52, –30, 44 | 3.18 |
| | IPS | L | –22, –72, 54 | 3.24 |

***Table 1:*** *All significant cluster-corrected results for the contrasts in Figure 1, B to D. In the column for hemisphere, dash indicates not applicable.*

# Discussion

Overall, the pattern of results converge onto a few key findings. Firstly, dACC appeared to encode many decision relevant pieces of information in a foraging framework and secondly, the vmPFC, albeit being very well suited for binary comparisons between concrete options, did not encode search value when it was chosen.

Comparison of average activity during foraging and decisions identified ACC among other regions (Figure 1). Usually in decisions, the most common signal observed in ACC is inversely related to the value difference between chosen and unchosen options. Such inverse value difference effects have been interpreted as indicating that ACC/dorsomedial frontal cortex is a "comparator" comparing choice values. According to this theory the region is more active when unchosen values are larger because a smaller difference between chosen and unchosen values means comparison takes longer before a choice is made (Basten et al., 2010; T. A. Hare et al., 2011) (Figure 1H). Related accounts emphasize an ACC role in monitoring for conflict between responses (Mattew M Botvinick, 2007).

However, our task also allowed us to test whether the ACC signal reflects the relative benefit of the alternative course of action or the value of exploring the environment. This hypothesis predicts that ACC, during forages, will stop reflecting the value of the unchosen option, and *always* represent the value of searching. We therefore refined the analysis and tested for a region that demonstrated both of these effects: Coding for the unchosen–chosen value difference during decisions but not forages (Figure 1B), and, on forages, instead coding for the search value(Figure 1F).

Both tests identified overlapping ACC regions. When these two effects were combined into a compound test (forage$_{\text{(search value–encounter value)}}$-decision$_{\text{(chosen value-unchosen value)}}$) the same ACC region was implicated (Figure 1D).

We analyzed foraging signal time courses in a region centered on the overlap between foraging search value and decision value difference effects (Figure 1C,D). ACC BOLD was positively correlated with the value of searching the environment, and negatively correlated with the value of engaging with the current encounter option, regardless of the choice participants ultimately made (Figure 1E,F). The frame of reference in which values are encoded in ACC is thus fixed in relation to response strategy, searching or engaging. This contrasts with vmPFC and other regions where value is encoded in a flexible reference frame tied to the choice taken or attended (Boorman et al., 2009; Lim, O'Doherty, & Rangel, 2011). Comparing search value signals in ACC, we found a more rapid increase (greater slope) on search than engage choices [t(17) = -2.54,p = 0.021] consistent with earlier, stronger signals in search decisions (Figure 2) and faster accumulation of search evidence in ACC on search choices (Hayden, Pearson, et al., 2011). In search choices there was also an effect of search cost (Figure 1F).

We next examined whether individual differences in ACC activity reflected differences in foraging. Both, behavioral variation in the influence of search value in promoting searches was correlated with neural variation in ACC search value effects(Figure 1gii) and behavioral differences in the influence of the lowest and highest alternative values were correlated with ACC activity (Figure 3). While average search value determined search choices (last chapter) it did not predict the rate at which participants repeatedly searched again and again in pursuit of the best alternative on each trial. Such perseverative search rates were, however, predicted by ACC responses to best alternatives (Figure 1Gi). Finally we looked at the decision phase; ACC activity still reflected the search value from the prior forage, as if still encoding how good it would be to search for alternatives (Figure 1H). Brain activity conveyed knowledge of environmental richness even during simultaneous binary decision-making when the signal was no longer relevant. Knowledge of environmental richness, which is normally pertinent to foraging but irrelevant to binary decision-making, impinges on, and impairs, simultaneous binary decision-making in behavioral experiments (Freidin & Kacelnik, 2011).

Despite their limitations (Mansouri, Buckley, & Tanaka, 2007) and alternative explanations of reward- and error-related activity in ACC (Mattew M Botvinick, 2007; Matthew F S Rushworth et al., 2011), conflict and comparator-based theories remain the most influential accounts of decision-related activity in ACC. However, the presence of an average reward signal (search value), a negative effect of search cost, anchoring of value representations with respect to search/engage strategies, differential rates of search signal accumulation on search and engage trials, and correlation, across subjects, between ACC signal variance and search choice variance (Figures 1,3) cannot be accommodated within comparator- and conflict-based ACC theories. Instead we suggest ACC codes the value of switching to a course of action alternative to that which is taken or is the default. ACC supplies such a signal even when subjects are not asked to forage but to make decisions. As soon as the subject switches to the alternative the signal dissipates but it is maintained if the course of behavior is maintained (compare red lines 1F versus 1E,H).

VmPFC encodes the value of chosen/attended options in comparison to unchosen/unattended options (Boorman et al., 2009; Lim et al., 2011; M P Noonan et al., 2010). During foraging, however, vmPFC activity only reflected the chosen option value when participants engaged and there was no representation of search value(Figure 6A). When subjects searched, the chosen search value was actually negatively correlated with vmPFC activity and there was no representation of encounter value. The absence of any representation of search value – the average value of the environment – and of search cost(Figure 6A) restricts any role vmPFC might play in foraging.

In contrast, seconds after foraging vmPFC played an important role in decisions. Comparison of average activity during decisions and forages and between decision and forage value differences(decision$_{\text{(chosen value-unchosen value)}}$-forage$_{\text{(chosen value–unchosen value)}}$) identified vmPFC (Figure 6B). It coded, negatively and positively, for values of unchosen and chosen options respectively. It effectively encoded the value difference between options. During the transition from foraging to decisions, vmPFC rapidly changed from positively encoding both components of encounter value, weighting both in the same way as participants did behaviorally (Figure 7), to representing the value difference between chosen and unchosen components in decisions (Figure 6A,C). The reference frame in which values are encoded in vmPFC is thus flexible and concerned with the value dimensions and contrasts most pertinent to decision-making. Such a reference frame makes vmPFC suitable for goal-based (Wunderlich et al., 2012) and multi-attribute (Fellows, 2006) decision making. Its importance during decisions was underlined by individual variation in vmPFC reward magnitude effects being correlated with decision accuracy (Figure 6D).

Reward prediction error signals associated with the ventral striatum, and its interactions with orbitofrontal cortex (M P Noonan et al., 2011), allow decision-making to change with experience. They occur even when there is little opportunity for learning (T. a Hare, O'Doherty, Camerer, Schultz, & Rangel, 2008) as in our task. We therefore examined whether forage prediction errors were also encoded by the striatum (Figure 5A) and its interactions with the ACC. Despite its weak activation with search value it exhibited post-search prediction error-like signals (positive effect of new encounter value, negative effect of previous search value:Figure 8A). It also responded to search costs (Figure 8B). The prediction error response had higher positive peaks in people who searched less (as if they had expected less:Figure 8Ci).

Across subjects, search costs activated striatum in proportion to the degree that they deterred searching (Figure 8 Cii).

An ACC region overlapping with, but anterior to, the search value effect (Figure 1C) was more coupled with left ventral striatum when search costs increased and search was chosen (Figure 8D). The coupling appeared related to disinhibition of effortful choices because the same ACC region was also more active in subjects more willing to overcome costs; individual differences in *foraging readiness* were associated with increased anterior ACC activation (Figure 8E).

VmPFC and ACC have been thought to operate in sequence during choice (T. A. Hare et al., 2011; M P Noonan et al., 2011) but our results suggest ACC represents choice in a manner at odds with intuitions of how comparative decisions are made. Because ACC value representations are anchored to response strategy (engage/search), our results confirm it is well placed to guide response selection. However, the different signals in ACC and vmPFC attest to independent roles in forages and decisions. The implication of ACC in foraging and encoding of the average value of the foraging environment may facilitate understanding of the reward signal it carries (Matthew F S Rushworth et al., 2011; Seo & Lee, 2007; Vickery, Chun, & Lee, 2011), its prominence during exertion of effort (Croxson et al., 2009; Peter H Rudebeck et al., 2006), in go-no-go decisions (K. Matsumoto et al., 2003), exploration (Daw et al., 2006; Quilodran et al., 2008) and in representing alternative and counterfactual choice values (Boorman et al., 2011; Hayden et al., 2009). Some action value learning tasks previously used to investigate ACC (Matthew F S Rushworth et al., 2011) may have been treated as foraging tasks and animals may have been choosing whether to stay with the current choice or switch to an alternative. Such a perspective also makes it possible to reinterpret ACC activation recorded during exploration tasks (Daw et al., 2006) as reflecting estimates of richness of alternatives in the environment. ACC activity is frequently recorded (Poldrack, 2006) and might reflect the value of alternative choices in other tasks and the inclination to refrain from engaging in the currently offered choice (Magno, Foxe, Molholm, Robertson, & Garavan, 2006). Foraging entails energetic costs and we found ACC activity also reflected the cost of foraging. ACC neurons have been shown to encode value signals that integrate both cost and reward (Steven W Kennerley et al., 2011). By contrast, vmPFC, a primate specialization (Wise, 2008), may underpin fine-grained, accurate, and flexible decision-making (T. A. Hare et al., 2011; Wunderlich et al., 2012).

# Chapter 4: Adaptive decision-making under changing risk pressure

“ There is a tide in the affairs of men.

Which, taken at the flood, leads on to fortune;

Omitted, all the voyage of their life

Is bound in shallows and in miseries.

[Brutus in *Julius Caesar Act 4*]”

## Abstract

*Sometimes when a choice is made, the outcome is not guaranteed and there is only a probability of its occurrence. Each individual's attitude to probability, sometimes called risk proneness or aversion, has been assumed to be static. Behavioral ecological studies, however, suggest such attitudes are dynamically modulated by the context an organism finds itself in; in some cases, it may be optimal to pursue actions with a low probability of success but which are associated with potentially large gains. We show that human subjects rapidly adapt their use of probability as a function of current resources, goals, and opportunities for further decisions.*

# Introduction

An understanding of risk and opportunity is essential for success and survival and there has been interest in the neural representation of risk, probability, and value (Michael L Platt & Huettel, 2008). We know individuals differ in attitudes to risk and probability. For example, people prepared to pursue a course of action that might lead to great potential gain (a large reward magnitude) even if there is a low probability of obtaining the outcome are said to be risk prone, while others are called risk averse. Such variation in attitudes is linked to individual differences in brain activity (Christopoulos, Tobler, Bossaerts, Dolan, & Schultz, 2009; Tobler, O'Doherty, Dolan, & Schultz, 2007). It is recognized that such attitudes differ depending on the type of prospect contemplated, for example, whether it is a potential gain or a loss (Kahneman & Tversky, 2000), but within a given frame there has been less investigation of how the use of probability to guide behavior changes with circumstances. Despite the existence of individual differences in risk attitudes it is possible that how each individual evaluates probability also changes with context.

It has been apparent to behavioral ecologists interested in Risk Sensitive Foraging Theory (RSFT) that dynamic changes in risk attitudes occur across time within individual foraging animals (Caraco, 1981; Hayden et al., 2009; Kacelnik & Bateson, 1997; McNamara & Houston, 1992; Real & Caraco, 1986). For example, during the day, warm-blooded animals pursue *safe* but small sized prey items – prey items that they probably will be successful in obtaining but which have a low food value. However, they may pursue *riskier* but higher value choices as evening approaches and *foraging opportunities* for the day decrease. This is particularly the case if their metabolic *resources* are low or if they need to gather enough food to meet a metabolic *target* to survive a cold night. In such circumstances, pursuing a *safe option* associated with a probable but small magnitude of food is, by evening, of little long term value because it will not be sufficient to guarantee the animal's survival through the night. Instead the animal should be biased towards *riskier options* associated with high magnitudes of food items even if they have a lower probability of success. The animal's attitude to probabilities is therefore, a function of its momentary resource budget and its longer term targets.

From this ecological perspective decisions are viewed as occurring within sequences and there is the possibility of adapting the decision-making strategy later within such a sequence depending on the outcomes of initial decisions. Therefore, when navigating an environment via a sequence of decisions, riskier choices can be seen as part of a particular strategy, influenced by past experiences and future prospects.

The first aim of the study was to test whether dynamic changes in decision strategy occur in humans as they make a series of decisions and to see whether they depended on a person's current resources as well as longer-term targets. Thus, while our approach borrows from RSFT, it differs from most accounts of "risk" and gambling prevalent in cognitive neuroscience because it recognizes that different use of probability – effectively different decision-making strategies – may be optimal in different contexts. Some contexts, such as the prospect of the cold night for the foraging animal described above, can be thought of as exerting a pressure to assume a more risky decision-making strategy. We refer to this contextual influence as *risk pressure.* Note, we use the term "risk" in the sense it is most commonly used, to refer to a choice's probability to yield no gain or a loss, rather than in the way it is sometimes used in neuroeconomics, to refer to the outcome variance of a choice (Michael L Platt & Huettel, 2008; Preuschoff, Bossaerts, & Quartz, 2006; Matthew F S Rushworth & Behrens, 2008).

We simulated the situation of a foraging animal pursuing an imperative longer term reward *target* by asking subjects to try and repeatedly collect a *target* number of points over a mini-block of eight decision trials. Subjects chose between safer "high probability" options and riskier but high magnitude options (Figure 1A). Reaching the *target* meant subjects kept the points they had won in that block but failure to reach the target meant all points in that eight-trial block were lost. Thus, the subjects accumulated *resources* in terms of points and with every decision their *foraging opportunities*, in terms of trials left in the block, decreased. Importantly, the safer choice normally had, on average, a higher value [on six out of eight times it had the higher expected value (probability x magnitude) and, in general, it was preferred by participants]. However, if a subject takes into account the sequential structure of the task as well as the contextual factors i.e. the target level, their current level of resources and the number of trials left, then it should motivate them to take the riskier choice instead. This is because, even if it is successful, the safer choice sometimes yields insufficient points to reach the target.

Our analysis focused on relating decisions and brain activity recorded with functional magnetic resonance imaging (fMRI) to two types of variables. The first type concerned specific decisions that participants made and the choice values that motivated those decisions. This part of the analysis often concerned the relative values of riskier and safer choices ($V_{riskier}$-$V_{safer}$). In the past, relative value signals have been used to identify neural mechanisms of decision-making (Boorman et al., 2009; Camille et al., 2011; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; N. Kolling et al., 2012; Lim et al., 2011; M P Noonan et al., 2010; Philiastides et al., 2010; Wunderlich et al., 2012). The second type of variable focused on the gradually changing context as participants moved through the block. For this we estimated three key parametrically varying quantities. First, trial number indexed how far through the block the subject had progressed. Second, *risk pressure*, was the difference between the subject's current *resources* and the imperative *target* scaled by the remaining *foraging opportunities* (equation 1; Figure 1B). *Risk pressure* should lead to a contextual modification of the options' values. Using a model, we formalized the amount of optimal modification in a given trial through the third key term – *risk bonus* (equation 6) – the degree to which *risk pressure* should optimally bias a person away from the safer choice, given the current offers' magnitudes and probabilities, as well as future decision opportunities.

# Methods

## Subjects

Eighteen subjects (nine female), aged 22-36 years, completed the task. They were paid £10 plus a performance-dependent bonus of between £15-30. Ethical approval was given by the Oxfordshire National Health Service Research Ethics Committee (07/Q1603/11).

**Training**: Before fMRI scanning every subject was instructed in the task and played a shorter version of the task used in the fMRI experiment for about ten minutes.

## Experimental Task

The behavioral task in the scanner consisted of 24 blocks each composed of eight trials (192 decisions in total), in which the subjects had to decide between a safer option with a higher reward probability but a lower reward magnitude and a riskier option with a potentially higher reward magnitude but lower reward probability. There were eight decisions and they were each presented once in each block in a randomized order that varied. In this way we were conclusively able to show that dynamic changes in decisions occur, because of sensitivity to *risk pressure*, even when the exact same options were presented. *Risk pressure* varied because all eight decisions were associated with different values and were presented in different orders, with different outcomes and in the context of different block target values (which the subjects had to reach in order to keep the points they won during the block).

Four target levels were used in the experiments. The different target levels helped ensure that *risk pressure* (see Introduction and the next section) had some parametric range. To equalize expected gains at the beginning of a block regardless of target level and to keep motivation relatively stable, we introduced a "multiplier" which was displayed on top of the "target" line. The multiplier indicated a factor that would be used to multiply the points subjects won before they were added to the subject's account if they reached the target. We chose the multiplication factor by applying our model (discussed in the next section) to generate equal expected gains at the first trial of a block. Simply put, if a participant had a high target to reach, all his points were multiplied (e.g. by 2), if they managed to reach it. Therefore, the subject should be equally motivated to perform the task when the targets were high because the average payouts were similar. The use of the "multiplier" procedure ensured that any neural *risk pressure* signals that we observed in the next chapter could not be explained away as a consequence of differing average reward expectations associated with different target levels.

More specifically, to induce variability in choices, we used four different target levels (595, 930, 1035, 1105 points) and a block length of eight trials [the pairs of options associated with each decision each had a reward magnitude in points, and a reward probability (indicated as percentage): 100p x 90% vs 265p x 35%; 180p x 60% vs 260p x 35%; 145p x 75% vs 245p x 35%; 145p x 55% vs 350p x 20%; 115p x 90% vs 240p x 45%; 150p x 60% vs 190p x 45%; 170p x 75% vs 245p x 40%; 120p x 80% vs 210p x 30%]. With increasing target level the overall chances of exceeding the required points in a block decreased. In parallel, the optimal number of riskier choices, as predicted by our model, increased. The target levels used led to a good spread in the number of optimal riskier decisions (1-4 riskier decisions) predicted at the beginning of a block and thus to a good spread of the model parameter "optimal *risk bonus scaling*".

The multipliers for the four target levels were: target level of 595 points – 1.1 multiplier; target level of 930 points – 2.3 multiplier; target level of 1035 points – 3.3 multiplier, 1105 points – 4.2 multiplier. Because we presented the same eight decisions in all blocks variability due to features of the choices per se was reduced. Additionally we decorrelated the decision parameters (reward probability, reward magnitude, value of option, difference between values) as much as possible to be able to use them as separate regressors simultaneously in the fMRI analysis. Moreover, the Pascalian values (probability x magnitude) of six of the eight decisions favored the safer choice. This meant that because variation in the target parameter drove decisions towards the riskier choice, participants' behavior would exhibit an approximate balance between both safer and riskier choice decisions, which ensured sufficient sample sizes of both decision types for fMRI analysis.

The behavioural task during the fMRI scan took approximately 55 minutes and was followed by a high-resolution structural scan and a field map acquisition for distortion adjustments. After scanning every subject completed a brief subjective rating questionnaire on the 'riskiness' of the eight decisions that had been presented in the fMRI task.

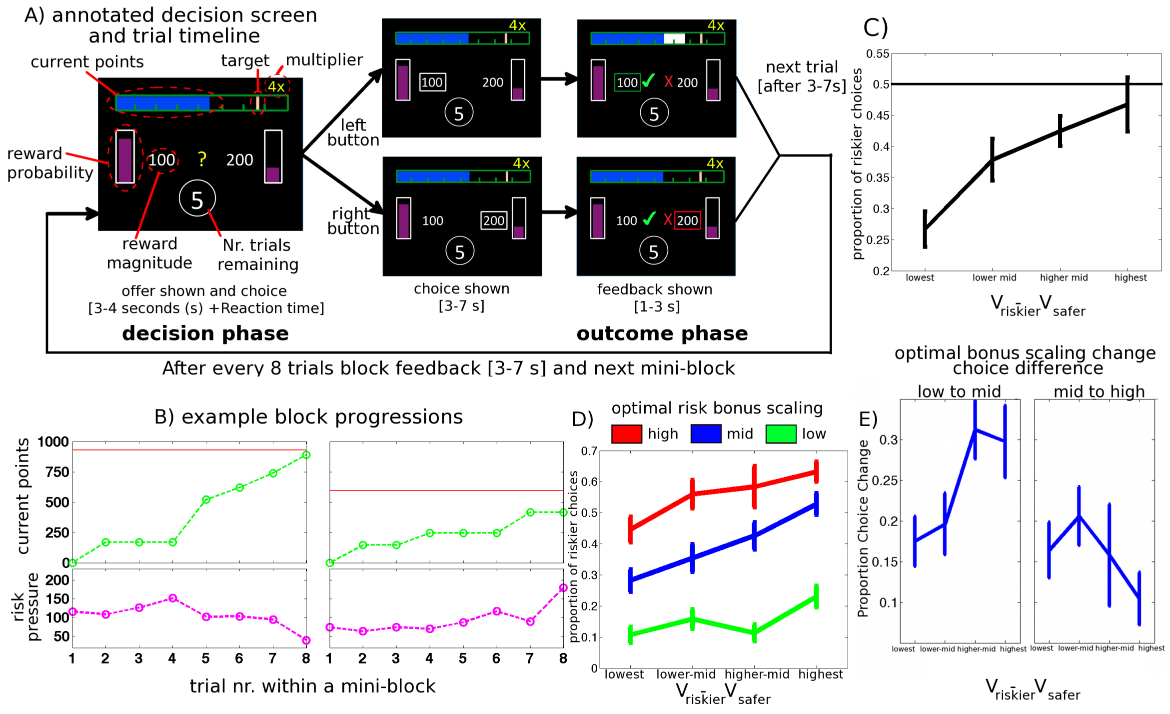

***Figure 1:A)*** *Trial timeline: At the start of trials subjects were presented with choices on the left and right of the screen. Each option was composed of a reward probability (height of purple bar fill) and magnitude (number next to each bar). This was followed by a choice cue (a yellow question mark) that instructed subjects to choose. Subjects chose between a more probable low magnitude option (safer option) and a less probable high magnitude option (riskier option) on each trial. After responding their choice was highlighted with a white frame and feedback was shown for both options (both the chosen option and the alternative). If the choice was rewarded then points were added to the blue "current points" bar at the top of the screen (white bar indicated added points), progressing it further towards the target. The number of trials remaining in the block was indicated in the circle at the bottom of the screen. After each trial the number of remaining trials was reduced by one. The target turned white if it was reached.* ***B)*** *Two examples of progressions through a mini-block. Points accumulated are shown in green with target level in red (upper panel) and the resulting* risk pressure *in magenta (lower panel). In the first example the target was relatively high and the* risk pressure *is highest before a big*

*win after the 4th decision, when the subject selected the less likely but more valuable riskier option. In the second example the pressure is lower at first but increases after a series of losses, until it actually exceeds the* risk pressure *experienced in the other block.* ***C)*** *Overall proportion of riskier choices as a function of increasing relative value of riskier choice ($V_{riskier}$-$V_{safer}$).* ***D)*** *Overall proportion of riskier choices split by optimal risk bonus scaling and binned by increasing relative value of the riskier choice ($V_{riskier}$-$V_{safer}$).* ***E)*** *(left) Differences in proportion of riskier choices between low and mid (green and blue in panel D) optimal risk bonus scaling and (right) between mid and high (blue and red in panel D)illustrating how changes in optimal risk bonus-scaling are associated with increased frequencies of riskier choices. Additionally, for the first change, choices with large $V_{riskier}$-$V_{safer}$ are affected more, whereas for the second change the more difficult decisions involving lower $V_{riskier}$-$V_{safer}$ are more affected. All errorbars are mean±SEM.*

## Behavioral analysis

As explained in the introduction, the target manipulation and block structure allowed us to compute a series of contextual variables. Trial number indexed how far through the block the subject had progressed. The second variable, *risk pressure* indexed how many points a subject needed to gain on average on each remaining trial in order to reach the target. *Risk pressure* thus took into consideration the subject's *resources* (the points they had earned prior to any given decision), the *target* number of points that had to be acquired in order to keep any earnings from the block, and the number of remaining foraging opportunities – the number of trials that remained in the block.

1) Risk pressure = (target points – points already earned)/trials remaining in block

In order to understand how *risk pressure* exerted an influence of decision-making it is first necessary to consider the relative value of riskier and safer options in the absence of any contextual modification. $V_{riskier}$-$V_{safer}$, the value difference favoring riskier as opposed to safer choices, was calculated as follows:

2i) $V_{riskier} = \text{normalised}(\text{magnitude}_{riskier}) + \text{normalised}(\text{probability}_{riskier})$

2ii)$V_{safer} = \text{normalised}(\text{magnitude}_{safer}) + \text{normalised}(\text{probability}_{safer})$

This is because I noticed that subjects acted as if they approximated $V_{riskier}$ and $V_{safer}$ by linearly combining each option's component magnitude and probability rather than multiplying them as would be optimal. This is apparent when the proportion of riskier choices made as a function of the relative value of the riskier choice ($V_{riskier}$-$V_{safer}$) is plotted when values are computed as a linear approximation or as the true product of the component reward magnitudes (Figure 2) in trials with little or no optimal risk bonus scaling according to the model.

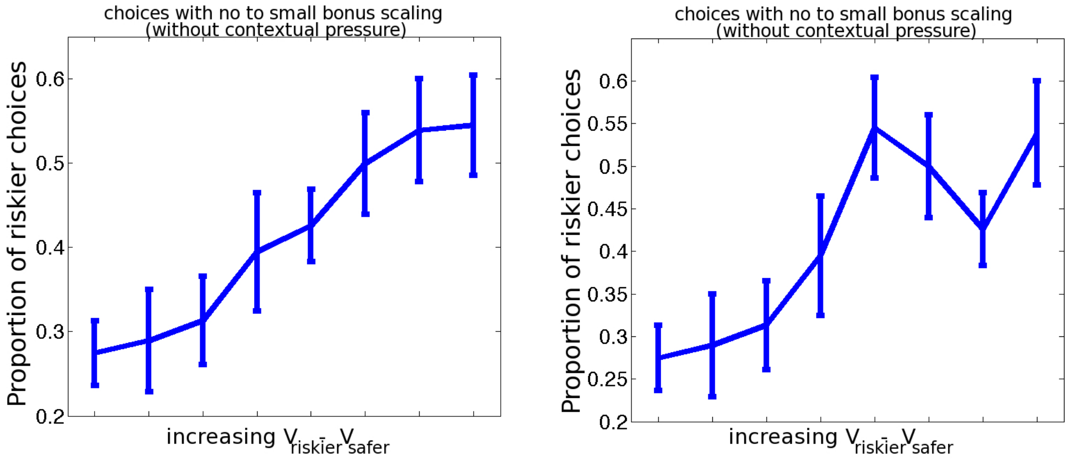

***Figure 2:*** *Proportion of riskier choices taken as a function of relative value of the riskier choice ($V_{riskier}$-$V_{safer}$) in trials with little or no optimal risk bonus scaling (blue line in figure*

*1D). $V_{riskier}$ and $V_{safer}$ were calculated as either the sum (left) or the product (right) of the options' reward magnitudes and probabilities.*

Nevertheless there was a correlation ($r>0.86$) between the value regressors we used and those we would have used had value been estimated multiplicatively. Note that both parameters (magnitudes and probability) were, separately, normalized by subtracting each mean and dividing by each standard deviation. Finally, importantly while I follow convention in referring to these terms as "values" it is of course the case that these values are inferred from subjects' choices. They are therefore likely to be predictive of choices but the question I investigate here is whether they are *sufficient*, in isolation, to explain choices or whether other contextual factors also influence decisions.

## Behavioural model

I, therefore, built a model examining the process of value modification due to contextual factors such as *risk pressure*. At its heart is the idea that, in the absence of *risk pressure*, it is optimal to combine information about both the probability and magnitude of a reward outcome associated with a choice but that with increasing *risk pressure* decision-making should be guided increasingly by just the potential reward magnitudes at stake. Although, I am not wedded to the precise parametrization of the model, the general aim of the approach is to find a principled and quantified way of modifying the decision rule, going from the unmodified decision rule that combines both reward probability and magnitude to a rule based exclusively on magnitudes. In the model we use a parameter –*risk bonus scale* – that moves from "0" to "1" as the decision rule is changed from the unmodified version to the increasingly contextually modified version. Such adjustments of a decision rule provide an intuitive way to think about how an agent adjusts their behavior in a new situation. The contextual parameter *risk bonus scale* therefore captured the insight that participants should opt for the riskier choice, even if its associated reward probability was low, if it was going to be difficult for them to reach the block's target level in the absence of that reward. At a *risk bonus scale* of 0 there is no modification of the option values shown in equation 2. At a *risk bonus scale* of 1, the options' values corresponded solely to their magnitudes.

The changes in the options' values were formalized by adding an *option bonus* to each option's raw value. This allowed estimation of a simple quantity that corresponded to how much an option's value increased for a given level of *risk pressure*. The size of the *option bonus* depended on both i) the *risk pressure* on a given trial but also on ii) the specific raw value of the option. The dependence on the specific raw value that each option possesses follows from the fact that high reward magnitude options, even when associated with low probabilities, have greater utility for reaching the target at the end of the decision sequence. The *option bonus* for a specific option A is calculated as:

3) $\text{option bonus}_A = \text{risk bonus scale} \times (\text{magnitude}_A - \text{magnitude}_A \times \text{probability}_A)$

The term in parenthesis on the right hand side of equation 3 can be thought of as an option specific component of the option bonus. It is the difference between the number of points that could potentially be gained from that option (its magnitude) and the average points expected from that option (magnitude x probability; note that the product of magnitude and probability corresponds to the average value of the options under this optimal model). We used the option bonus to calculate modified model values of the options:

4) modified model $\text{value}_A$= ($M_A$ x $P_A$) + option $\text{bonus}_A$

Where $M_A$ and $P_A$ correspond to the magnitude and probability of rewards associated with option A. Alternatively the modified value option A can be written:

modified model $\text{value}_A$ = ($M_A$x$P_A$)+ risk bonus scale*($M_A$ – ($M_A$x$P_A$))

or again alternatively as:

modified model $\text{value}_A$ = ($M_A$ x $P_A$) + (risk bonus scale x ($M_A$ x (1 - $P_A$ )))

And decisions should be made as follows:

**If** (modified model $\text{value}_A$ > modified model $\text{value}_B$) **then** chose A

**Else if** (modified model value < modified model $\text{value}_B$) **then** chose B

So far I have explained how *risk bonus scale* was used in conjunction with the option's reward probability and magnitude to estimate an *option bonus* for each option. It is necessary now to explain how the *optimal risk bonus scaling* itself was calculated. I simulated for every trial all unique decision sequences each associated with a different *risk bonus scale* by calculating their modified values and using the above decision rule (Figure 3). For every unique decision sequence, generated with the value modification model, I could compute an end of block expected value. The *optimal risk bonus scaling* was defined as the *risk bonus scale*, which led to the decision sequence with the highest end of block value. Importantly, when doing so, I took into account that all net outcomes that fell short of the target value had a value of zero. It is not necessary to assume that participants were able to track the exact *optimal risk bonus scaling*, for it to serve as an approximation of how the values of specific choices should be modified as a result of the context on a given trial. Task parameters were chosen to maximize the parametric range of such modification.

It is furthermore possible to calculate the risk bonus scale that leads to the point of equivalence for a given pair of options. In other words at an optimal risk bonus scaling equal or above this value for an option pair, the riskier option should be preferred:

5) $\text{equivalence risk bonus\_scale}=(M_S x P_S - M_R x P_R) / (M_R x(1-P_R) - M_S x(1-P_S))$

OR

$\text{equivalence risk bonus\_scale}=(M_S x P_S - M_R x P_R) / ( (M_R - M_S) - (M_R x P_R - M_S x P_S) )$

Where $M_R$, $M_S$, $P_R$, and $P_S$ refer to the reward magnitudes associated with the riskier and safer options and reward probabilities associated with the riskier and safer options respectively.

By computing this value for all remaining decisions and rank ordering decisions from the least to the most risky we could estimate the value of all unique decision sequences and select the one that led to the highest end of block value. In all neural and behavioral analyses the *risk bonus scale* used is therefore equal to the *optimal risk bonus scaling* in a given trial, i.e. the *risk bonus scale* that generates a sequence of future decisions that would lead to the highest expected value at the end of the block, taking into account the current context (*risk pressure*) and future prospects (set of options left and the pair presented).

The *optimal risk bonus scaling* is, therefore, a contextual parameter reflecting the degree of bias towards riskier choices that is optimal for a given context and applies to both options in a trial in the same way. The option bonus becomes larger for riskier choices compared to safer choices as the *optimal risk bonus scaling* increases, reflecting the riskier choices' increased utility for reaching the target. Therefore, the option bonus can be understood as a combination of *risk pressure* to take a riskier choice and features of the specific option at hand.

Finally, I used one more parameter I refer to as *risk bonus* (as distinct from *optimal risk bonus scaling*) that was used in neural and behavioral analyses. This was the differences in value modification in favor of the riskier choice compared to the safer choice. It was calculated using the *optimal risk bonus scaling* as:

6) $\text{Risk bonus} = \text{option bonus}_{\text{riskier}} - \text{option bonus}_{\text{safer}}$

Therefore, *risk bonus* reflects the relative change in value of the riskier choice compared to the safer choice that occurs as a function of *risk pressure* and the magnitude and probability characteristics of both choices in a given trial. It should be noted that in this regard the model is an optimal model that serves to motivate definitions of terms but that real subjects may not be completely optimal. For example, if instead *option bonuses* were only adjusted as function of their reward magnitudes (rather than as a function of both reward magnitudes and probabilities – equation 3) then the resulting *risk bonus* regressor would be correlated at r=.96 with the regressor used.

## Illustration of model

As mentioned above, figure 3B-D depicts how the optimal bonus scale was determined based on the decision sequences for an example trial. Figure 5A depicts the relationship between the optimal bonus scale and speculation pressure. Although the optimal bonus scale has a strong relationship with the speculation pressure, the relationship is non-linear; variations in the speculation pressure do not always lead to a change in optimal bonus scale, for example several levels of speculation pressure are associated with an optimal bonus scale of zero indicating that there should be no contextual modification of the options' values. Small changes in speculation pressure, when speculation pressure is already high, lead to large increases in optimal bonus scale.

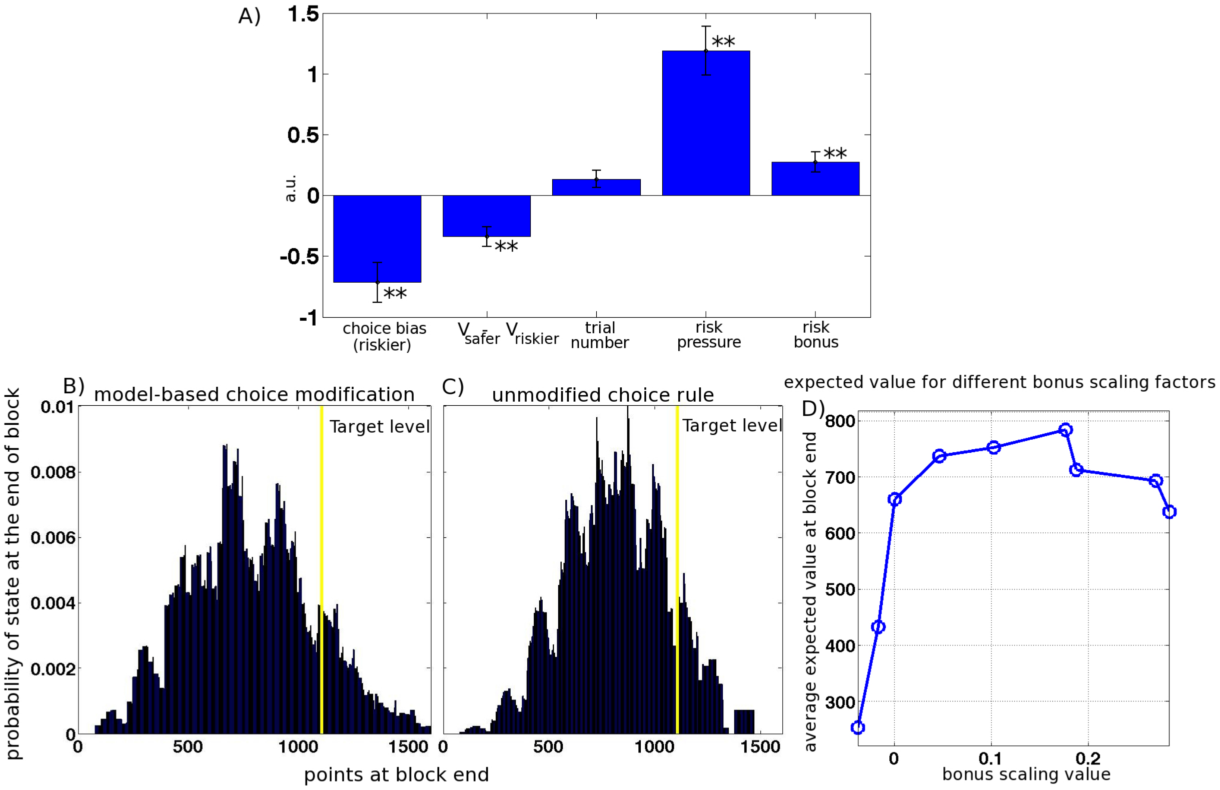

***Figure 3:*** *A) Behavioral general linear model including both risk pressure and risk bonus. Although much of the behavioral variance is captured by including risk pressure as a regressor, risk bonus has a further and significant influence on behavior. (** p<0.01). Illustration of the results of decision sequences through the rest of a mini-block and calculation of optimal risk bonus scaling for an example trial. B and C) Histograms showing the probability distribution of block outcomes for two specific decision sequences (for illustrative purposes with a 20 point smoothing to represent uncertainty about precise final state). We considered only the part of the distribution that falls above the target level to calculate the expected value of a decision sequence, multiplied with a constant dependent on target level. The histogram in B) is based on the decision sequence with the maximal expected value for that trial, and therefore the one with the optimal risk bonus scaling, while the histogram in C) is based on the decision sequence suggested by the raw values of the choices. Note the difference in distribution ratio that falls to the right of the yellow target line. D) A plot showing the expected values of all unique decision*

*sequences suggested by different risk bonus scaling values for a given trial. In this example a risk bonus scaling of around 0.18 is optimal.*

The value difference term ($V_{safer}$-$V_{riskier}$) is the fixed difference in preference between the safer and riskier option and *trial number* refers to the current trial's position in a block. The *risk pressure* term summarizes how the environmental context should lead to riskier choices. The *risk bonus* term, however, encapsulates a model-based interaction of both the contextual factor of risk pressure and features of the specific choice options available on the current trial. It predicts how participants should take risks that are dependent on how much the risk pressure on a trial should lead to a difference in evaluation of the specific current options also taking into account future possibilities for taking such risks. It can be seen that it explains additional variance in behavior over and above that explained by risk pressure. Another illustration of this can be seen in figure 1 E.

## Optimality of the model

The model is very nearly optimal. The decisions taken by our model after taking into account risk pressure and by a truly optimal model differ in less than 7% of cases. More importantly the divergences in expected value at the end of the block that would follow from taking a sequence of actions according to our risk sensitive model or according to a truly optimal model are very small; the expected values of the choices our model generates, after taking into account risk pressure, are highly correlated (r>0.98) with the expected values of the choices generated by a truly optimal model (in which all possible decision sequences are computed and compared according to their expected values). The reason we use our model and do not simply take the completely optimal decision sequence, is that our model offers a simple account of how context might lead to a simple modification of a choice's value that might in turn lead to a change in the decision that is taken.

The reason a small number of minor divergences occur between our model and the completely optimal decision sequence has to do with the hard target constraint; in some cases very specific outcome combinations add up to generate a slightly higher overall expected value than the model. This is because the model is constrained to envisage that the level of risk-taking in the rest of the block will be consistent with the level of risk-taking in the current trial when it computes the level of risk-taking that is appropriate in the current trial. Thus it computes a uniform level of risk-taking for the entire sequence of decisions that remain before the block end. By contrast an optimal model is not constrained by any concept of risk taking and simply finds the decision sequence that gives the highest expected value. In this way a small number of specific decision sequences become very slightly more optimal because they lead to an expected value that just exceeds the target. However, because the gain in expected value is rather minimal and can most likely only be computed with a very precise knowledge of all specific magnitudes and probability in the decision sequence we thought that it was unlikely that participants could make use of this information. Moreover I found no empirical evidence that they could do so. Indeed, on average, for the 7% of decisions in which the optimal sequence and our model diverged from each other, about half were made in accordance with our model and half in accordance with the optimal sequence.

In summary, the approach allows me to: 1) examine decision-making in the context of the varying impact of *risk pressure*; 2) conceive the impact of *risk pressure* as a quantifiable modifying influence on a default decision-making process. However, I explore an alternative approach as well that considers how an agent with sufficient experience of a set of contexts may use a reinforcement learning model to estimate the values of choices. A number of links between the approaches are identified and discussed.

## Reinforcement learning approach

An alternative approach to the task is a reinforcement learning one. Such an analysis has the appeal that it aligns naturally with other past studies using state-based reinforcement learning to explain neural data and behavioral performance.

It would be feasible to use a reinforcement learning approach to analyze the behavior recorded in our paradigm. Such a model approach might represent all combinations of the current number of points acquired and of trial number and it would estimates expected overall block end values separately for each target level. However, when related reinforcement learning-based approaches have been used in the past they have typically not incorporated knowledge of which option pairs have actually been observed in past decisions; they therefore incorporated no expectations about how options might be paired in future decisions or knowledge of the specific sequence of decisions in a block). In order to incorporate this information it would be necessary additionally to hold independent state space representations for every possible combination of sequences.

I therefore used a state space-based reinforcement learning model of this type. The starting point of the model is to estimate the values of all possible block end states taking into account the four different target levels, i.e. making all below target values zero. I subsequently computed the value of all earlier states in blocks, one trial back at a time, giving equal likelihoods of all option pairing appearing in any decision, by looking up the values of the later states that would follow on from them as a function of the state space map. Furthermore, I made the assumption that in every decision the option that would lead to the higher next state would be selected. Doing so from the last to the first trial, it is possible to give the current state a value defined by the state (trial number and point count) as estimated from the model and also find such a state estimate for each the two options if they are chosen in the decision. Since the future state is different for losses and wins, I need to compute two future state space values for each option. The two future state values are then averaged by using the probabilities of winning for each option i.e. by multiplying the value of each state with its probability, which is the expected future state value for both available options, The difference between both options' future state values is the state value difference and, according to such an approach, should be the decision variable. The model assumed equally likely redraws of all eight trial types for every future trial.

I then compared several key value estimates and decision variables using this state-based reinforcement learning model with those derived from our risk pressure-risk bonus model which is explained in Methods and above. In order to make the comparison fair, I used a slightly modified version of our risk pressure-risk bonus model, which also assumed equally likely redraws of all eight trial types for each future decision (r=0.79 between future block end value differences). I estimated:

1) ***Future state space value difference/future block end value difference:*** The future state space value for each of the two options on each decision under the reinforcement learning model and I then determined the difference between the future state space values. I then found the comparable term using our risk pressure-risk bonus based model by determining the expected value at the end of the block for each choice that might be taken in a decision and then

determining the difference between the expected values for each pair of options offered in a decision. There was a high correlation between reinforcement learning-based future state space value differences and risk pressure-risk bonus model based future block value differences (r=0.95).

***2) Overall expected value at the end of the block:*** The overall expected value at the end of the block estimated under the reinforcement learning model and risk pressure-risk bonus model were highly correlated (r=0.97).

***3) Q-action values and option values for specific choices:*** The Q-action values associated specific options under the reinforcement learning approach and the option values used by the risk pressure-risk bonus model were highly correlated (r=0.97).

Thus both approaches generate almost equivalent estimates of ultimate or long term value expectations at the block end and something like such estimates were found to be represented in the dACC (see next chapter Figure 4). However, only the risk pressure-risk bonus approach also has parameters relating to risk pressure and the risk-based value modification of option values in a specific decision and such terms were useful for describing both behavior and the pattern of activity recorded in dACC, vmPFC,PCC, and IFG (see next chapter).

## Results

I tested whether the frequency of riskier choices was simply driven by $V_{riskier}$-$V_{safer}$ or whether it also reflected the *risk pressure* associated with the context in which the decision occurred using a logistic regression analysis (Methods: Behavioral Analysis).

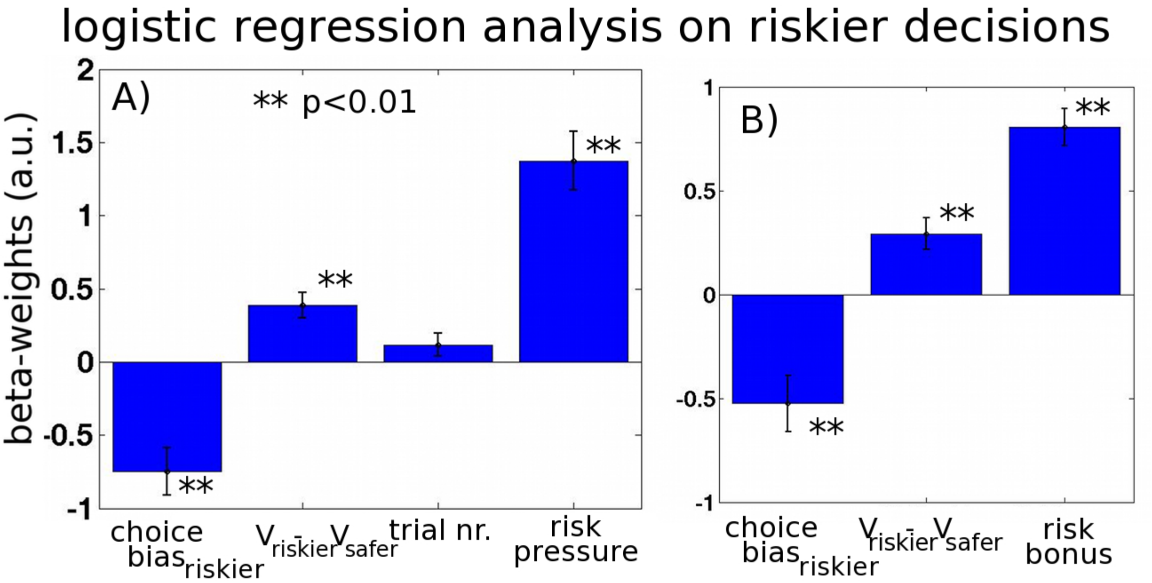

***Figure 4:*** *Logistic regression regressing riskier choices against parameters defining each decision.* ***A)*** *GLM incorporating relative value of riskier choice ($V_{riskier}$-$V_{safer}$), number of trials already performed in the current block (trial nr.) and* risk pressure*. Increases in both $V_{riskier}$-$V_{safer}$ and* risk pressure *were associated with significant increased riskier choices. The constant term from the GLM, however, indicates a bias against riskier choices (left hand bar with negative value).* ***B)*** *An alternative analysis used risk bonus (reflecting the model-based impact of the current* risk pressure *on $V_{riskier}$-$V_{safer}$) and again increases in this term were associated with significant increases in riskier choices. All errorbars are mean±SEM.*

$V_{riskier}$-$V_{safer}$ exerted a significant influence ($p<0.001$, $t(17)= 4.48$) but this is obviously expected given that the estimates of the subjects' values are based on their choices (equation 2). What is important to note, however, is that it was not *sufficient* to explain choices; *risk pressure* exerted an additional effect ($p<0.001$,$t(17)= 6.88$) (Figure 4A). An alternative logistic regression looked at riskier choices as a function of the *risk bonus* on each trial (this term expresses how the relative value of the riskier option as opposed to the safer option changes as a function of *risk pressure* and the options specific magnitudes and probabilities; equation 5). The risk bonus on a trial exerted a significant impact on riskier choice frequency ($p<0.001$, $t(17)= 9.03$; Figure 4B). Note that both analyses included a negative constant term (negative-going bar on left side of figures 4a and 4b, in form of the intercept of the regression model) meaning that subjects were biased against riskier choices and their default approach was to take safer choices although they do so less when $V_{riskier}$-$V_{safer}$ or *risk pressure* is higher. As an additional control analysis a GLM with both risk pressure and risk bonus reveals that both regressors explain unique behavioral variance (Figure 3).

To illustrate the results of the GLM's I binned the behavioural data in several different ways. Firstly, subjects had a baseline tendency towards risk aversion but they took more risky choices as *risk pressure* increased. This is apparent when trials are binned into four levels according to $V_{riskier}$-$V_{safer}$ and the frequency of riskier choices are plotted (Figure 1C). Overall, participants were less likely to take riskier choices at all levels of $V_{riskier}$-$V_{safer}$ but the effect was smaller when $V_{riskier}$-$V_{safer}$ was larger.

Secondly, riskier choice frequency can be examined not just as a function of $V_{riskier}$-$V_{safer}$ but also as a function of the *optimal risk bonus scaling*; *optimal risk bonus scaling* is one of the parameters derived from our model that expresses the approximately optimal degree to which participants should be biased towards riskier choices as *risk pressure* increases independent from the specific options presented in the trial (Figure 1D). Three equally sized bins of trials were created using the *optimal risk bonus scaling* factor for a trial. Within each level of *optimal risk bonus scaling* I examined the effect of $V_{riskier}$-$V_{safer}$. Participants took more risky choices when $V_{riskier}$-$V_{safer}$ was larger even when the *optimal risk bonus scaling* was lowest. On trials with little or no *optimal risk bonus scaling*, participants did not, on average, prefer riskier choices, even when $V_{riskier}$-$V_{safer}$ was high (there was no significant preference with a one-tailed t-test against 0.5, see Figure 2). However, participants began taking more risky choices even when $V_{riskier}$-$V_{safer}$ was in the lower mid range when *optimal risk bonus scaling* was high. A change in *optimal risk bonus-scaling* from low to mid levels (Figure 1E, left) and from mid to high levels (Figure 1E, right) is associated with an increased frequency of taking riskier choices. In the first case decisions with large $V_{riskier}$-$V_{safer}$ are affected whereas in the second case the more difficult decisions involving lower $V_{riskier}$-$V_{safer}$ are more affected.

Another way to examine how participants shifted away from a baseline tendency to risk aversion is to compare their behavior to the predictions of our model which, as already noted, makes decisions that are close to optimal. Participants were more likely to make model-conforming safer choices than they were to make model-conforming riskier choices (Figure 5B). However, riskier choices were still more likely than not to conform to model predictions. This means that even though participants were not completely optimal, they integrated over choice value and contextual factors in a way predicted by our model, with a slight overall bias against the riskier option..

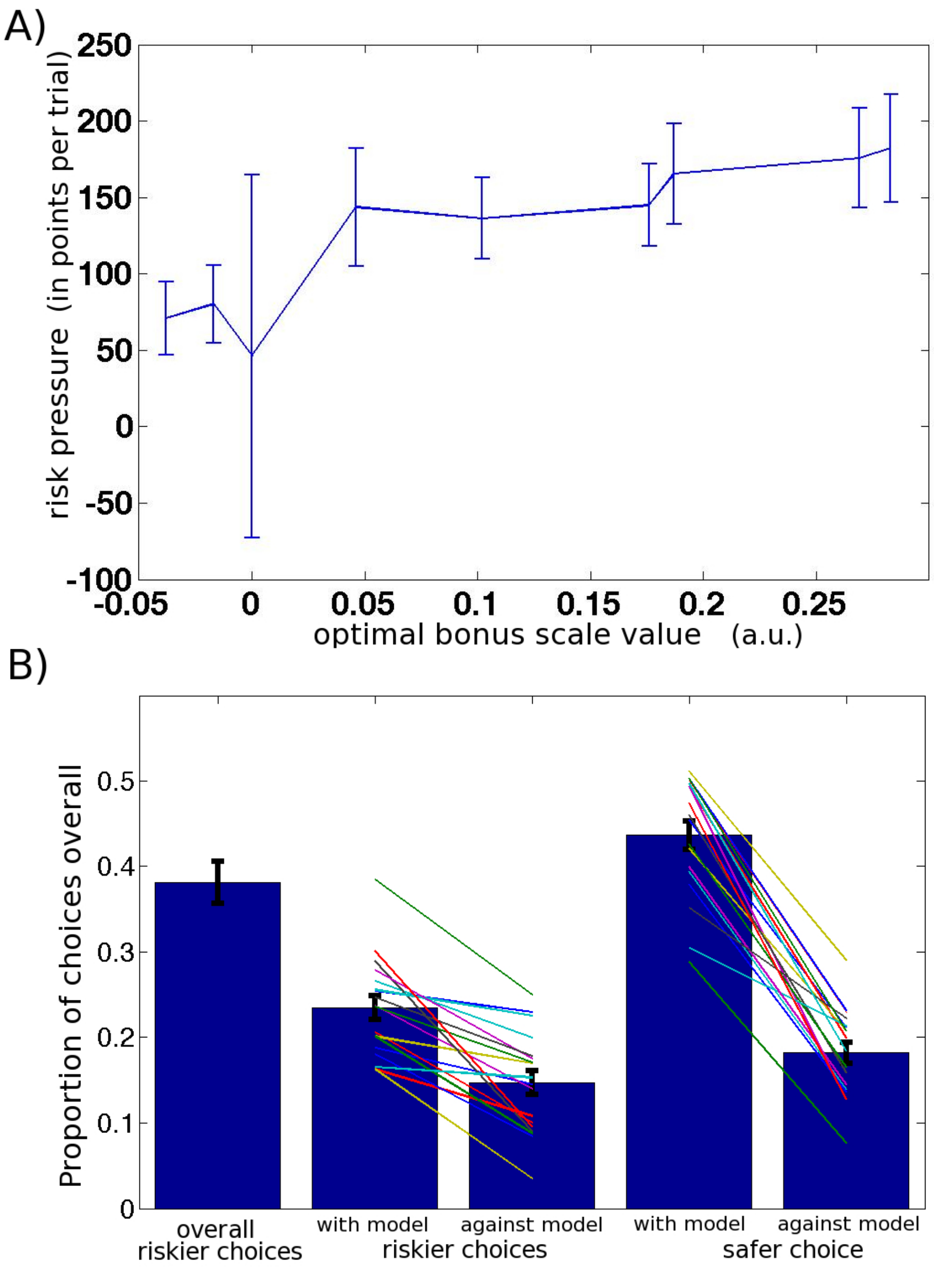

*__Figure 5:__ A) The relationship between a trial's risk pressure and the optimal risk bonus scaling (as determined using the maximal expected value at the end of a block depicted in Figure 3). There is a clear relationship between both quantities although it is non-linear and not one-to-one in nature. A number of risk pressure levels entail no need to move towards a more risk-oriented decision strategy, whereas relatively small variations in high risk pressures should lead to quite large changes in the optimal risk bonus scaling and therefore predicted changes towards a more risk-oriented decision strategy. B) Overall proportion of riskier choices as well as a breakdown of decisions into four categories. Model conforming / opposing riskier and safer (lines indicate individual*

*subjects).*

# Discussion

## Dynamic changes in the use of probability

Instead of assuming attitudes to probabilities reflect stable individual differences, a behavioral ecological approach to decision-making suggests animals should adapt decision-making strategies as a function of their current resources, resource targets, and the opportunities that remain for foraging (Caraco, 1981; Hayden et al., 2009; Kacelnik & Bateson, 1997; McNamara & Houston, 1992; Real & Caraco, 1986). I argued that these factors can be integrated to determine the current *risk pressure* – the degree to which it might be adaptive to adjust decision-making towards pursuit of low probability but potentially large reward magnitude outcomes. The combination of *risk pressure* with the precise values of the specific options that might be chosen in a given decision determine a *risk bonus* – an increase in value that accrues to the low probability but potentially large magnitude option in a decision. I designed a decision-making task for humans (Figure 1) that manipulated these factors, changing resource levels, target levels, and opportunities for further foraging. Human subjects were sensitive to *risk pressure* and the risk bonus; increases in each factor lead to more frequent riskier choices (Figure 1, 4). Although adding a risk bonus to the values of choices made in the context of *risk pressure* provides an intuitive way to think about how decision-making strategies can be rapidly updated there are, nevertheless, links between several of the concepts used in my approach and those that can be derived from a reinforcement learning-based approach.

There is also a link to previous studies that have shown subjects often have biases towards certain decisions and that activity in some brain regions is associated with taking decisions that do not conform with the default strategy (Venkatraman, Payne, Bettman, Luce, & Huettel, 2009; Venkatraman, Rosati, Taren, & Huettel, 2009).

Individual differences in risk-taking behavior may, in extreme cases, be associated with pathological gambling (Clark & Limbrick-Oldfield, 2013). While pathological gambling may be linked with a baseline change in risk proneness/aversion my results raise the possibility of a link with individual differences in how decisions are influenced by context. An approach focusing on changing sensitivity to contextual factors such as *risk pressure* may elucidate aspects of developmental change in risky behavior (Blakemore & Robbins, 2012; Paulsen, Platt, Huettel, & Brannon, 2012). Assaying response strategies with low likelihoods of success but with the potential for delivering great gains may be imperative at some points in adolescence.

Overall it is not too surprising that participants were able to modulate their risk taking behaviour by taking their current environment into account, as this is a very important form of behavioural flexibility ecologically (Caraco, 1981; Kacelnik & Bateson, 1997; McNamara & Houston, 1992; Real & Caraco, 1986). However, it is a significant departure from how risk taking is thought of generally within decision neuroscience. Normally, risk preferences are thought of as static or changing between different forms of frames of biases rather than rapidly as a function of current constraints. Investigating risk taking in a dynamic context with quantitative modifications of the evaluative processes required a novel kind of behavioural model that allowed me to index the environmental changes and longer term values as well as resulting modifications of decision strategy for a given decision between two options. Validating such a novel model behaviourally is important when using it for understanding the neural mechanisms the model is meant to approximate.

Having focused on the behavioural models of adaptive decisions under changing risk pressure in this chapter, in the next one, I will focus on its neural correlates. More, specifically I will test what brain regions are sensitive to the continuous tracking of changing context that in turn effected evaluation of specific choices. Additionally, I can measure what effect this changed evaluation has on the neural correlates of decision-making generally and the network properties of regions known to play roles in reward and decision-making, more specifically.

# Chapter 5: Multiple neural mechanisms of decision-making and their competition under changing risk pressure

## Abstract

*The ability to take risks flexibly is essential for survival. As discussed in the last chapter, optimal foraging theory as tried to address the influence of context on risk taking in animals. However, the neural mechanisms of the ability to dynamically and contextually assess risks and payoffs are largely unknown. Using my paradigm which induced behavioural flexibility in risk taking, I can demonstrate that dorsal anterior cingulate cortex (dACC) in humans carries signals indexing the pressure to pursue unlikely choices and signals related to the taking of such choices. Furthermore, dACC exerts this control over behavior when it, rather than ventromedial prefrontal cortex, interacts with posterior cingulate cortex.*

# Introduction

Investigations of the origins of choice variability have always been an integral part of decision making research. By understanding what parameters influences decisions and how those factors differ between individuals, we can infer much about the evaluative process in general. Furthermore, how individuals differ in their use of choice relevant information is fundamental for our understanding of an individuals behaviour. Since people vary particularly strongly in their use of probability information and such differences have great implications for society, risk taking behaviour has been a particular focus regarding measurements of choice variability. As one of the first areas of research within decision neuroscience, many studies have tried to look at the neural correlates of static risk preferences. The goal of those studies was to find out the neural correlates of a being more risk prone or averse (Huettel, Stowe, Gordon, Warner, & Platt, 2006). Other studies have looked at particular frames or fallacies (De Martino et al., 2006; Hsu et al., 2005) that make people take risks they would otherwise not. However, it is equally interesting to investigate how a person is able to modify their propensity for risk taking dynamically as a function of the environment they find themselves in. Finding the neural correlates of dynamic risk taking and the mechanisms of tracking environmental risk pressures as well as enabling such modifications, is crucial for understanding the neural mechanisms of behavioural flexibility that makes us very well adapted to a variety of environments. Furthermore, differences in perception of the risk environment itself or impairments in modulatory control could contribute to individual differences in risk taking behaviours, rather than only differences in processing of probability per se.

Indeed, seeing risky behaviour as a function of internal and external context is exactly how it is conceptualized in the optimal foraging literature (Caraco, 1981; McNamara & Houston, 1992; Real & Caraco, 1986).

Using the task described in the last Chapter, I can look at the different neural correlates, i.e. brain states, as people are taking risks dynamically, while they are also considering their environmental constraints. Furthermore, my computational model allows me to specifically test how increasing risk boni, i.e. value modification by the changing risk pressure, affects neural activity.

All the regressors used in a given whole brain analysis shared less than 25% of their variance making it possible to identify variance in the fMRI-recorded activity related to each (Figure 1). FMRI analysis focused on two frontal areas – ventromedial prefrontal cortex (vmPFC) and dorsal anterior cingulate cortex (dACC) implicated in decision-making (T. A. Hare et al., 2011; N. Kolling et al., 2012; M. F. Rushworth et al., 2012).

# Methods

## Subjects

Eighteen subjects (nine female), aged 22-36 years, completed the task. They were paid £10 plus a performance-dependent bonus of between £15-30. Ethical approval was given by the Oxfordshire National Health Service Research Ethics Committee (07/Q1603/11).

## Image data acquisition

Structural and functional MRI (fMRI) measurements were taken using a Siemens 3 Tesla MRI scanner. For the fMRI, I used a Deichmann echo-planar imaging (EPI) sequence (Deichmann et al., 2003) (time to repeat (TR): 3000 ms; 3**x**3**x**3mm voxel size; echo time (TE): 30ms; flip angle: 87°; slice angle of 15° with local z-shimming) to minimize signal distortions in orbitofrontal brain areas. This entailed orienting the window at 30° with respect to the AC-PC line.

Additionally for each participant, anatomical images were acquired with a T1-weighted MP-RAGE sequence, using a GRAPPA acceleration factor of 2 (TR: 2200 ms; TE: 4.53 ms; inversion time: 900ms; voxel size: 1**x**1**x**1 mm on a 176**x**192**x**192 grid) [same protocol as in Chapter 3].

## Trial time-course

Each trial started with the presentation of the safer and riskier choice options. After 3-4 seconds (Poisson distributed), a response cue (yellow question mark) indicated participants should make their decision (last chapter Figure 1A). The response cue stayed on until participants made a response using a button box. Reaction time (RT) was registered, defined as the time period between onset of the response cue and button press. We refer to the time period from the presentation of the decision to when participants made their choice as the decision phase. After a variable jitter of 3-7 seconds (Poisson distributed), in which the chosen option was highlighted on the screen, feedback on both options was presented for 1-3 seconds (Poisson distributed). If the participant's choice was rewarded, then the points earned were graphically “added” to the “current points” bar (last chapter Figure 1A) indicating the amount of points gained so far in the block. We refer to this last time period in a trial as the feedback/outcome period.

An average trial took about 12 seconds and the inter-trial interval was jittered between 3 and 7 seconds (Poisson distributed). Within a block, each trial followed the previous one until the last trial of the block was reached. Afterward, a block outcome screen appeared for 4-7 seconds (Poisson distributed) that either displayed the gains of the participant in the block (if the target had been reached) or showed no gain for the block when the target was not reached.

## Image data analysis

I used FMRIB’s Software Library (FSL) (Smith et al., 2004) for image pre-processing and analysis. Functional images acquired were first spatially smoothed (Gaussian kernel with 5mm full-width half-maximum) and temporally high-pass filtered (3 dB cut-off of 100s). Afterward, the functional data were manually denoised using probabilistic independent component analysis (C. F. Beckmann & Smith, 2004) identifying and regressing out obvious noise components (Kelly et al., 2010). We used the Brain Extraction Tool (BET) from FSL (Smith, 2002) on the high-resolution structural MRI images to separate brain matter from non-brain matter. The resulting images guided registration of functional images in Montreal Neurological Institute (MNI)-space using affine registrations (7 degrees of freedom). The data was pre-whitened before analysis to account for temporal autocorrelations (M W Woolrich et al., 2001). Statistical analysis was performed at two levels. At the first level, we used an event-related general linear model (GLM) approach for each participant. On the second level, we used FMRIB's Local Analysis of Mixed Effects (FLAME) (C. F. Beckmann et al., 2003) with outlier de-weighting and tested the single group average. The main effect images are all cluster-corrected results with the standard threshold of $z=2.3$ and corrected significance levels of $p=0.05$.

## FMRI analysis

I acquired up to 1200 volumes with 45 slices per subject. Constant regressors, modelled as boxcar functions, captured the three critical time phases that occurred in each trial: the decision phase, the outcome phase and the time period after every eighth trial in which the block outcome was presented. On each trial the time period indexed by the decision regressor began when the two options appeared on the screen and the duration lasted until the subject pressed a response button. Similarly, onset and duration of trial feedback and block outcome regressors covered the exact time the respective information was displayed on screen. I ran three separate GLM's. From the first:

$$\beta_0 + \beta_1(V_{riskier}-V_{safer}) + \beta_2(V_{chosen}-V_{unchosen}) + \beta_3(V_{riskier}+V_{safer}) + \beta_4(\text{risk bonus}) + \beta_5(\text{trial number}) + \beta_6[\log(\text{reaction time})]$$

we derived the $V_{riskier}-V_{safer}$ signal shown in Figure 3B. The second GLM

$$\beta_0 + \beta_1(M_{riskier}-M_{safer}) + \beta_2(P_{riskier}-P_{safer}) + \beta_3(\text{risk pressure}) + \beta_4(\text{trial number}) + \beta_5[\log(\text{reaction time})]$$

was very similar but now $V_{riskier}$-$V_{safer}$ effects were split by choice (riskier or safer) and probability ($P_{riskier}$-$P_{safer}$) and magnitude ($M_{riskier}$-$M_{safer}$) differences were modeled separately. Although modeling riskier and safer choices separately revealed interesting results that are reported (Figures 3C, 4A), the independent modeling of probability and magnitude did not lead to new insights. In a third similar GLM we focused on splitting the trials with active risk bonus by choice (riskier or safe) for the value and risk bonus regressors. For each trial type (safer option chosen or riskier option chosen) we therefore had the following GLM:

$$\beta_0 + \beta_1(V_{riskier}) + \beta_2(V_{safer}) + \beta_3(\text{risk bonus}) + \beta_4(\text{trial number}) + \beta_5[\log(\text{reaction time})]$$

This revealed the risk bonus effect in vmPFC (Figure 2A) and effect of the number of remaining trials (Figure 2B). In all analyses I excluded all decisions taken after the target was reached as well as decisions when the target was beyond reach.

All dACC activity time-courses were taken from a single dACC ROI (MNI coordinates x=-2, y=28, z=36) where the group level effects of risk pressure and $V_{riskier}$-$V_{safer}$ overlapped (also peak of overall $V_{riskier}$-$V_{safer}$ effect). PCC, vmPFC, and IFG ROIs were centered at the group peak contrast effect that first identified them. The GLMs used in constructing activity time courses during decision periods included similar regressors to those used in the whole brain analyses: $V_{riskier}$-$V_{safer}$, trial number, log(reaction time), risk pressure and risk bonus. Figure 1 is a correlation matrix of those parameters. The highest $r^2$ correlation between the conceptually related risk bonus and risk pressure is still below .4.

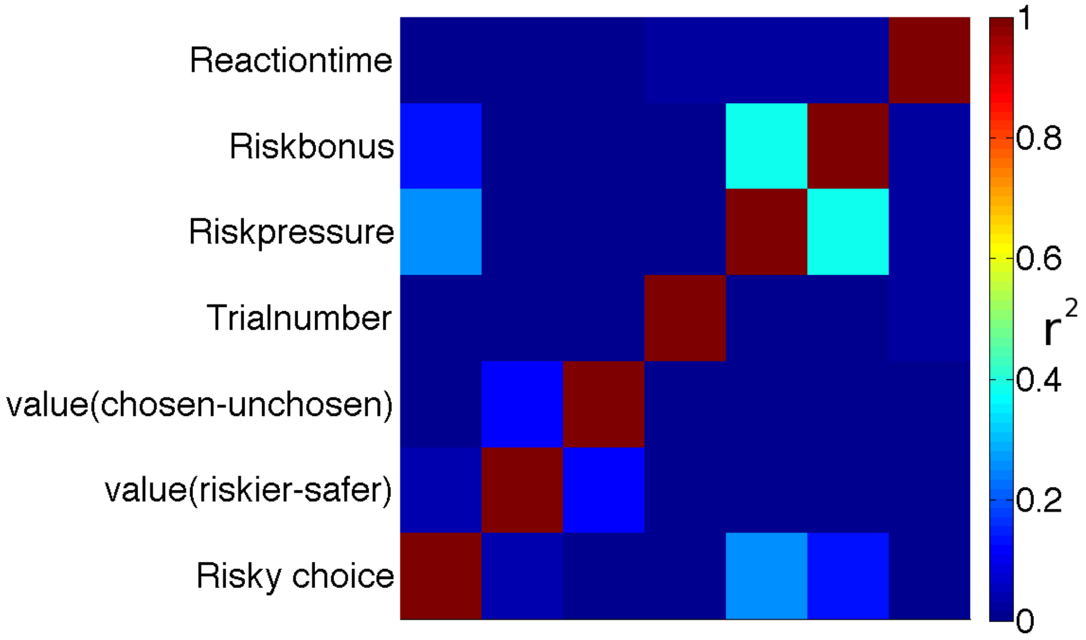

***Figure 1:*** *overall correlations ($r^2$) between all regressors at the decision stage. The highest correlation is between the conceptually related factors of risk bonus and risk pressure. However, because risk bonus contains model based information about the optimal consequences of the risk pressure context on the evaluation of two specific choice options it is not confounded with the risk pressure itself.*

I also conducted PPI analyses (Friston et al., 1997) to investigate functional connectivity between vmPFC and PCC and between dACC and PCC during the decision period. For the first of these analyses I generated a demeaned BOLD time-course regressor from the vmPFC as well as an interaction term with the demeaned psychological regressor [I used the inverse of the risk bonus regressor because vmPFC had been shown to be inversely related to the risk bonus in an earlier analysis (Figure 2A)] separated according to choice (riskier or safe), to generate two PPI interaction terms corresponding to the trials in which riskier and safer choices were made. The main effects of the same two psychological regressors were also included in the GLM. In addition we also included the relative value regressor $V_{riskier}$-$V_{safer}$ split by choice and log (reaction time). The PPI analysis for the dACC PPI analysis was conducted in an analogous manner but using a physiological regressor derived from the dACC ROI and a psychological regressor of $V_{riskier}$-$V_{safer}$. For the PPI time-courses, I subtracted the global mean (the average from all voxel at a specific time point) from the individual time-course of each region investigated, to avoid positive results due to global correlations.

To investigate interactions between dACC-PCC and vmPFC-PCC as a function of IFG activity we took the difference between the dACC-PCC PPI regressor and the vmPFC-PCC PPI regressor and multiplied it with the normalized, time-course from the left IFG, seeded at the group peak effect of the $choice_{riskier}$ – $choice_{safer}$ contrast (Figure 3A).

## General linear models of fMRI analysis in ROIs

The time-courses were derived from regions-of-interest (ROI), a sphere of 3 voxel radius, identified in Montreal Neurological Institute (MNI) standard space on the basis of the whole group analysis, calculating a mean time-course within a ROI in each subject individually and the coordinates were then transformed to individual subject space by using the same linear registration as in the initial analyses. We then oversampled the time-course by ten and created epochs from the beginning of an event onward and applied a GLM to every pseudo-sampled time point separately. By averaging the resulting β weights across subjects we created the time-courses shown (the standard errors were calculated between subjects). The approach is identical to Chapter 3.

# Results

## Contextual modification of value

To look at the impact of context we split all trials into those where the context meant that there was a *risk bonus* and those where there was none (see Methods). First I looked at the main effect of the *risk bonus*, in other words I looked at the model-based modification of each trial's option values away from the default safer choice in favor of the riskier choice as a result of *risk pressure*.

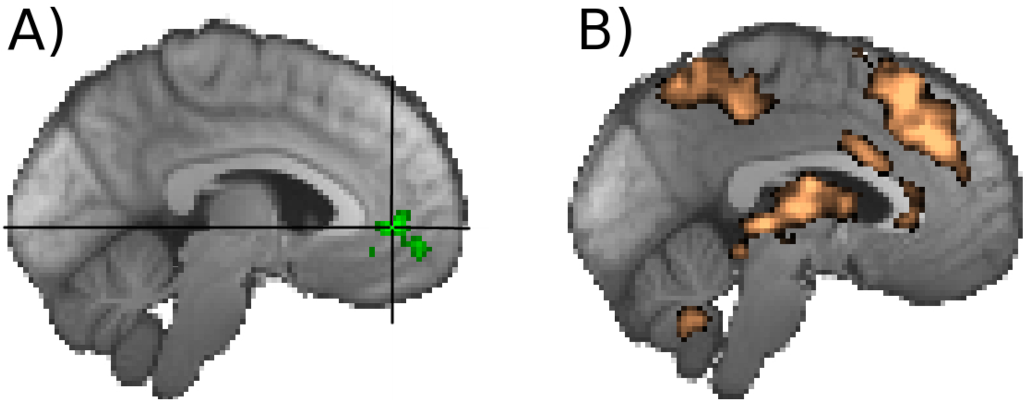

***Figure 2: A)*** *Decreasing risk bonus was associated with increased vmPFC activity. The impact was present regardless of subjects' choices.* ***B)*** *Activity increases in dACC and elsewhere during the decision phase as number of trials remaining decreased.*

I observed a relative *decrease* in vmPFC activity as risk bonus increased that was independent of which choice, riskier or safer, subjects ultimately made (Figure 2A). In other words vmPFC activity is negatively related to the *risk bonus*. Beyond this choice independent decrease I was unable to find any choice-related value signals, either "raw" ones (equation 2) or contextually modified ones (equations 3,4 and 5) (such as an absolute or relative chosen value signal). This is in stark contrast to most other studies which have suggested vmPFC codes the value or relative value of potential or attended choices, (Boorman et al., 2009; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; N. Kolling et al., 2012; Lim et al., 2011; Philiastides et al., 2010; Wunderlich et al., 2012). In summary, while vmPFC may normally track choice values during decision-making it does not do so in the current paradigm in which both immediate value and current risk bonus had to be integrated to make appropriate choices. Instead vmPFC's activity decreased if the context meant that there was a risk bonus and subjects increasingly biased their decisions towards the riskier choice and away from the default of taking the safer choice.

## Sequential progression and the generation of riskier choice decisions

Progression through the eight-trial mini-blocks had a strong impact on activity in dACC, and other regions (Figure 2B). Crucially, in addition to this effect, dACC was sensitive to another piece of information more directly related to the decision strategy subjects used; its activity was correlated with *risk pressure*, the average points per trial needed to reach the target (Figure 3C). However, the effect of *risk pressure* on dACC activity reversed depending on choice. A positive effect of *risk pressure* on dACC activity was apparent when subjects chose the safer option, whereas a negative effect was apparent when subjects chose the riskier option.

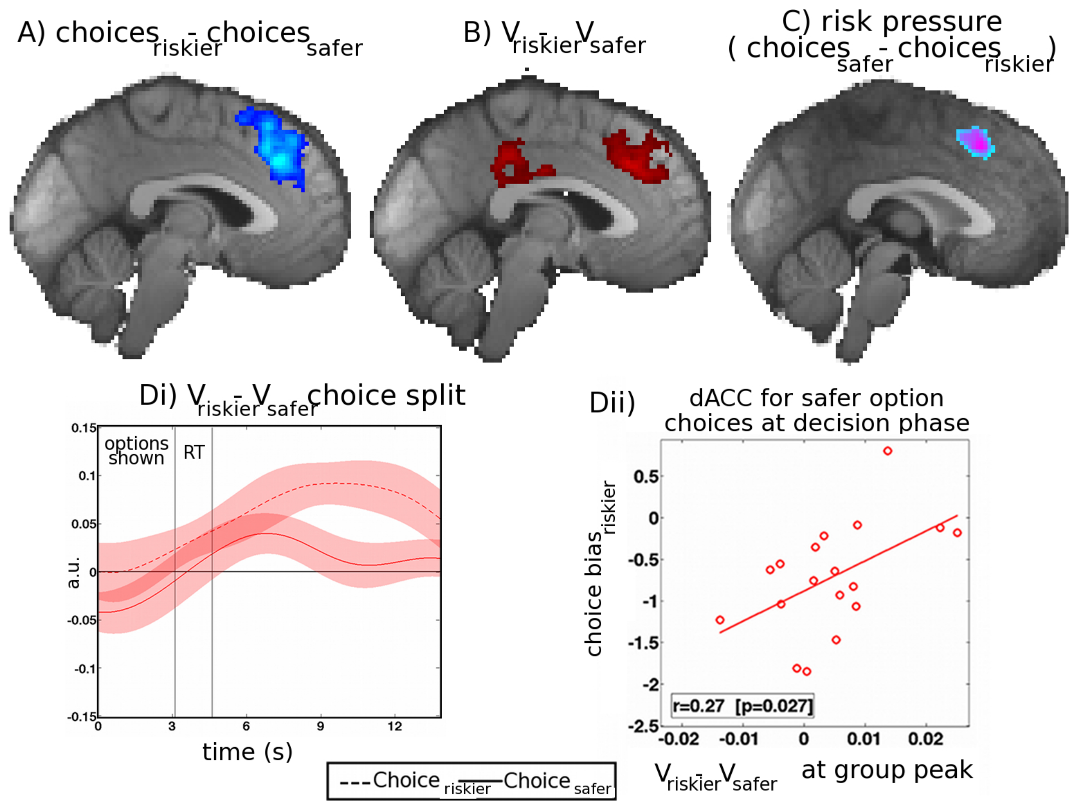

***Figure 3: A)*** *When just the main effect of choice is considered then activity in dACC and adjacent dorsomedial frontal cortex increased for riskier as opposed to safer choices.* ***B)*** *Again the same dACC region and adjacent dorsomedial frontal cortex exhibited increased activity as a function of the relative value of riskier choices* ($V_{riskier}$-$V_{safer}$). ***C)*** *dACC activity during the decision phase increased as a function of* risk pressure *when subjects did not succumb to it and instead made riskier rather than safer choices; Di) dACC group time-course of* $V_{riskier}$-$V_{safer}$ *effect shown separately for riskier (continuous line) and safer choices (dotted line).* ***Dii)*** *Subjects that were less biased against riskier choices exhibited a higher dACC* $V_{riskier}$-$V_{safer}$ *effect at the peak of the group time-course.*

In other words, dACC activity increased with increasing *risk pressure* when choices went against the prevailing *risk pressure* but decreased with increasing *risk pressure* when subjects chose in agreement with *risk pressure* (Figure 3C; 4A).

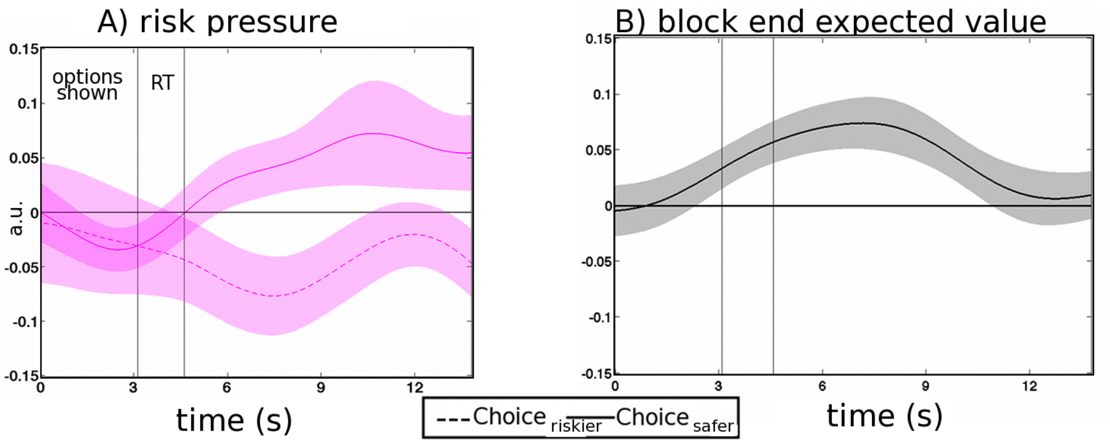

***Figure 4: A)*** *the risk pressure signal in dACC increased when subjects did not act in accordance with it but it had a negative effect when subjects did act in accordance with risk pressure and took the riskier option.* ***B)*** *Independent of choice (riskier or safer), higher expected value at the end of the block, as estimated using our model, was related to increased dACC activity. See also Figure S6.*

The dACC *risk pressure* signal cannot be explained away as a signal indexing approach towards a reward that might be delivered at the end of the block (Croxson et al., 2009; Shidara & Richmond, 2002) because progress through the sequence of trials itself was present as a separate regressor in the general linear model (GLM) and associated with an independent effect on dACC activity (this is the effect already shown, Figure 2B). The *risk pressure* signal cannot be explained away as a consequence of differing average reward expectations associated with different target levels because of the use of a "multiplier" procedure (Methods; Experimental Task) ensured that average reward expectations were the same at the beginning of a block regardless of the target. It is, however, the case that expectations about the reward that would be received at the end of the block (as opposed to just the current trial within the block) began to diverge as soon as participants began to make choices and were either lucky or unlucky. However, when I included an additional term in the GLM indexing the expected value of the reward at the end of the block I found it had an independent effect on dACC activity (Figure 4B). No similar signal was observed in vmPFC (Figure 5 A and B).

Moreover, none of these areas appeared to carry a future state space value difference code. In other words, I did not find an area that represented the difference between the expected value at the block end if one option or another were taken on the current trial. This could be due to the feedback given after every trial which might make the actual end of block states seem relatively unpredictable in comparison to the outcome for each component decision in the block and which might have led to a decision-making process that was more guided by current option values and a contextually based value modification. Nevertheless a reinforcement learning and state-based approach might provide additional important insights under different conditions, for example if subjects always progressed through the same fixed order of decisions.

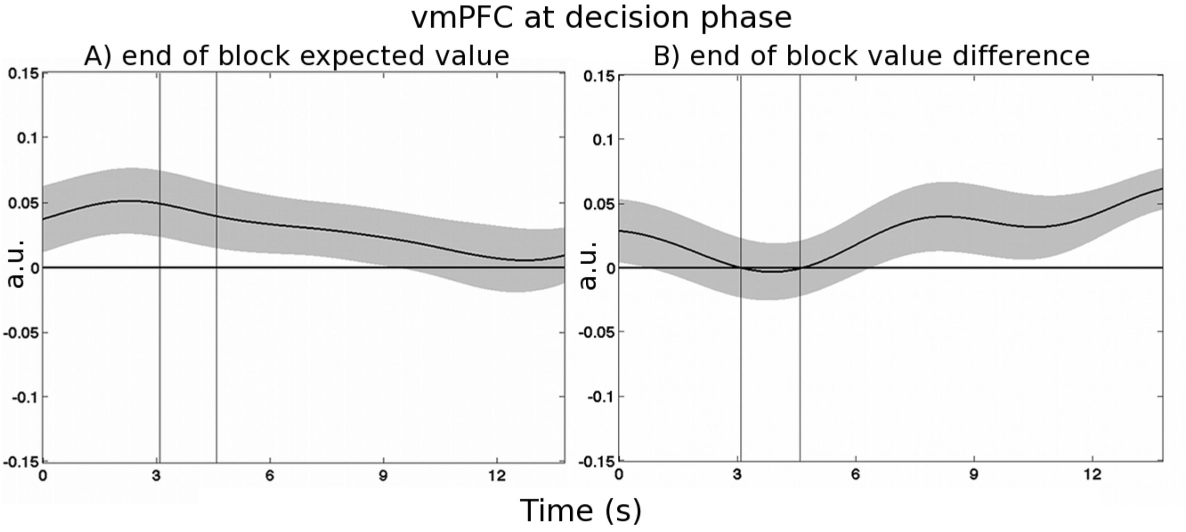

***Figure 5:** In vmPFC there was no effect of the overall expected value at the end of the block as estimated by our model (left), nor was there any effect of the difference in the block's expected value as a function of the current choice being taken and the alternative choice (right).*

In summary, dACC exhibited a number of signals related to progress through the sequence of decisions, the expected reward at the end of the sequence, and a *risk pressure* signal indexing the need to take riskier choices as a function of contextual factors (accumulated resources, target and remaining foraging opportunities). The *risk pressure* signal flipped with the decision strategy subjects pursued (safer versus riskier); it was positive when subjects needed to change their behavior and switch to riskier choices as opposed to the default safer choice.

In addition to these contextual effects the same dACC region also exhibited activity that was tied to specific patterns of choice and choice valuation. DACC activity was higher in decisions in which the riskier rather than the safer choice was taken ($\text{choice}_{\text{riskier}}$-$\text{choice}_{\text{safer}}$, Figure 3A). Not only did the main effect of $\text{choice}_{\text{riskier}}$-$\text{choice}_{\text{safer}}$ activate dACC, but so did the relative value of the riskier choice ($V_{\text{riskier}}$-$V_{\text{safer}}$; Figure 3B). Moreover the signal encoding $V_{\text{riskier}}$-$V_{\text{safer}}$ was stronger on trials on which subjects actually took the riskier choice, although it was also present when subjects took the safer choice (Figure 3Di). Individual variation in $V_{\text{riskier}}$-$V_{\text{safer}}$ signal size at the group peak coordinate in dACC when taking the safer choice was related to how frequently subjects took the riskier choice (Figure 3Dii), suggesting that variation in this aspect of dACC activity is intimately related to decision-making.

Activity increases related to the $\text{choice}_{\text{riskier}}$-$\text{choice}_{\text{safer}}$ contrast were also apparent in the inferior frontal gyri (IFG) and frontal operculum (see Table 1) while the $V_{\text{riskier}}$-

$V_{safer}$ contrast was also associated with activity in posterior cingulate cortex (PCC) (Figure 3B and Figure 8A,B) and dorsolateral prefrontal cortex (dlPFC). I propose an explanation of IFG and PCC activity in a later section. In summary, one region – dACC – encoded the expected reward at the end of the sequence of decisions, progress through the sequence of decisions, *risk pressure*, the taking of riskier and opposed to safer choices, and the relative value of riskier choices. The time-course analyses shown in figures 3D and 4 are all from the same region of interest (ROI) with Montreal Neurological Institute (MNI) coordinates x=-2, y=28, z=36.

| **Contrast** | **Z-coordinates MNI (mm)** | **Label** | |
|---|---|---|---|
| $V_{riskier}$-$V_{safer}$ | 12,34,54 | Right DLPFC | |
| | 54,-60,38 | Right PPC | |
| | 0,-34,40 | PCC | |
| | -50 -64 38 | Left PPC | |
| | -2,28,36 | dACC | |
| $Choice_{riskier}$-$Choice_{safer}$ | -2,34,36 | dACC | |
| | 52,24,4 | Right IFG | |
| | -28,24,-8 | Left anterior insula | |
| | -48,30,-8 | Left IFG | |

***Table 1:*** *List of peak activations for the corrected significant clusters for the whole brain contrasts.*

Although further experiments are needed to determine quite why the impact of *risk pressure* on dACC activity changed depending on whether or not subjects acted in accordance with it or not, it is worth considering that it may reflect the operation of an evidence accumulation process to threshold that finally results in a riskier choice being taken. This would be consistent with the observation that actually taking a riskier choice activates dACC (Figure 3A) as does the evidence advocating such a choice ($V_{riskier}$-$V_{safer}$; Figure 3B). If an accumulation process is taking place before riskier choices are generated then it seems *risk pressure* increases such activity (Figure 3C). However, once such a process has hit its bound, triggering the taking of a riskier choice, further activity increases related to the *risk pressure* are not observed. Although there is evidence for the operation of accumulation processes in dACC (Hayden, Pearson, et al., 2011; N. Kolling et al., 2012) further experiments are needed to determine whether *risk pressure* is contributing to such a process.

In the past, another region, the lateral frontal pole, FPl, has been associated with tracking the values of alternative courses of action (Boorman et al., 2009, 2011). FPl activity also increased as a function of $V_{riskier}$-$V_{safer}$ (Figure 6) and individual differences signal strength were related to individual differences in the degree to which subjects modulated their behavior according to the model-based risk bonus (Figure 6Bii). Unlike in dACC, FPl signals tracking *risk pressure* and $V_{riskier}$-$V_{safer}$ were apparent regardless of which choice, riskier or safer, subjects took (Figure 6C,D). In other words, FPl provides a constant signal, regardless of current choice type, of how necessary it is to adjust choice strategy away from the default safer choice and towards the riskier choice in the face of *risk pressure*.

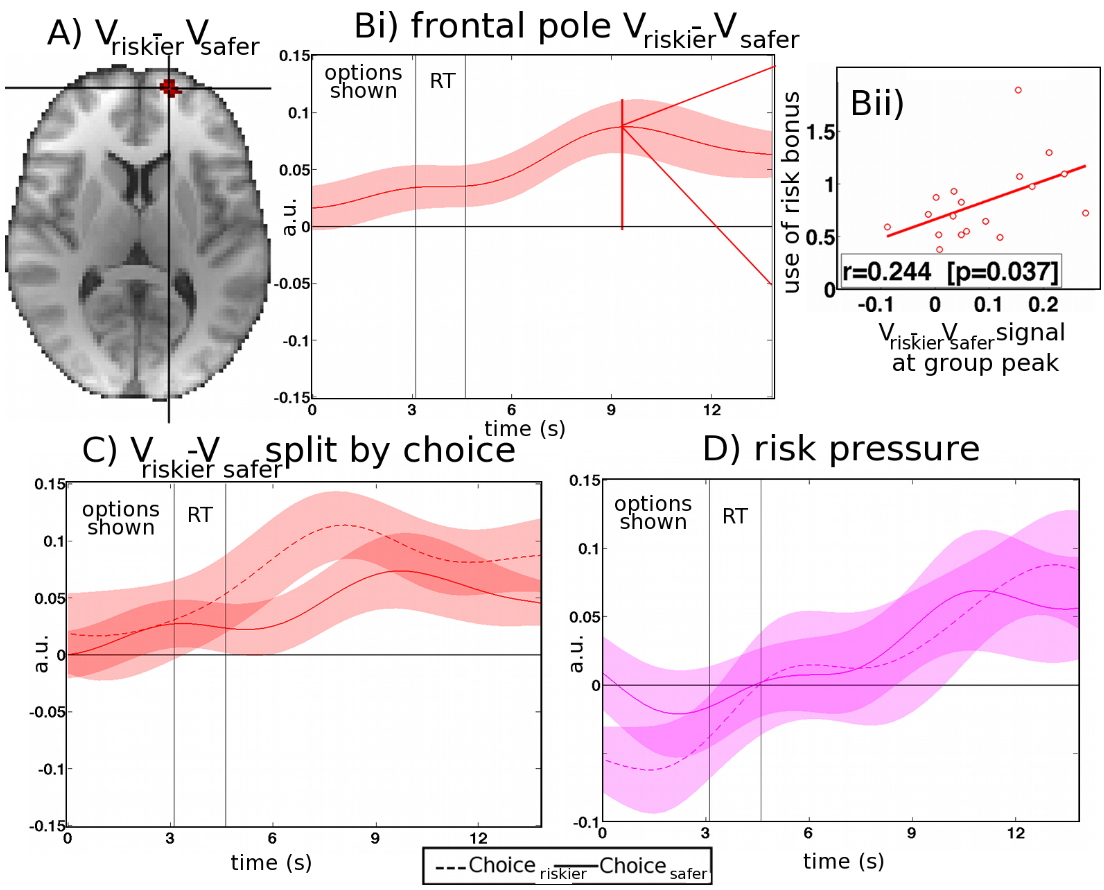

***Figure 6:A)*** *The FPl carried a* $V_{riskier}$*-*$V_{safer}$ *signal regardless of which choice subjects ultimately made.* ***Bi)*** *The group time-course of the* $V_{riskier}$*-*$V_{safer}$ *signal (regardless of whether riskier or safer decisions were made). Bii) Individual subject* $V_{riskier}$*-*$V_{safer}$ *signal effect sizes at the peak of the group time-course predicted individual behavioral sensitivity to risk bonus;* ***C)*** *The* $V_{riskier}$*-*$V_{safer}$ *signal was present in FPl both for riskier (continuous line) and safer choices (dotted line).* ***D)*** *Risk pressure activated PFl similarly for riskier choices (continuous line) and safer choices (dotted line).*

## Outcome related signals

So far I have shown dACC is more active when a riskier choice as opposed to a safer choice is made (Figure 3A) and that dACC activity reflects the relative value of riskier choices (Figure 3B) and *risk pressure* (Figure 3C). Next we consider whether dACC also contains signals related to evaluation of the success of riskier choices when their outcomes are revealed. Subjects can update their estimate of *risk pressure* or the likelihood that they will reach the target when they see the outcome of their choice. Therefore we tested whether dACC activity was related to changes in *risk pressure* at the time of outcome presentation.

To do this I plotted the effect of decision outcome on dACC activity after safer and after riskier choices. In addition we also binned the outcome effects according to three levels of the change they caused to *risk pressure*. In other words, we examine the effect of two factors, choice type (riskier versus safer) and the size of impact of outcome on *risk pressure* (three levels: low, medium, high).

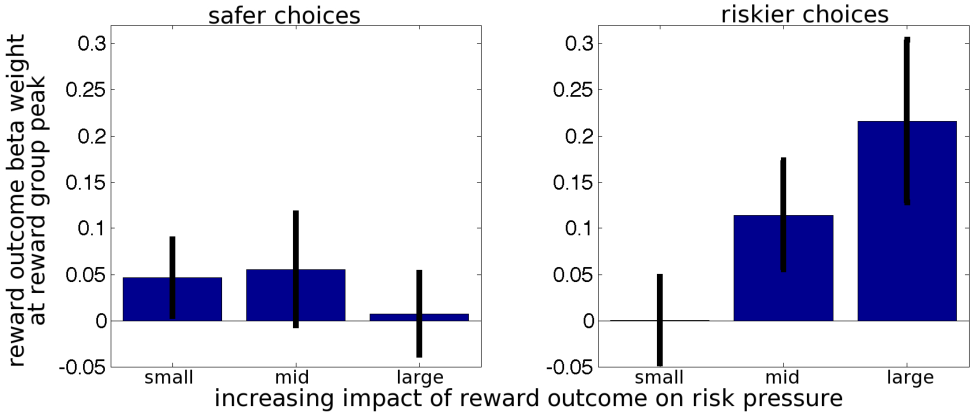

***Figure 7**: The effect of outcomes on activity in the dACC after riskier or safer choices. The effect of outcomes is not only dependent on choice (riskier or safer), but also on how much impact they have on future* risk pressure *(greater on the right than on the left of each panel).*

There was a significant interaction ($F_{2,\ 34}$=3.417, p=0.044) between the two factors on outcome-related dACC activity. As the outcome's impact on *risk pressure* increased so did the outcome's impact on dACC activity but this was only the case when riskier choices were taken (Figure 7). After safer choices there was no increase in the impact an outcome had on *risk pressure* (in fact if anything there was a slight decrease).

The results remain the same even after controlling for the expected value of the whole block ($F_{2,34}$=4.352,p=0.021) and outcome surprise ($F_{2,34}$=3.848,p=0.031). At the time of outcomes dACC is not simply encoding prediction errors in value (Jocham, Neumann, Klein, Danielmeier, & Ullsperger, 2009; Steven W Kennerley et al., 2011; M. Matsumoto et al., 2007) but the impact riskier choices have on reducing *risk pressure*.

## Functional connectivity and networks of choice

A large body of work has implicated vmPFC in reward-guided decision-making but it was deactivated in the current experiment when the subject's context meant that the default safer choice should not be taken and the riskier choice should be taken instead (risk bonus effect; Figure 2). By contrast dACC activity increased with *risk pressure* and was greatest when subjects chose the riskier choice (Figure 3). It therefore seems that the two frontal brain regions, vmPFC and dACC, may mediate decisions in different situations. If there are two systems competing to control behavior then it is not clear how the competition is resolved or if there is any critical area that mediates both types of decisions.

One region that may be a nexus for both types of decision modes is the PCC. In many neuroimaging studies it carries a value difference signal like that seen in the vmPFC (Boorman, Rushworth, & Behrens, 2013; FitzGerald et al., 2009; N. Kolling et al., 2012). However, a series of single neuron recording studies have emphasized the similarities between the parameters both it and dACC encode (Pearson, Heilbronner, Barack, Hayden, & Platt, 2011) and in the current study it, like dACC, was sensitive to the relative value of riskier choices ($V_{riskier}$-$V_{safer}$)(Figure 3B). The PCC region that was active in this contrast probably includes areas 31 and 23 but it also the caudal cingulate motor areas which lie in the cingulate sulcus at the point of its inflection into its marginal ramus (Céline Amiez et al., 2012; M. Beckmann et al., 2009).

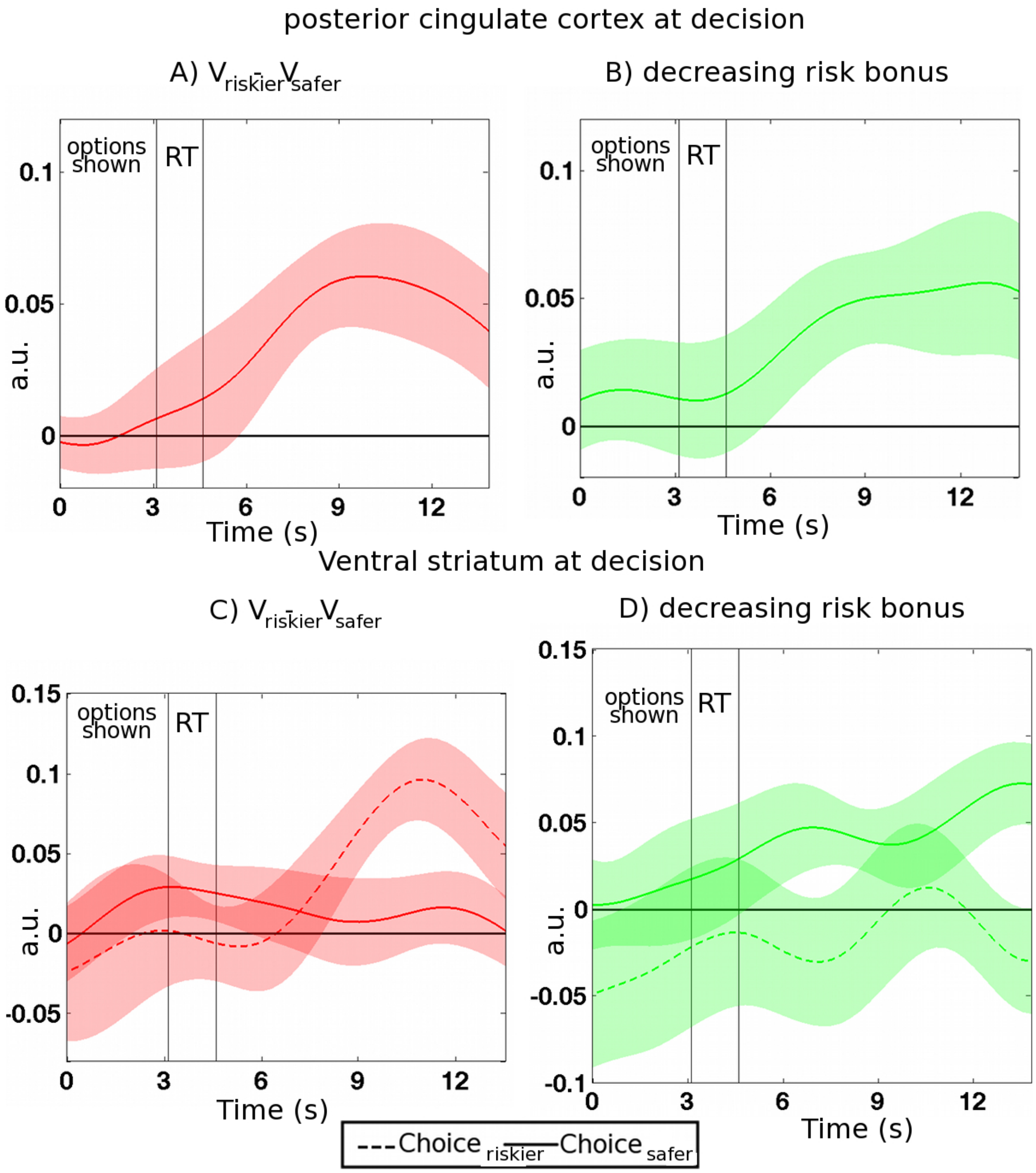

*Figure 8*: *Posterior cingulate cortex (PCC) for all decisions. A) Activity increased with the relative value of the riskier choice ($V_{riskier}$-$V_{safer}$) as in the dACC. B) However, PCC also activated when the risk bonus was reduced as did vmPFC. Time-course of the ventral striatum at decision split by choice (riskier or safer). C) The relative value of riskier choice ($V_{riskier}$-$V_{safer}$) only activated the ventral striatum when it was chosen. D) The reduction of*

*the risk bonus activated the ventral striatum only when the safer choice was taken.*

In macaques the caudal cingulate motor area projects to both primary motor cortex and ventral horn of the spinal cord (Dum & Strick, 1996) and so it may be involved in making the movement needed for implementing a particular choice. In macaques it is connected to dACC, vmPFC and adjacent parts of PCC (Morecraft & Van Hoesen, 1993; Parvizi, Van Hoesen, Buckwalter, & Damasio, 2006) and so it is therefore a region through which vmPFC, dACC, and PCC might interact and influence action movement selection.

I conducted a psychophysiological interaction (PPI) test of whether vmPFC and dACC activity were coupled with PCC activity in different ways as a function of choice (riskier or safer) and their relative values ($V_{riskier}$-$V_{safer}$). There was greater coupling between dACC and PCC as a function of $V_{riskier}$-$V_{safer}$ but only when the riskier choice was chosen (Figure 10B). In other words, PCC's coupling with dACC increases as a function of the decision variable, $V_{riskier}$-$V_{safer}$, that predisposes participants to take riskier choices (last chapter) and which influences dACC activity (Figure 3). By contrast, vmPFC was more coupled with PCC when the default safer choice was taken and as a function of risk bonus being low (Figure 10A). In other words, PCC's coupling with vmPFC increased in inverse relationship with the decision variable risk bonus. The inverse of risk bonus was associated both with lower vmPFC activity (Figure 2A) and higher frequencies of taking the default safer option (last chapter). PCC carries signals that are more similar to either vmPFC or dACC depending on the prevailing context at the time of each decision and depending on the choice subjects actually took (for coupling pattern of the ventral striatum see Figure 9).

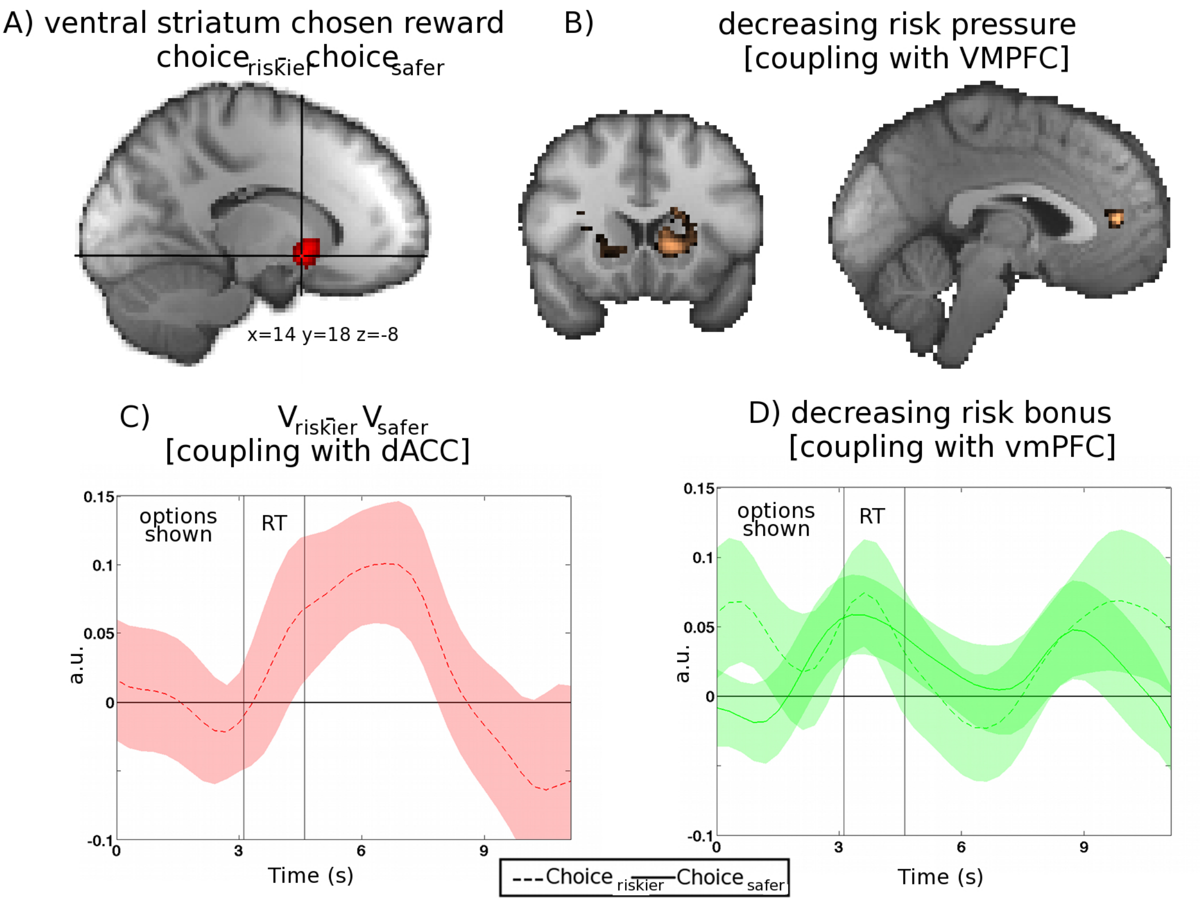

***Figure 9:*** *A) The ventral striatum was more sensitive to rewards after taking the riskier choice than after taking the safer choice. It only encoded* $V_{riskier}$*-*$V_{safer}$ *when the riskier choice was taken and decreasing risk bonus when the safer choice was taken). B) Physiological psychological interaction (PPI) analysis based on the time course of vmPFC activity from the ROI in figure 2A and inverse risk pressure [we used the inverse of the risk bonus regressor because vmPFC had been shown to be inversely related to the risk bonus in an earlier analysis (Figure 2A). Both, ventral striatum and pgACC (as opposed to dACC) exhibited higher functional connectivity with vmPFC when risk pressure was eased during the decision period. C) PPI between the ventral striatum and dACC (point of group peak activation related to* $V_{riskier}$*-*$V_{safer}$ *(see Figure 3). Note that the functional connectivity only changes as a function of* $V_{riskier}$*-*$V_{safer}$ *when the riskier choice is taken. D) PPI between the ventral striatum and the vmPFC region from figure 2A. The vmPFC region was more*

*active with decreasing risk bonus and it was also more coupled with the ventral striatum.*

Finally, I looked for evidence of a brain area that might resolve competition between dACC and vmPFC and determine which couples with PCC. I focused on IFG because I had noticed its activity changed with choice in the current experiment even though it did not carry a value signal (see above) and because it has been argued that it or adjacent regions exert a regulatory influence over vmPFC activity in other situations (Baumgartner, Knoch, Hotz, Eisenegger, & Fehr, 2011; T. a Hare, Camerer, & Rangel, 2009). I carried out a further PPI analysis, which, once again, tested vmPFC-PCC and dACC-PCC coupling but this time I examined vmPFC-PCC and dACC-PCC coupling as a function of IFG activity. PCC's coupling with dACC versus vmPFC, was related to IFG activity when the riskier choice was chosen (Figure 10C). In other words, with increasing IFG activity the relative strength of dACC-PCC coupling increased (which was also, as described above, a function of the $V_{riskier}$-$V_{safer}$ value difference) as opposed to vmPFC-PCC coupling (which was also, as described above, a function of low risk bonus).

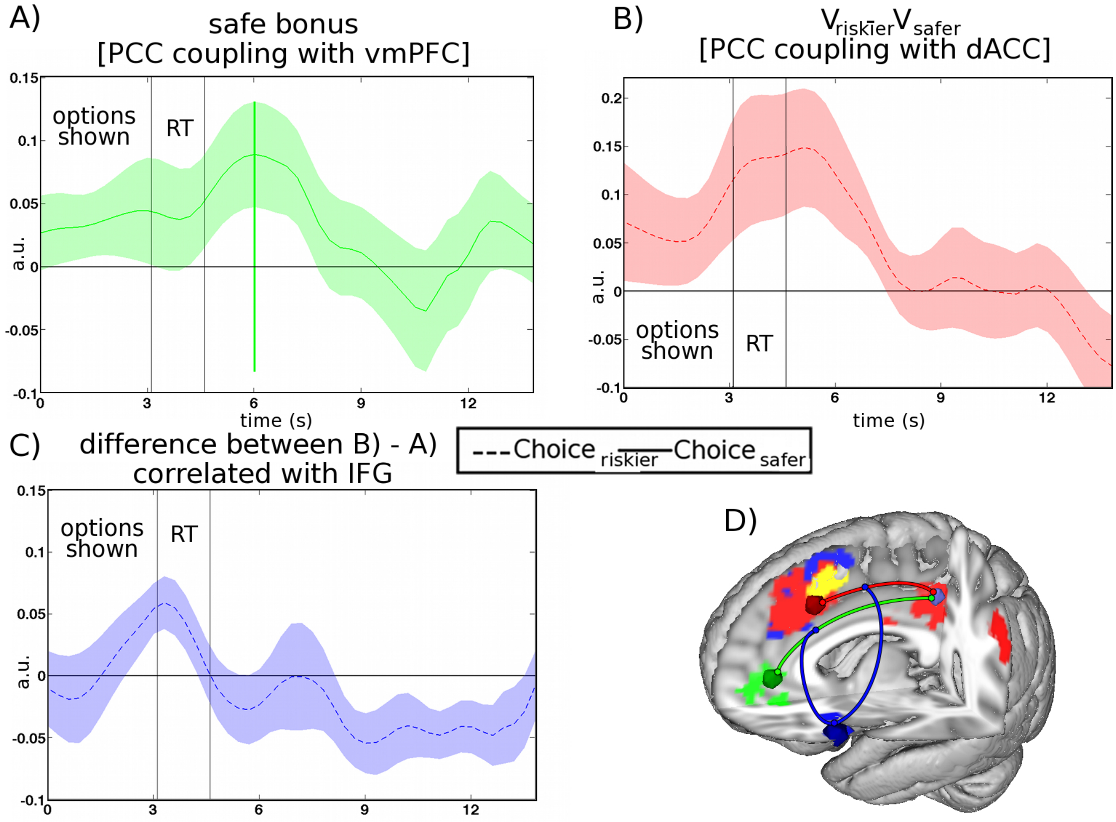

***Figure 10:*** *PPI analyses demonstrated vmPFC and dACC interactions with PCC during different types of decisions and the relationship with IFG activity.* ***A)*** *Time-course illustrating PPI between PCC and vmPFC as a function of decreasing risk bonus on trials when the safer option was taken.* ***B)*** *Time-course of PPI between PCC and dACC as a function of $V_{riskier}$-$V_{safer}$ on trials on which the riskier option was chosen (dotted).* ***C)*** *Time-course of PPI between PCC, left IFG, and the differing effects of vmPFC and dACC (the two regressors from panels B and A).* ***D)*** *Illustration of effects in panels A, B and C. The PCC couples with dACC and vmPFC during decisions in which the riskier option (red) and safer option (green) were taken respectively. Left IFG may regulate PCC's interactions with the vmPFC and dACC by increasing the relative degree of coupling to the former as opposed to the later during riskier choices.*

Such a pattern of results is consistent with a controlling function for IFG, not just of activity in other brain regions but of the interconnectivity between other brain regions. A clear demonstration of the causal direction of effects, however, would require showing that IFG disruption affected the coupling patterns.

# Discussion

Having shown my model was a good description of the behavioural changes due to changing environmental risk pressure in the last chapter, this chapter was about relating such adaptive decisions to their possible underlying neural mechanisms. Using the model estimates as well as the risky and safe value parameters and contextual pressure itself, I was able to describe decisions, environmental factors and to model their relationships with one another and with various neural structures and networks during the taking of a choice and when monitoring the outcome of the choice. Most of these structures have before been implicated in decision making, but the changing nature of the way participants made decisions during this task, allowed me to dissociate their roles further.

## Neural systems for decision-making

VmPFC and dACC might constitute two distinct decision-making systems rather than components of a single serial system for decision-making (Boorman et al., 2013; N. Kolling et al., 2012; M. F. Rushworth et al., 2012). There was evidence for vmPFC and dACC acted in independent or even opposite ways in the current study.

Although there has been particular interest in the role that vmPFC plays in valuation and decision-making (Boorman et al., 2009; Camille et al., 2011; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; N. Kolling et al., 2012; Lim et al., 2011; M P Noonan et al., 2010; Philiastides et al., 2010; Wunderlich et al., 2012) vmPFC did not mediate the influence of the contextual variable of *risk pressure* on decision making. Instead vmPFC became less active as risk bonus increased (Figure 2A). Both lesion and neuroimaging evidence suggest that, in addition to its role in valuation and decision-making, vmPFC mediates the repetition of a previously successful choice or the taking of a default choice (Boorman et al., 2013; M P Noonan et al., 2010; M. P. Noonan et al., 2012) and the pattern of activity recorded in vmPFC suggests it was similarly concerned with default responses in the present task.. This interpretation is suggested by the following observations. On average subjects were risk averse and defaulted to taking the safer choice. This was most true on trials on which the *risk pressure* was low (see last chapter) and it was on just such trials that vmPFC activity was greatest (Figure 2A). Note that in this task default choices occur when decision-making is less constrained by context

Instead of vmPFC, both dACC, FPl were pre-eminent in tracking the *risk pressure* afforded by the evolving decision context (Figures 3,4,6,7). FPl and dACC have been co-activated in other studies (Boorman et al., 2011; Daw et al., 2006) and together they constitute another neural system important for decision-making. In macaques FP and dACC are monosynaptically interconnected (Petrides & Pandya, 2007). There is evidence that FPl, unlike more medial FP cortex, is only found in humans and not in other primates but that it remains interconnected with dACC (Neubert et al., 2014).

In FPl signals indicating both *risk pressure* and $V_{riskier}$-$V_{safer}$ were present regardless of which choice (riskier or safer) subjects took. By contrast, in dACC both signals changed as a function of choice and the taking of riskier choices was associated with additional activity (Figures 3, 4). These observations suggest dACC was more closely related to the actual decision to take a specific riskier option while FPl had a more consistent role in tracking the contextual variables that guided decisions. Individual variation in the sizes of both FPl and dACC signals were predictive of subjects' sensitivities to the risk bonus and their predispositions to make riskier choices (Figure 3Di, 6Bii). Individual variation in the $V_{riskier}$-$V_{safer}$ signal in dACC, when the safer choice was taken, predicted how frequently subjects rejected the default safer choice and took the alternative riskier option. This is consistent with the idea that, when one course of action is being pursued or is the apparent default course of action, dACC is tracking the value of switching to an alternative (N. Kolling et al., 2012; M. F. Rushworth et al., 2012). In a previous study dACC also encoded the relative value of switching away from the current default choice to explore a foraging environment (N. Kolling et al., 2012). An "inverse value difference" signal is often seen in dACC (N. Kolling et al., 2012; M. F. Rushworth et al., 2012); when a decision is being made dACC activity increases as the value of the choice not taken increases and decreases as the value of the choice that is taken increases. This signal is opposite to the one seen in vmPFC. One simple interpretation of the dACC inverse value signal is that it is encoding the value of switching away from the current choice to an alternative one.

So far I have focused on dACC signals that are recorded at the time decisions are made but dACC activity is also observed subsequently at the time of decision outcomes. Outcome-related dACC signals can also be interpreted in a similar framework and related to the need to switch away from a current choice and to explore alternatives (Hayden et al., 2009; Hayden, Pearson, et al., 2011; Quilodran et al., 2008).

A notable feature of dACC activity in the present study was that, unlike vmPFC activity, it reflected the longer term value of a course of action, progress through the sequence of decisions, and the evolving level of *risk pressure* (Figure 2B, 3C, 4).

Boorman and colleagues (Boorman et al., 2013) have also argued dACC reflects the longer term value of a choice and not just its value at the time of the current decision that is being taken. Not only does dACC carry signals related to the longer term and contextually modified value of a choice but it also encodes the approximate value of a number of potential alternative courses of action (N. Kolling et al., 2012). By contrast, vmPFC is more concerned with the valuation of specific aspects of specific choices. Value-related activity in vmPFC is most prominent when the choices' values are determined by multiple attributes and it is necessary to identify the attribute currently most relevant for guiding a choice (Fellows, 2006; Hunt et al., 2012).

One prominent account of dACC function has emphasized its role in detecting response conflict (Mattew M Botvinick, 2007). Although some features of the dACC results are consistent with the response conflict account, other features, such as the value difference signal ($V_{riskier}$-$V_{safer}$) in dACC are not easy to interpret within the framework offered by the conflict account; the dACC $V_{riskier}$-$V_{safer}$ signal encodes the relative value of the riskier choice as opposed to the safer choice but it was stronger when that very same choice, the riskier choice, was being made and when, because of its relatively greater value, the decision should have been relatively easy to take.

## PCC as final common pathway for decision-making

It has not previously been clear how the two distinct decision-making mechanisms associated with vmPFC and dACC might interact. The present study suggests PCC is part of a final common pathway to action selection employed by both systems.

The PCC region probably included areas 31 and 23 but it extended to the main branch of the cingulate sulcus at the point of its inflection into its marginal ramus (Figure 3B) where the caudal cingulate motor area is situated (Céline Amiez et al., 2012; M. Beckmann et al., 2009). Activity in this region, or just caudally, has been reported to resemble both that in vmPFC (Boorman et al., 2013; FitzGerald et al., 2009; N. Kolling et al., 2012) and in dACC (Pearson et al., 2011). In this study, it was more closely coupled with vmPFC when risk bonus was low and the safer choice was taken, but it was more closely coupled with dACC when the riskier choice was taken and when the relative value of the riskier choice ($V_{riskier}$-$V_{safer}$) increased (Figure 10 and 11). In other words, which region PCC couples with during a decision is related to the signal it carries and the choice subjects ultimately make. This means that while there may be two parallel decision-making circuits dependent on dACC and vmPFC, both circuits have a serial element that converges in or just posterior to the caudal cingulate motor area in PCC.

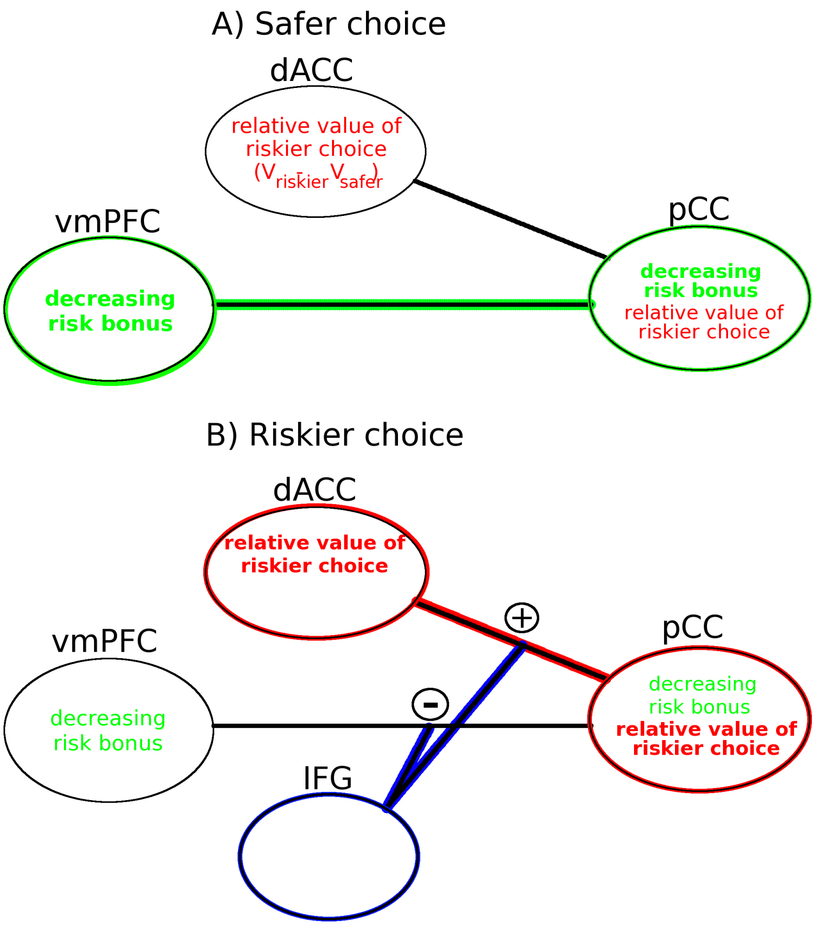

***Figure 11***: *A) The proposed network involved in generation of decisions in favor of the*

*riskier choice. The three key brain areas, vmPFC, dACC, and PCC are indicated and the decision variable that was prominent in each area during the task is summarized. The brighter green font indicates that a negative risk bonus signal (mostly a negative effect on trials with a positive risk bonus) was present in vmPFC and this was associated with increased coupling with PCC when the safer choice was taken; in other words, as the risk bonus decreased, coupling of vmPFC with PCC increased when the safer choice was made. The paler red writing indicates that the signal representing the relative value of the riskier choice ($V_{riskier}$-$V_{safer}$) in dACC was not coupled with the PCC in the same decisions. B) The network for the generation of riskier choices. Now the brighter red font indicates that the PCC is more coupled with dACC, compared to vmPFC, as a function of $V_{riskier}$-$V_{safer}$ and left IFG activity.*

Crucially, the competition between the two mechanisms associated with vmPFC and dACC was modulated by a third frontal region, the IFG (Figure 10 and 11). The IFG has often been identified with executive control (Swann et al., 2009) and the current results suggest one way in which it might exert control is to regulate the relative activity in two parallel systems for decision-making and the manner in which they interact with PCC. How exactly IFG is involved in the evaluative process itself, is still unclear. Our results suggest that if it has a causal role in promoting non-default riskier choices then its disruption would lead to taking safer, default choices. In agreement with the possibility that IFG or an adjacent lateral frontal region is involved in dynamic, context dependent changes in decision making, one recent study applied transcranial magnetic stimulation (TMS) in this vicinity and found subjects were more likely to make socially unbiased decisions and to integrate considerations of reward magnitudes in the standard manner (Baumgartner et al., 2011), rather than taking the social context into consideration.

# Chapter 6: Prospective foraging behaviour

“Im Reich der Ideen haengt alles von Begeisterung ab – in der realen Welt ruht alles auf Ausdauer.”

In the realm of ideas everything depends on enthusiasm… in the real world all rest on perseverance

-Johann Wolfgang von Goethe

## Abstract

*The value of an action can be a consequence of its being a part of a longer sequence with a specific goal or lie in the immediate positive outcomes following it. During foraging, longer term value can be due to an understanding of one's own future environments and the actions one is likely to generate in them. Such prospective foraging has received very little attention, in part because it is difficult to isolate a prospective element from a decision. I designed a task that allows me to look at such prospective motivations of foraging independently from other more myopic forms of foraging value and showed that participants were able to estimate such a quantity to guide their choices. Furthermore, they become more biased toward searching as a function of already committed searches suggesting a form of retrospectively driven over-perseverance or effort fallacy in addition to the more prospective influences.*

# Introduction

The decision of whether to engage with a currently encountered offer, or to reject it in order to search elsewhere for better alternatives is ubiquitous in nature and modern life. It is also the crucial comparison in foraging decisions. Although there has been extensive interest in the neural mechanisms of choice and reward guided comparisons (Boorman et al., 2009; Hunt et al., 2012; Rangel & Hare, 2010), only very little work as been done to directly address this unique set of foraging-type decision making problems (Hayden, Pearson, et al., 2011; N. Kolling et al., 2012) despite their clear ecological relevance (Stephens & Krebs, 1986). Foraging decisions, compared to most other choices, require a wealth of environmental information and potential outcomes to be tracked and compared with specific choice and/or goal directed strategies.

Beyond treating rejection, search or foraging as an active, value driven decision with potentially unique underlying neural mechanisms, this perspective generates a variety of new questions about the implementation of different aspects of foraging value. One of the most fundamental is the prospective component of foraging and search that includes an ecological model about future rewards in the world but also potential future actions of the agent within it. In other words, for a sophisticated prospective foraging strategy it is crucial to represent and use ecological models about the current environment. Furthermore, this needs to be combined with an understanding of one's own future behaviour and options in potential future environments, integrating agent and environmental models with each other. Such integration would allow more adaptive decisions in many ecological contexts.  If such a form of decision-making is possible then it may build on pre-existing neural mechanisms for simpler foraging choices. However, it might also require an interaction between the neural mechanisms for foraging that have been describe in previous chapters with neural mechanisms concerned with cognitive control such as those associated with the inferior frontal gyrus (IFG) (Swann et al., 2009) and with sequential organization of behaviour in lateral frontal pole (lateral FP), and choice simulation in dorsomedial prefrontal cortex (dmPFC) (Nicolle et al., 2012) .

In fact, I expect at least two distinct lateral prefrontal brain regions to emerge in the dorsolateral prefrontal cortex (dlPFC) and the lateral FP. Furthermore, parts of IFG, might also participate in modulating the current decision strategy of our subjects, similarly to last chapters results. However, whereas longer term value in the last two chapters was a matter of making a sequence of simple decisions according to a modified decision strategy with lateral FP activity, in this experiment I was particularly interested in the use of information about future decision strategies in conjunction with a variable number of decision opportunities and self-termination. It is possible that this more prospective and free modulation of value by future longer term value additionally recruits other, more lateral prefrontal brain regions such as the dlPFC, in parts to motivate volitional persevering search behaviours.

It has been shown that different environmental parameters (e.g. variance and mean value) and meta-parameters (e.g. volatility) in the environment can be represented in dACC during decision making (Chapter 3) and learning (Behrens et al., 2007), but a lot less is known about the interaction of such environmental reward information with insight about own future decisions to implement truly prospective foraging behaviour.

It is further known that medial frontal cortex and ACC is also concerned with social cognition, i.e. that it might be involved in tracking complex information about not just the non-social reward environment, but also subjective perspectives and social attributes (Behrens et al., 2008). This might allow it to integrate agent-specific with environmental information for the purpose of prospective foraging. Note that the simulation of one's own future actions in order to carry out such a foraging task would require meta-representational processes, the sorts of processes that are prerequisite for theory of mind. The use of such meta-representational abilities to solve such foraging problems may have preceded their use in social interaction.

Additionally, prospective value is a particularly important concept for foraging behaviour, because more complex foraging often requires goal-directed, persistent behaviour to lead to net gains, rather than a focus on immediate pay-offs. Thus, sequential behaviours are crucial for foraging and an essential function of ACC (Procyk et al., 2000; Shidara & Richmond, 2004), but are also associated with other brain regions, including lateral prefrontal cortex. Those areas most likely interact to generate more complex forms of foraging value.

In this experiment I wanted to probe people's insight about the possible future benefits and costs of their own behaviour in foraging style decisions. In other words, they were asked to make prospective foraging decisions that sometimes entailed the possibility of a further sequence of decisions that might yield benefits. This meant the value of the first decision was not exclusively related to the benefits that might immediately follow it but instead its value was related to the states that the agent might enter as a consequence and the value that they might have for procuring a reward outcome. I hypothesize that this particular form of prospective planning i.e. exploiting complex information about the environment and one's possible future actions within it, is one key function of the dACC, in interaction with other lateral prefrontal brain regions such as the dorsolateral prefrontal cortex.

To summarize, neurally, I want to investigate how during a distinct type of decisions, of whether to forage or not, prospective elements and simulations of own future behaviour are computed and integrated with other, more basic elements, such as the potentially immediate gains of foraging within the prefrontal cortex.

For this purpose, I designed a behavioural task to measure prospective foraging behaviour in humans. Subjects made decisions about whether to forage for a better offer or to keep the currently encountered one, factoring different levels of costs of search (similar to the task in Chapter 2). Crucially, in this experiment the properties of the alternatives, as well as maximal number of opportunities for foraging changed on a trial-by-trial level (Also see Methods). I therefore parametrically varied the potential rewards, opportunities and costs of search independently from each other, allowing me to compute the value of the prospective element of searching and test its behavioural and neural correlates. Furthermore, I also independently manipulated the alternatives complexity and difficulty to generate complex but non-prospective opportunities for searching.

In this task I was particularly interested in whether 1) participants were taking into account the value of the search option as well as its costs and compare it with a current offer. 2) Whether they were, more specifically, also sensitive to the prospective element of the search decisions. 3) How participants' subsequent decisions diverged from the intentions implicit in their initial choices. For example, if they were drawn to an option partly because of the number of subsequent search choices it afforded then did they take full advantage of all those opportunities or were they "under-persevering". Alternatively, are subjects "over-persevering"? Does the fact that they have already spent time engaged in a lengthy sequence of decisions make them persevere with further decisions for longer than they ideally ought to? Such an effect resembles a "sunk cost" effect in economics also observed in ecological environments in other species (Pompilio, Kacelnik and Behmer 2006 Science); 4) and is these effects behaviourally and neurally dissociable from effects of complexity or difficulty of foraging environments.

I both have a model-based estimate of prospective value, as well as a model independent one. Although they are very similar and highly correlated, those two approaches to the analysis illustrate the relationship between our model and the environmental parameters its estimates are based on.

# Methods

## Subjects

23 subjects (11 female), aged 21-36 years, completed the task. They were paid £10 plus a performance-dependent bonus of between £5-11. Ethical approval was given by the Oxford University Central University Research Ethics Committee (CUREC) (Ref-Number: MSD-IDREC-C1-2013-095).

**Training**: Before the main task began, participants were given written instructions and trained on 20 trials with the experimenter in the room, allowed to ask questions. As soon as they understood the task they began the main part of the experiment consisting of 160 trials.

## Experimental Task

Subjects had to make decisions between choice options represented by visual tokens shown on a computer monitor that were associated with different values of monetary payment to be paid at the end of the experiment. On each trial subjects were shown one option that I refer to as the "offer". Its value was defined by the position of the dial of a clock-like stimulus. This was the default option that the subjects would have unless they alternatively decided to search for a better alternative instead. They were told that if they decided to search the alternative would be redrawn from the foraging environment that was also shown on the computer monitor.

If subjects decided to stay with an offer by pressing the corresponding button, they would proceed to the next stage of the trial. If, however, they pressed the other button and thus decided to redraw, one out of the alternatives was presented to them as their new offer. The probability of each alternative being drawn was equal to the proportion of the overall area for the alternatives occupied by that particular alternative. In other words, alternatives with big areas were more likely to be drawn than the ones with small areas.

Every time the subjects opted to search they incurred a fixed cost with such costs varying in three levels, non (0), low (6) and high (12). The costs were indicated to subjects by “Costs:non”, “Costs:low” or “Costs:high” being shown on the computer monitor.

Subjects could maximally search only a fixed number of times indicated by a number next to the alternatives. This factor I refer to as max. searches from now on, varied between trials and effectively limited the number of opportunities for the participant to redraw away from particularly bad offers or in pursued of a particularly good one.

At the decision stage the alternatives and the max. searches were presented together and the costs and current offer were presented separately. The order in which the information was presented was randomized across trials. Each piece of information was presented for a certain duration before the next piece was also revealed, with the duration between presentations drawn randomly from a Poisson distribution that depended on the feature revealed, mean duration for alternatives and max. searches is

4.2s (range 3.8 – 4.8s, std 0.2s), 1.2 s for costs and for the current offer each (range 0.8 – 1.7s, std 0.2s). Only after all three pieces of information had been present could a choice be made when a question mark appeared on the screen.

After the first decision, if the offer was selected, the frame surrounding the offer would turn yellow and the frames around the alternatives turned red to indicate the choice the subject had taken. It remained this way for, on average, 4 seconds (std =1.8s). If however, the search was chosen then, on average after just below 4 seconds (std =1.8s), the new offer would be revealed and subsequently a series of further search decisions could be made at 200ms intervals until all the opportunities for searching had run out or the subject had decided to keep the offer.

Regardless of whether searches had been made or not, in the last stage of the trial the offer currently held was shown alone in the middle of the screen for a further 3.9s (range 2 -7.2s, std 1.3s) after which the outcome was presented (win or lose) for another 1.5s (range 1.1 – 2.1s, std 0.2s). There was always a 50% chance of winning in this stage that was independent of the subjects' previous choices. Subjects were instructed about this contingency and furthermore to ensure its credibility, all past wins and losses were counted and presented to the subjects at the outcome time, to reassure them that losses and wins averaged out at about half/half.

The Inter-trial-interval (ITI) until the next trial started was on average 4 s (range 2 – 7.2s, std 1.3s).

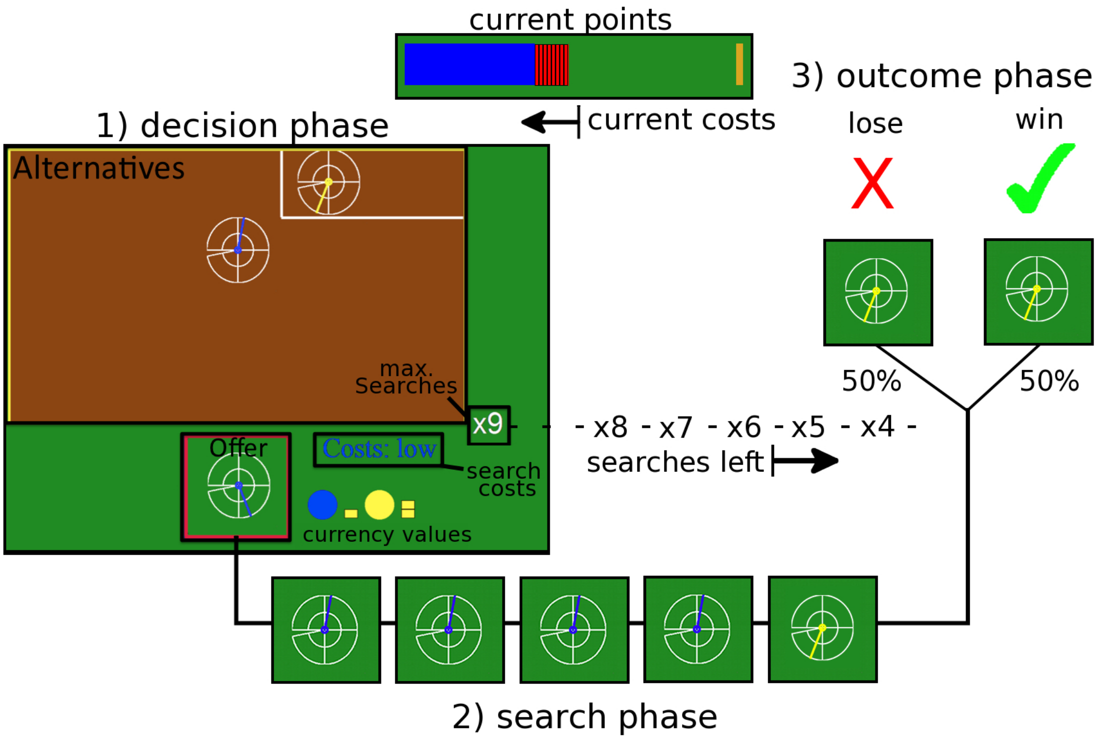

***Figure 1:*** *An illustration of the task using an example trial: At the beginning of every trial subjects saw the alternatives options that they might encounter if they foraged in the brown rectangle in the upper left. In this case the foraging environment consists of two alternatives. One option only occupies approximately a tenth of the foraging environment and so if a subject chose to search then there was a corresponding one in ten chance that they would encounter this option. The other option in the foraging environment occupied nine tenths of the foraging environment and subjects could expect to encounter it on nine tenths of searches. The foraging environment was displayed together with the maximum number of searches allowed (in this example they were allowed nine opportunities to search through the alternatives), the cost of search (in this case the cost is indicated as "low"), as well as the current offer that the subjects would receive if they chose not to search. The maximum number of searches, the search costs, and the current offer were all presented in a randomized order. Subjects could make a*

*choice of whether to stay with the default offer or leave to search in the foraging environment, after a yellow question mark appeared on the screen. The lowest value was indicated by the dial being in a horizontal location on the left with increasing value indicated by the dial rotating counter clock-wise. Since the dial could cover almost the whole circle the most valuable stimulus was closest to the least valuable as the rotation came full circle. Furthermore, the dial itself could be one of two colours, blue or yellow, with yellow stimuli being worth twice as much as their blue counterparts for the same position of the dial. In this example trial the subject chose to search five times, until encountering the less likely but more valuable alternative. With each search the number of remaining searches left decreased by one and the new costs were subtracted from the current point count. After accepting an offer, or when all searches had to been used up, the trial proceeded to the next stage in which the subject either received the reward associated with the stimulus or not, each with a 50% probability. All the way through the trial, participants were reminded about the value of both currencies in the bottom of the screen left and the button assignments on the right (not shown). The black writing is not shown to the subjects and is for illustrative purposes only.*

## Behavioral analysis

I used logistic regression analyses on the choices to determine which factors impacted on a subjects' decision to search/forage as opposed to stay with an offer.

Several terms together reflected the overall and immediate average expected value of opting to search: the average value of all alternatives in the foraging environment (taking into account the likelihood with which they would be encountered), the cost of search, and the value of the offer. Additionally, I used a parameter, which was the interaction between the variance of the values of the alternative options in the foraging environment and the logarithm of the max. searches. This factor can be contrasted with the first three because, rather than being the immediate value of searching they index the prospective value of opting to search in terms of the possible additional further search opportunities that might be available and which would be forgone if the subjects simply opted to take the offer instead. The parameter provides a simple model free estimate of prospectivity. In the next section I explain a slightly more sophisticated index that depends on a particular model of decision-making. Furthermore, I also included, as additional main effects, the number of searches and the variance of the alternatives, as well as the complexity of the alternatives (the number of options it comprised), to test whether these potentially confounding factors were actually driving participants' foraging choices and any neural activity that was recorded.

## Model of prospective foraging

Apart from the model-free logistic regression, I also built a future choice model, which allowed me to estimate the overall expected value of the alternatives factoring in properties such as mean and variance, costs and number of searches.

The model had a simple threshold rule to determine the expected gains of searching assuming a participant would continue searching unless an offer equal or exceeding the currently held threshold was drawn or the maximal number of searches had already been reached. In other words, a subject should not search further but rather take the newly drawn offer ($\text{offer}_{\text{new}}$), i.e. take a STAY choice, if its value exceeded the threshold ($\text{threshold}_{\text{stay}}$) otherwise the subject should make a SEARCH choice and opt to continue.

```
If $\text{Offer}_{\text{new}} > \text{Threshold}_{\text{staying}}$ then STAY
else SEARCH
```

This corresponds to the insight subjects should have that if they have the opportunity to search repeatedly within the foraging environment, they are not necessarily stuck with whatever option they initially draw in the foraging environment. Instead it is, partly the subject's own decision rules that determine the ultimate outcome of searching. Since the set of possible final offers was determined by the participant's behaviour each threshold level will give a unique set of probabilities of receiving a particular value. Taking this set of possible outcomes, discounting by their probabilities, and subtracting the costs, gives the overall expected value of a specific behavioural strategy indexed by its associated threshold.

To find the optimal threshold for every decision, the model selected the threshold that led to the highest average expected value at the end of a trial. (See figure 2 for an example trial with its optimal threshold).

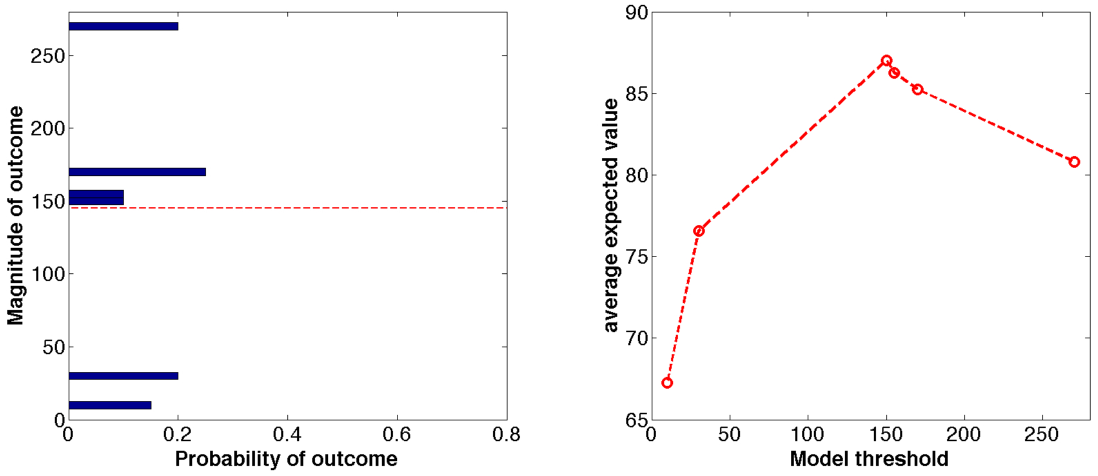

***Figure 2:*** *(Left) example trial with six possible alternative outcomes on the y-axis and their probability on the x-axis. The dotted red line is the optimal threshold for rejection/acceptance of a new offer. If the subject forages in the foraging environment and encounters an offer that exceeds this threshold then the subject should STAY and take the offer. (Right) the average expected value estimated from model by employing all possible thresholds. The model estimate is highest at around 150, which is therefore the model estimate for this particular trial.*

Since I limited the maximum number of searches a participant could make in a trial, participants and the model had to stay with a drawn offer encountered in the foraging environment if they persisted by taking every possible forage opportunity, even if it was below the threshold value. Knowing this and using that information appropriately is key for a principled model based prospective foraging strategy.

Mathematically the prospective part of the expected gain of search is an interaction between the maximal number of allowed searches and the properties of the alternatives such as their mean and variance as well as the costs of searching.

To show how the max. searches manipulation affected the model estimated probability of a specific number of searches being made I plotted the average probability of end states after every possible number of searches separately for each maximum; By definition a line in figure 3, where each colour represents the possible end states for trials involving a particular number of max. searches, needs to end when the maximum has been reached. Note that, therefore, the last possible end state is markedly more likely than the state before. In other words, the model can estimate the odds of being stuck with a lower end offer than it would have excepted if there were no limit to the searches.

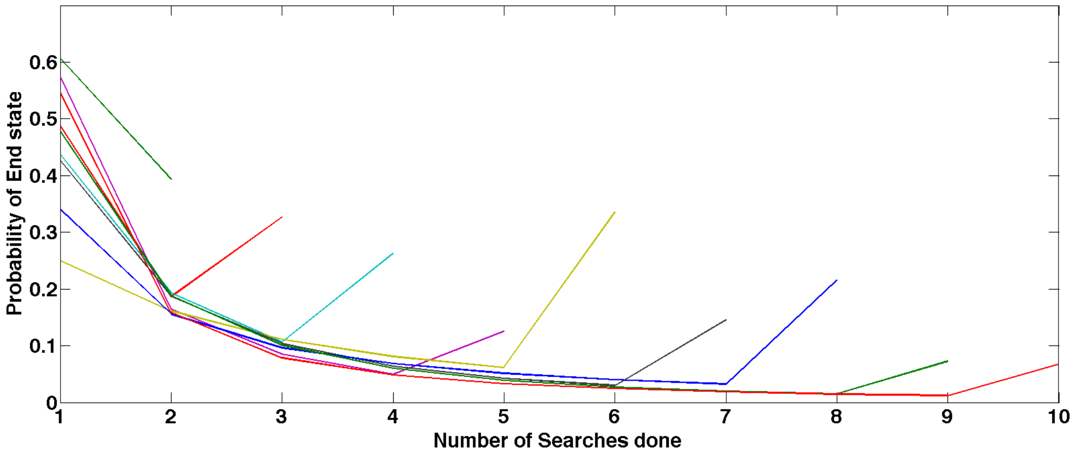

***Figure 3:*** *Average probability of having searched a certain number of times on with a separate line for each different max. searches, one to ten. The spike in probability for each line is due to the maximal number of searches imposing a limit on the number of search decisions.*

The model allows me to compute the overall expected gains given the optimal threshold. By subtracting the myopic/shortsighted expected gains of the outcomes available immediately on current choice itself I can then isolate the prospective component of value. The myopic component itself was computed by using the average expected gains of the first outcome of foraging subtracting its costs.

2) myopic foraging value = (mean(alternatives) – offer)*p(win) - costs

In other words, if the subject only has one potential search left, there is no added benefit to having prospective insight and the prospective element is zero. If however, the possible number of searches is above one then the model takes this into account and its expected gains exceed those of a completely myopic agent. Thus, the difference between having only one search and multiple ones is equal to the prospective element of the model's estimated gains.

Apart from generating a direct estimate of optimal prospective value (pValue) the model also contains a threshold parameter for every decision, which allows me to test whether subjects were sensitive to it, neurally and behaviourally.

Apart from the basic model, with one threshold for search over stay, I also made a version that allowed for variable i.e. collapsing threshold to be fit to the model. This encapsulates the insight that at decision point, you are not stuck with executing all future search decisions using the same threshold value, but instead allow for accepting lower offers in the future, as search opportunities run out. However, as such a model created very highly correlated estimates of prospective value, I will not discuss the variable threshold version in particular here.

## Results

Firstly, I ran a logistic regression analysis on choice behaviour. This revealed a significant impact of all the search value relevant parameters. The offer value led to fewer searches (t(22)=-13.7,p<0.0001) and so did costs of searching (t(22)=-10.83,p<0.0001), whereas the mean value of the alternatives made people more likely to search (t(22)=11.49,p<0.0001) and importantly so did the estimate of prospective value, here represented by a model free simple interaction between logarithm of max searches and the variance of the values of alternatives (t(22)=3.86,p<0.001). Furthermore, the number of searches already done in a trial, made subjects more likely to search, independently from the other value related factors, i.e. having made two (t(22)=5.24,p<0.001), three(t(22)=4.5,p<0.001), or many (many is more than three) searches (t(22)=5.99,p<0.001). Alternatively one can include the number of searches as one parametric regressor into the GLM, rather than treating each number as a separate condition, and also find a significant impact on choices (see Figure 10).

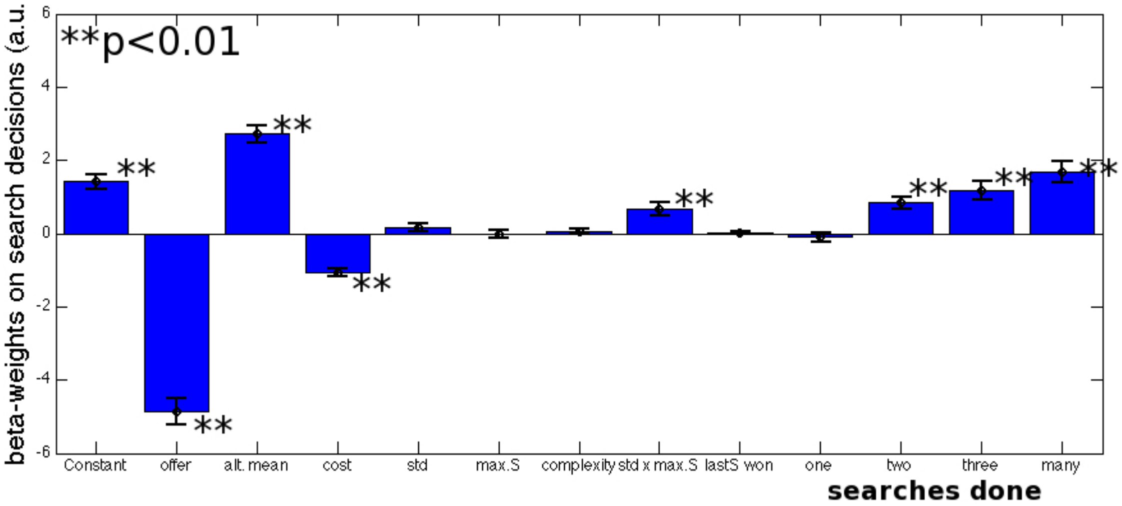

***Figure 4:*** *Logistic regression analysis on all search choices using offer value and the properties of the alternatives. Std refers to the standard deviation of the possible alternatives and max. S is the logarithm of the maximal searches available. Its multiplication (std x max. S) was predictive of search choices, rather than the components in isolation. Additionally, I tested whether the complexity of the alternatives (the number of options) or whether subjects had won their last stimulus after searching, had an impact on their decisions, which it did not. However, the number of already committed searches in a trial had a strong impact on search decisions, here split separately for one, two, three or many (more than three) searches already done.*

To illustrate the effect of the model-based estimate of prospective value on foraging behaviour I binned model-based estimates of prospective value and plotted the proportion of search choices. This shows a very clear increase of search choices with model-based prospective value. I furthermore split the trials into equally sized bins according to different myopic value differences (see Equation 2 in methods), again separating each level of myopic value difference (blue, green, red and cyan) into four equally sized bins of increasing prospective value (1/4 to 4/4 lowest to highest quarter) (Figure 5 right).

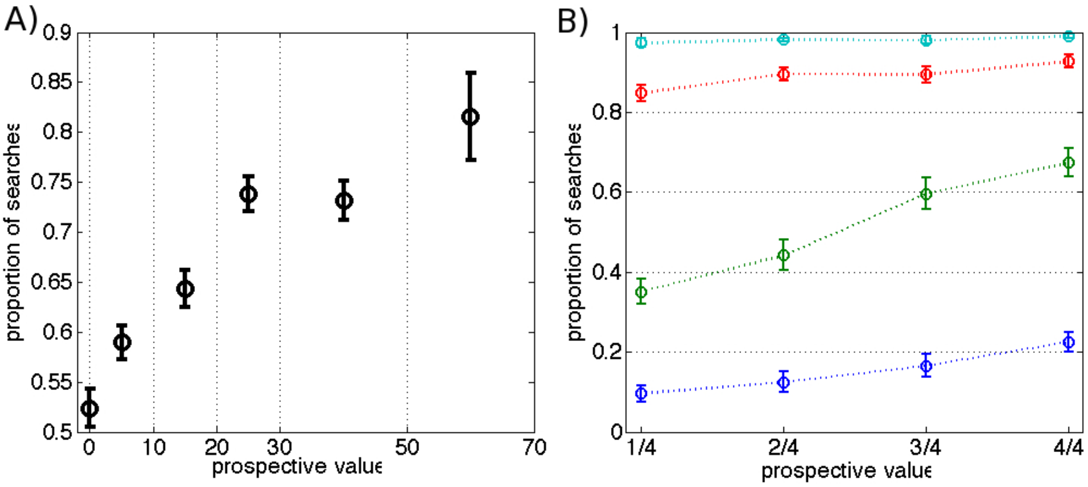

***Figure 5:*** *Binned probability of search plotted against only the prospective value element of the model (A). The effect of model based prospective value (four bins) on the probability of making search decisions for increasing levels of myopic value difference (separate lines) (B). Errorbars are standard error of the mean (SEM) between subjects.*

I also fit a softmax function to the subjects' search behaviour using the model-based estimate of expected value against the current offer's value (offer*p(win)) with a subject specific equivalence point and temperature (Figure 6). Overall the equivalence points were slightly above zero (t(22)=2.09,p=0.0485) [mean=3.37, standard error=1.61]. All Softmax fits are plotted as dotted lines in Figure 6.

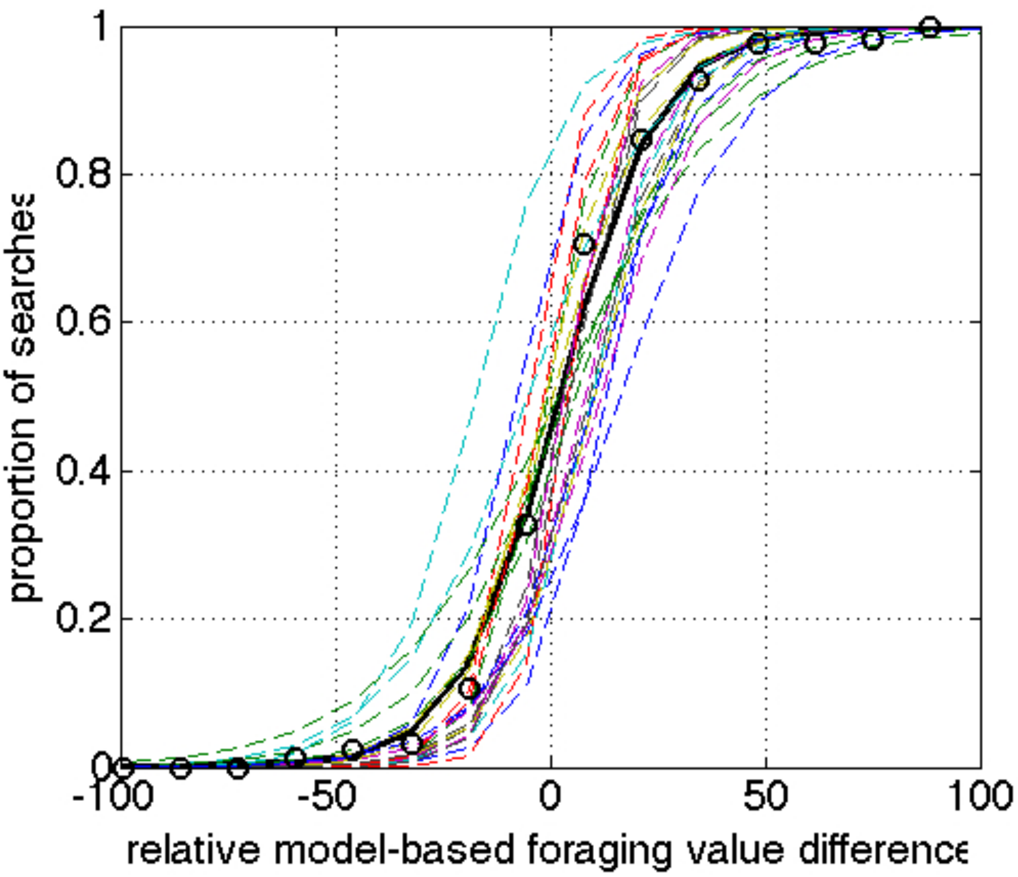

***Figure 6:*** *Softmax curve fitted to each subject individually (thin dotted lines), and for all subjects together (thicker black line) on the overall model-based value difference between foraging and staying with an offer against probability of searching. As a group, subjects were slightly biased against searching.*

To illustrate the effect of searches already done on subsequent search behaviours, i.e. possible over or under-perseverance as a function of past decisions, I plotted the accuracy of participants according to the model (Figure 7), split for trials in which subjects should have searched (red) or stayed (blue). Furthermore, by separating the accuracy by the number of searches done itself, it is clear that, as shown by the GLM above, subjects begin to make more search decisions after having already searched more than once. Figure 7 further shows that according to the model, subjects begin to search more frequently when they should stay, as the accuracy for trials in which they should search begins to diverge from trials in which they should stay after two (t(22)=4.62,p<0.001), three (t(22)=3.18,p<0.001) or many (more than three) searches (t(22)=6.07,p<0.001) (all paired two-tailed t-tests).

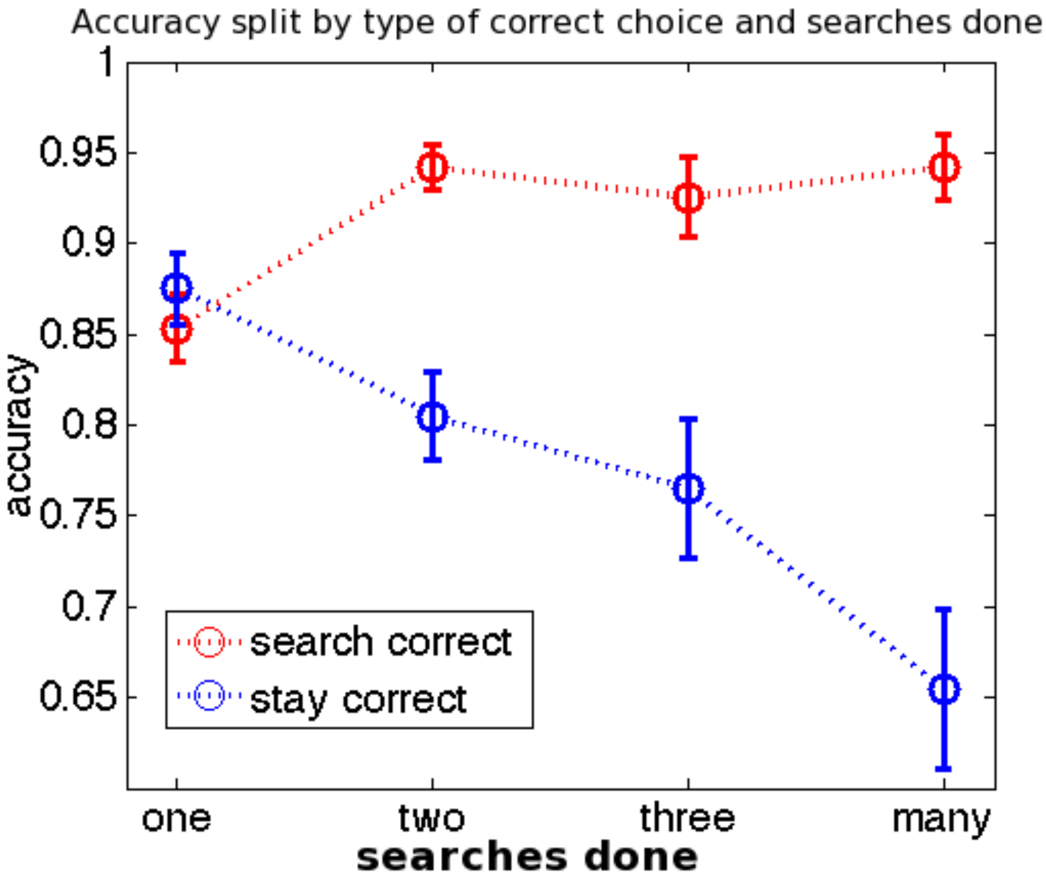

***Figure 7:*** *Accuracy of subjects' decisions when searching (red) or staying (blue) was correct according to the model as a function of searches already done. As more and more searches had already been committed, subjects started making more searching decisions, even when those were incorrect. For this analysis all trials in which participants reached the maximum of searches possible were excluded.*

I ran a separate logistic regression analysis with the model-based parameter of prospective value, and again, offer value, the alternatives' mean value, and the costs of searching. The model-based prospective value parameter was predictive of search decisions independent of the other factors that together comprised the myopic value estimate (Figure 8).

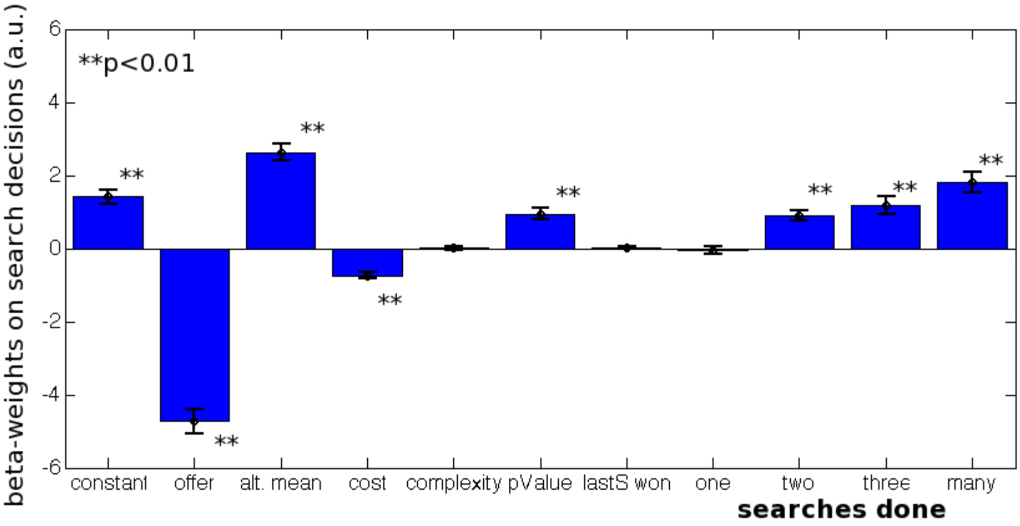

***Figure 8:*** *Logistic regression analysis on all search choices using the model prospective value (pValue), offer value and the other properties of the alternatives. Again, I tested whether the complexity of the alternatives (the number of options) or whether subjects had won their last stimulus after searching, had an impact on their decisions, which it did not. The number of already committed searches had a strong impact on search decisions, here split separately for one, two, three or many (more than three) searches already done.*

Reaction times (RTs), a measure of decision difficulty, on the other hand were affected by different factors. Although I had a relatively long monitoring phase during which no decisions could be executed, participants still took longer to make the first stay/search decision as a function of task parameters. Particularly, when complexity was high (t(22)=7.05,p<0.001), participants were slower. Furthermore, if the mean of the alternatives was higher (t(22)=-9.03,p<0.001), offer value was low (t(22)=6.81,p<0.001) or costs were low (t(22)=2.72,p≤0.01), participants made decisions more rapidly (Figure 9 A). This is in contrast to the later searching decisions (in other words decisions taken after an initial search decision had been taken), in which the complexity still had a big impact on RT (t(22)=7.42,p<0.001), but now only the number of searches already made and prospective value (t(22)=-5.92,p<0.001) appear to further influence RTs (Figure 9 B).

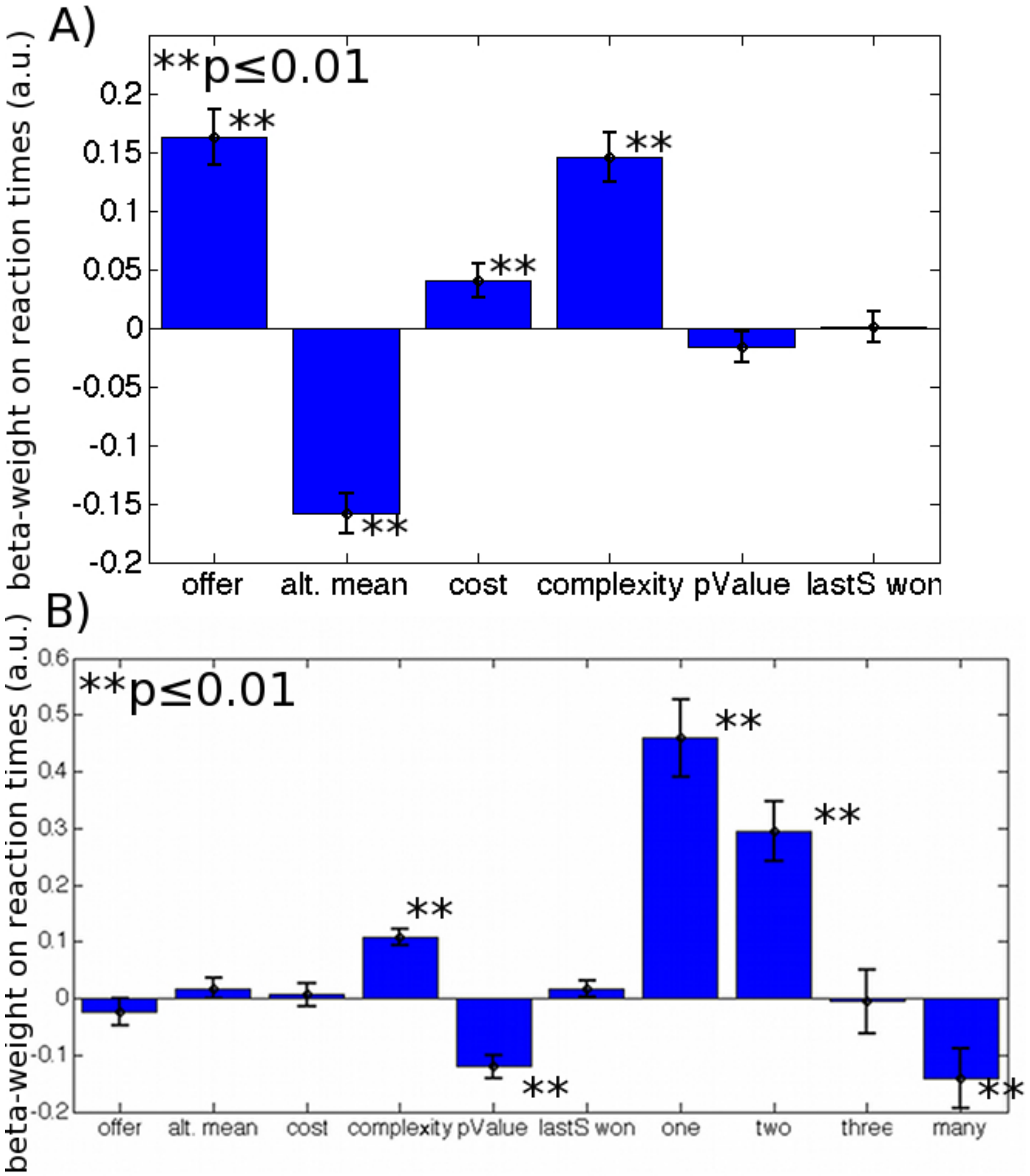

***Figure 9:*** *Multiple regression on the logarithm of reaction times for the first decisions (A). The offer values made subjects slower, as did more complex alternatives i.e. alternatives with more possible outcomes and search costs. On the other hand only a high alternative*

*option mean made participants faster. Analyzing decision reaction times after the first search decision and with only a 200ms delay until choices can be made, show different influences on RT in those events (B). Here, complexity still has an impact on RT and so does the number of searches already made. However, other than that, only prospective value made subjects faster when making their decisions.*

Another way to analyze the behavioural data with the model is using the threshold explicitly. The threshold is the self-imposed point above which future offers will be accepted or not. As such there is an initial threshold at the start of a trial for the first decision, and a change of such a threshold when the model is re-applied after each search. It is important to note that the threshold and threshold change are taken from the model as it is applied to the current choice, i.e. the initial model threshold is subtracted from the current model threshold. In actuality the threshold in the model is a rule for the future outcomes, not the current offer. However, this analysis is trying to test whether subjects stick to an initial rigid threshold from the beginning of the trial until the end, or whether they alternatively, modify their thresholds, i.e. reduce it as they continue searching.

Using this way to analyze the behaviour, offer ($t(22)=-23.7, p<0.001$), initial threshold ($t(22)=15.33, p<0.001$) as well as the threshold change ($t(22)=5.06, p<0.001$) and the logarithm of the number of searches ($t(22)=5.67, p<0.001$) all had a significant impact on search decisions. Note that the result regarding the number of searches is very similar to the effects shown in Figure 4 and 8 and is only meant to illustrate that an alternative approach, one using a logarithm of the number of searches, approximates the effects analyzed separately for different number of searches previously (one, two, three and many).

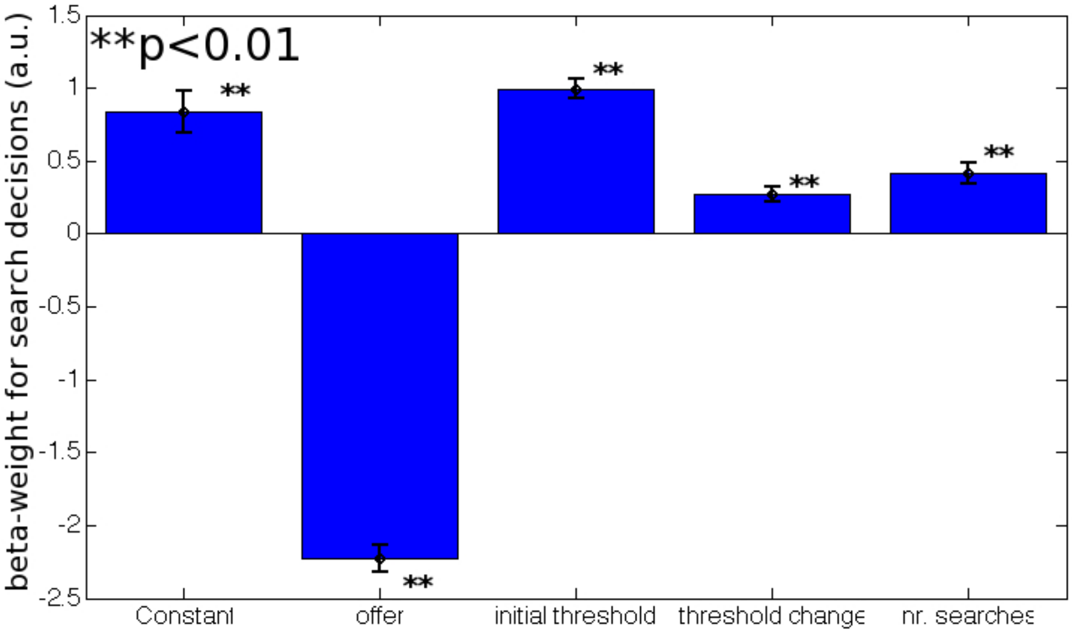

***Figure 10:*** *Logistic regression analysis on all search choices using the model's threshold, as well as the change in threshold during the course of searches. Furthermore, the current offer that is compared to the threshold and logarithm of the number of searches themselves were predictive of search decisions.*

## Outcome stage

Recall that when participants actually opted to stop searching and to take an offer then there was still only a 0.5 probability that they would receive the reward associated with the option. Whether participants were lucky and received the reward did not affect their subsequent choice behaviour on the next trial either by affecting reaction times or decisions. Furthermore, winning after having searched did not make subjects more likely to search in the next trial, as seen in the behavioural GLMs in Figure 4 and 8.

# Discussion

Overall, I found that subjects were able to track the prospective element of value. They were sensitive to the number of forage opportunities available and in a manner that was weighted by the variance of option values in the foraging environment. The effects of prospective value on decision making were dissociable from what were termed the myopic foraging values – the value of the immediately available outcomes and costs on the current trial. There were also dissociable from other, possibly retrospective influences of past searches on search decisions and decision accuracy.

This experiment was designed to investigate prospective foraging in humans during a sequential decision-making task. By making subjects search for their desired goal within different environments that had varying possibilities for using prospective reasoning and variable benefits and costs associated with such behaviour, I was able to show that subjects engaged in prospective foraging to improve their longer term gains. I was furthermore able to dissociate this element of their decision from the overall complexity of the foraging environment. Furthermore, I found that on average participants over-persevered as a function of already having committed searches. Thus, I found prospective as well as retrospective influences on subjects' behaviour that can both be investigate further neurally.

In this paradigm, different from most other decision tasks, a currently encountered offer has to be compared against more elusive concepts of alternative values, goals and whole sequences of actions. Since this kind of decision problem is strongly associated with ecological kinds of behaviours such as those needed for dynamic foraging and planning and which should be ubiquitous in natural and modern environments, solving them well ought to be a primary concern for a neural mechanism of choice. However, most decision making experiments in cognitive neuroscience have focused on a specific set of decision problems in which participants have to compare a set, often two, concrete offers with each other to select one of them, because of its higher utility in one or more dimensions such as reward size, probability or delay. Thus, I will be able to test the neural correlates of those different, direct, and indirect retrospective and prospective decision parameters during foraging decisions and contrast them with the correlates of other types of decisions.

## Prospective foraging

By applying a simple model capturing subjects' prospective insight into future behaviour in a particular environment, I was able to describe participants' choices better than with models that only employed "myopic" value estimates that only considered the most immediate consequences of choices. To do this, I used regressors indexing the value of prospective insight. To approximate behaviour, the model used a simple threshold rule for search/stay decisions on future search outcomes. It assumed that if a criterion was met, participants would decide to stay otherwise they would reject the offer in pursuit for superior alternatives. I was able to show that participants' behaviour could be described appropriately using such a threshold rule and that subjects were sensitive to changes during the course of a trial (see Figure 10).

This prospective foraging perspective can be considered to be a specific form of model-based choice, requiring ecological models of future environments and also of one's own ability to make decisions within it. The strength of our paradigm lies in the fact that the model is based on a specific and ubiquitous decision problem and its components are therefore likely to map onto distinct brain functions (see Chapter 1) dedicated to solving these ecological kinds of choices. Accordingly, these models are based on foraging behaviours, which I have shown to engage a distinct decision making mechanism centred on the dACC.

This paradigm is therefore combining the finding of a possible specific and novel decision making mechanism in dACC, with more sophisticated decision making strategies that require insight into future behaviour and the prospective of future decision opportunities, in the pursuit of longer term value/goals. Excitingly, while the ACC has been implicated in the pursuit of goals over a sequence of actions (Procyk et al., 2000; Shidara & Richmond, 2004), effortful decisions (Croxson et al., 2009; Peter H Rudebeck et al., 2006; M. E. Walton, Kennerley, Bannerman, Phillips, & Rushworth, 2006) as well as in negotiating changing environmental constraints such as risk pressure, to make longer term decisions (Chapter 4 and 5). This however, is not implemented by the dACC alone, but in unison with a frontal polar circuit tracking the need for dynamic change and in competition with other ways of making decisions in the vmPFC.

Similarly my hope is to further understand how prospective foraging is supported by those brain regions using information about the environment and one's own possible future simulated behaviour within it. The neural mechanisms of prospective foraging might require additional areas such as the dlPFC to implement self-terminating volitional sequential foraging behaviours and decision strategies, which have to be maintained over an extended period.

## Over-persevering search

Although participants showed insight into their future opportunities, because they were clearly guided by prospective values, they also, on average, over-persevered when they had pursued their goal through more than one search, suggesting they were not completely rational in their attempts to acquire their goals. This retrospective influence on subjects' behaviour can be clearly seen to effect choices (Figure 4 and 8) as well as to impact on accuracy (Figure 7). However, one caveat of the behavioural analysis is that, it is not possible to dissociate between whether the participants' searches themselves led to increased future searches, or whether differences in subjects' states on some trials which might already have been present at the beginning of a trial, lead to over-persevering behaviours. Thus, neural analysis might be very informative in testing whether there is neural activity at the start of a trial or even the preceding outcome phase that predicts over-perseverance, or whether instead processes during the course of the searches themselves, predict over-persevering behaviours.

Either way, I will be able to look at the representation of information necessary for constructing and using prospective value, testing modifications of decisions by prospective influences, independently from other modifications of decision strategy. Furthermore, I will investigate the contextualization by prior search decisions, and test whether preceding neural activity during earlier decisions or even further in the past, at the last outcome, are in fact predictive of those behavioural effects. If so, I will be able to show the neural correlates of retrospective biases on sequential decision making, when people persevere more than they would if they had not previously been searching.

Conceptually, the retrospective effects found here are similar to sunk costs and effort fallacies observed in behavioural experiments (Pompilio, Kacelnik, & Behmer, 2006; Shidara & Richmond, 2004). Such retrospective sunk cost effects could ecologically be important as long as there is a consistent association between costs and particular outcomes. Then outcomes that require more effort to be experienced also give valuable information about possible rewards, but only if the agent perseveres. Otherwise, no knowledge is gained about potentially very valuable outcomes that could aid in an organism's survival. In other words, outcomes that are not experienced unless much effort is spent are also more rare and rare outcomes are more informative than outcomes that can be sampled very readily. Secondly, in a competitive environment, rewards, which require effort should, especially when they also need a certain amount of minimal skill, be less likely to be depleted than other more readily available alternatives. Thirdly, a bias towards increasingly sequential behaviours, beyond what is otherwise optimal, might serve an improved exploration of new sequential behaviours or maintenance of already initiated sequential strategies. Thus, commitment to sequential behaviours could both serve improved skill acquisition as well as increased success rates of already acquired behaviours. To reiterate, favouring options that one has already invested in could be a mechanism for 1) sampling those options even if they are currently not perceived as on average more valuable, to acquire more costly information and 2), especially in a competitive environment, acquiring outcomes that have a decreased likelihood of depletion by rival foragers. Additionally, 3) they might support a greater reliance on skill acquisition and maintenance, which could ultimately, despite costing more in the short-term, be more beneficial overall.

In theory, although beyond the scope of this specific study, all three arguments could be tested empirically, by varying information content of outcomes, competitiveness or depletion rates of the environment or the skill component of the task in general.

## Conclusion

In this study I was able to show prospective foraging as well as over-perseverance in my human subjects. They clearly took into account both their current environment as well as their future opportunities and costs, when making decisions of whether to search or not. The way they used both types of information in conjunctions, suggests they used prospective value contingent on their own future behaviour given the set of potential outcomes. The next step is to understand how they generate such prospective values neurally and in how far a real simulation of future decisions is made prior to committing to search. I can also look at the effects of such commitment at a neural level, as there was a behavioural effect of having searched before on the current decision.

In the previous chapters, I have argued that the dACC is involved in the integration of non-arbitrary information in a specific and ecologically relevant way for the generation of complex foraging behavior, and other contextualized goal driven behaviours, leading to long term values. In this experiment I showed that participants can make prospective foraging decisions, which should engage the same neural mechanism of choice, as well as rely on additional ecological models possibly computed elsewhere; a strong candidate for such computation of longer term/prospective value is the dlPFC together with the lateral FP, both implicated in enabling more strategic value driven behaviours. Whereas the lateral FP might represent the longer term value of prospective foraging more generally, the dlPFC might be more interested in maintaining voluntary sequential strategies during this self-terminating dynamic foraging.

Overall, the concept of prospective foraging and over-perseverance both pose exciting questions about the roles of lateral and medial prefrontal cortex in the generation of prospective values and foraging. Answering them will hopefully tell us more about how humans are able to implement and maintain such sophisticated behavioural strategies over time.

# Chapter 7: General Discussion

The neural mechanisms of how humans' decide between apples and oranges or other concrete everyday comparisons are increasingly well understood. I, however, was interested in a different set of decision making problems in which a currently encountered offer has to be compared against more elusive concepts of alternative values, the environment, current goals and whole sequences of actions. Such decisions are crucial for ecological kinds of behaviours such as dynamic foraging and planning and ubiquitous in natural and modern environments, solving them well, ought to be a primary concern of neural mechanisms of choice. I drew inspiration from optimal foraging theory, which has frequently concerned itself with those kinds of behavioural problems to understand how animals navigate through an ecological environment adaptively.

Such an approach requires going beyond discussing behavioural change exclusively in terms of simple switching, rejection or disruption of currently performed task sets. It rather proposes an intricate comparison process leading to what I have referred to as foraging or searching, which needs to take into account many different types of information such as environmental value as well as the way values are distributed within the environment. My first two experimental chapters showed that many of signals indexing many of thse aspects of value exist in dACC. Foraging and other search behaviours can often be the result of longer term motivation, perseverance or dynamic contextual decision strategies, rather than being based on random exploration or simple stochasticity. In my third and fourth experimental chapter I was able to show that environmentally driven strategic modification might involve dACC in interaction with other prefrontal areas such as vmPFC, lateral FP and IFG using the concrete example of contextual risk taking. Non-random exploration can be very prospective and deliberate, taking into account future opportunities i.e. decisions as well as the current environment in which future decisions are made. In my last experimental chapter, I was able to show that participants could engage in prospective foraging behaviours, although they also showed a distinct over-commitment to searching once they decided in its favour. This effort fallacy could have a variety of ecological origins, biasing humans increasingly toward sequential perseverance in their pursuit of their longer term goals.

I hope that optimal foraging theory will be able to give a novel perspective to a series of key discoveries within the field, reconciling many past findings especially about the dACC being involved in representing a wealth of information about a changing external reward environment as well as much about how agents' actions are based on considerations of such values as discussed in the Introduction. It also turned out to be very fruitful for generating novel models and concrete hypotheses to test on neural activity for different important cognitive phenomena, such as foraging decisions, environmentally sensitive risk taking or prospectively oriented search.

Regarding the overall organization of the frontal lobes, the results of my thesis suggest not only an interesting perspective on dACC function, but also on the greater network organization within the frontal lobes for the generation of flexible and contextually varied choice behaviours.

In both neural analyses the concept of at least two competing medial prefrontal comparison mechanisms emerged, suggesting one possible way to implement adaptive decisions in the context of multiple competing constraints on behaviour. Different neural systems could concurrently evaluate different courses of actions according to their different reference frames, representations and sources of information. They interact and often compete to in order to determine value-guided choices. Figure 1 suggests a concrete version of one possible competitive interaction between the vmPFC and ACC partly based on my work (Figure 1C). This idea is in contrast with other suggestions of one serial pathway going all the way from sensory processing and valuation to simple comparison between arbitrary options in a universal currency in one unitary decision system, which then triggers the execution of resulting actions (Figure 1A). Such a serial system, despite having great parsimony and elegance, has never received much support from human or animal lesion studies, which suggest very robust choice capacities if only single neural systems are affected, despite there being more specific and subtle decision problems.

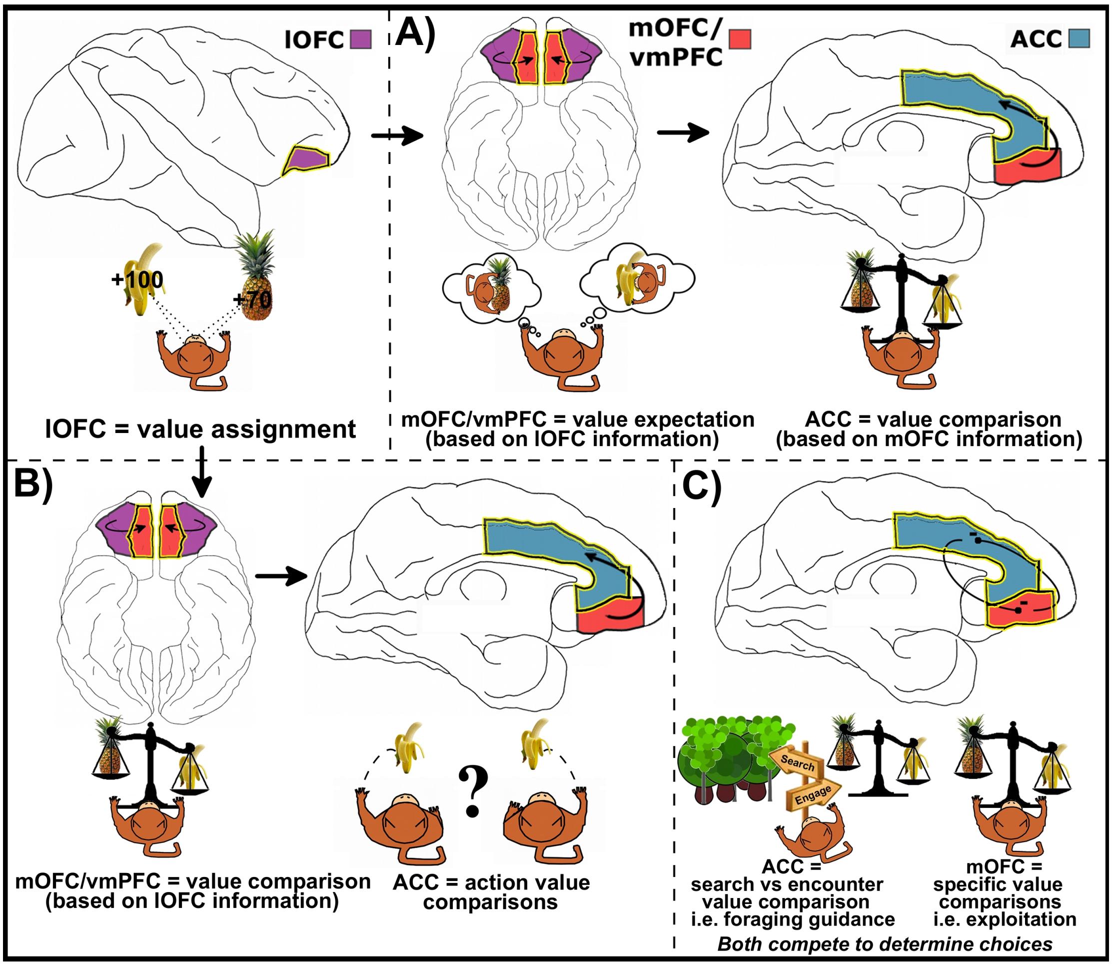

***Figure 1**: Illustration of different theories of prefrontal reward-guided decision mechanisms. A) Proposes that the medial OFC or vmPFC is concerned with representing and imaging value based on learned expectations, possibly generating much of its*

*subjective desirability. Then ACC uses this information to compare all values which each other to derive a decision in a competitive accumulative process. B) Alternatively vmPFC has been suggested to be involved in object-based comparisons, whereas ACC could implement decisions between different actions. C) A new model of parallel decision making systems in the vmPFC and ACC that compete with each other to determine a decision. ACC guides value driven foraging behaviours to optimize longer term value, whereas vmPFC is concerned with comparing different concrete options which each other, integrating across many possible dimensions to find the best offer currently available (Figure taken from* (M. F. Rushworth et al., 2012)*).*

Having made the suggestion of interacting circuits it is crucial to further investigate these interactions and how they are changed and controlled. In this thesis I tried to make some first steps using functional connectivity analysis in chapter 5. I found IFG activity could be related to differences in connectivity between value systems in the vmPFC and dACC and other more posterior motor related areas which could implement specific actions based on choices made.

Furthermore, my thesis has mostly focused on initial foraging decisions, but implementation of longer term behaviour is an equally important part of foraging. I have begun to look at persevering search and prospectively driven foraging in my last experimental chapter, but there are many open questions about the maintenance of foraging behaviours over time. Intriguingly, a lot of evidence from the literature suggests a role of ACC also in the invigoration of foraging behaviours among other multi-stage or sequential goal directed behaviours. The role of neurotransmitters such as dopamine in general invigoration (Niv, Daw, Joel, & Dayan, 2007; Niv, 2007) or incentive salience (Berridge & Robinson, 1998) has long been investigated. Recently, Howe and colleagues showed, using fast-scan-cyclic voltammetry, how tonic or at least ramping dopamine changes in the ventral striatum could underlie motivation toward a distant goal and closely follow proximity and value of such future rewards (Howe Graybiel Nature 2013). This marks a great advance in our understanding of dopaminergic functions. However, how dopamine might be involved in maintaining persistent foraging using the dACC specifically or how the striatum and ACC might interact functionally to implement effortful behaviours, is largely unknown. Additionally, the implementation and strategic modification of foraging behaviours might best be understood by taking into account the already very solid evidence for dACC's role in performance monitoring for the purpose of behavioural adjustments (Markus Ullsperger et al., 2014) in conjunction with the proposals brought forth in this thesis.

In the future, there may be many further implications from a ecological foraging perspective and novel hypotheses that need to be tested, to see whether such an approach is useful for understanding overall human and animal brain function.

Further work will have to go into investigating possible biophysically plausible mechanisms that could underlie proposed dACC function, as has been done very convincingly for other circuits such as the LIP or vmPFC. However, one possible distinction between ACC and other neural networks could be that it does not unconditionally accumulate evidence for or against a certain choice, such as left or rightward sacccades, but instead, constitutes an asymmetric accumulator, which is only engaged in evidence accumulation under very specific circumstances. Understanding what those circumstances are and how evidence accumulation is initiated will be key to predicting ACC activity in general as well as important for predicting the precise details of the temporal profile of its activity.

It will be crucial to investigate the temporal profile of the neural correlates of foraging-style behaviours. Many suggestions in this thesis predict specific and informative temporal profiles, especially during situations of transient competition between different neural systems of evaluation and concurrent network interactions. Additionally, establishing the causal relevance of the neural correlates of decisions using interference methods will be essential. There also needs to be an investigation of the neural correlates and neural activity causal to participants' ability for perseverance and in some cases over-perseverance.

Regardless of whether the proposed perspective will pass the test of time, it has already been useful in generating a new debate over the function of one of the most enigmatic brain regions, anterior cingulate cortex. It has also proved very useful for describing conceptually intriguing types of behaviour and cognitive processes as well as its neural correlates, whether one calls them foraging or not.

# Bibliography

Abe, H., & Lee, D. (2011). Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*, *70*(4), 731–41. doi:10.1016/j.neuron.2011.03.026

Adams, G. K., Watson, K. K., Pearson, J., & Platt, M. L. (2012). Neuroethology of decision-making. *Curr Opin Neurobiol*, *22*(6), 982–989. doi:10.1016/j.conb.2012.07.009

Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, *82(4)*, 463–496.

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–44. doi:10.1038/nn.2921

Amiez, C., Joseph, J. P., & Procyk, E. (2006). Reward encoding in the monkey anterior cingulate cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *16*(7), 1040–55. doi:10.1093/cercor/bhj046

Amiez, C., Sallet, J., Procyk, E., & Petrides, M. (2012). Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration. *NeuroImage*, *63*(3), 1078–90. doi:10.1016/j.neuroimage.2012.06.023

Barraclough, D. J., Conroy, M. L., & Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nature Neuroscience*, *7*(4), 404–10. doi:10.1038/nn1209

Barron, H. C., Dolan, R. J., & Behrens, T. E. J. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience*, *16*(10), 1492–8. doi:10.1038/nn.3515

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21767–72. doi:10.1073/pnas.0908104107

Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., & Fehr, E. (2011). Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nature Neuroscience*, *14*(11), 1468–74. doi:10.1038/nn.2933

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (n.d.). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8039375

Bechara, A., Damasio, H., Damasio, A. R., & Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(13), 5473–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10377356

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science (New York, N.Y.)*, *275*(5304), 1293–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9036851

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (2005). The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends in Cognitive Sciences*, *9*(4), 159–62; discussion 162–4. doi:10.1016/j.tics.2005.02.002

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, *20*(2), 1052–63. doi:10.1016/S1053-8119(03)00435-X

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*(2), 137–52. doi:10.1109/TMI.2003.822821

Beckmann, M., Johansen-Berg, H., & Rushworth, M. F. S. (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization.

*The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(4), 1175–90. doi:10.1523/JNEUROSCI.3328-08.2009

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*, 245–249. doi:10.1038/nature07538

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*, 1214–1221. doi:10.1038/nn1954

Bernacchia, A., Seo, H., Lee, D., & Wang, X.-J. (2011). A reservoir of time constants for memory traces in cortical neurons. *Nature Neuroscience*, *14*(3), 366–72. doi:10.1038/nn.2752

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research. Brain Research Reviews*, *28*(3), 309–69. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9858756

Blakemore, S.-J., & Robbins, T. W. (2012). Decision-making in the adolescent brain. *Nature Neuroscience*, *15*(9), 1184–91. doi:10.1038/nn.3177

Blanchard, T. C., & Hayden, B. Y. (2014). Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*, 646–55. doi:10.1523/JNEUROSCI.3151-13.2014

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*, *62*(5), 733–43. doi:10.1016/j.neuron.2009.05.014

Boorman, E. D., Behrens, T. E., & Rushworth, M. F. (2011). Counterfactual choice and learning in a Neural Network centered on human lateral frontopolar cortex. *PLoS Biology*, *9*. doi:10.1371/journal.pbio.1001093

Boorman, E. D., Rushworth, M. F., & Behrens, T. E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(6), 2242–53. doi:10.1523/JNEUROSCI.3022-12.2013

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 356–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18189009

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control, *108*(3), 624–652. doi:10.1037//0033-295X.I08.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–46. doi:10.1016/j.tics.2004.10.003

Braver, T. S. (2012). The variable nature of cognitive control: a dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–13. doi:10.1016/j.tics.2011.12.010

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science (New York, N.Y.)*, *307*(5712), 1118–21. doi:10.1126/science.1105783

Buckley, M. J., Mansouri, F. A., Hoda, H., Mahboubi, M., Browning, P. G. F., Kwok, S. C., … Tanaka, K. (2009). Dissociable components of rule-guided behavior depend on distinct medial and prefrontal regions. *Science (New York, N.Y.)*, *325*(5936), 52–8. doi:10.1126/science.1172377

Bussey, T. J., Wise, S. P., & Murray, E. A. (2002). Interaction of ventral and orbital prefrontal cortex with inferotemporal cortex in conditional visuomotor learning. *Behavioral Neuroscience*, *116*(4), 703–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12148937

Cai, X., & Padoa-Schioppa, C. (2012). Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(11), 3791–808. doi:10.1523/JNEUROSCI.3864-11.2012

Cai, X., & Padoa-Schioppa, C. (2014). Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron*, *81*(5), 1140–51. doi:10.1016/j.neuron.2014.01.008

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(42), 15048–52. doi:10.1523/JNEUROSCI.3164-11.2011

Caraco, T. (1981). Energy budgets, risk and foraging preferences in dark-eyed juncos (Junco hyemalis). *Behavioral Ecology and Sociobiology*, *8*(3), 213–217. doi:10.1007/BF00299833

Carmichael, S. T., & Price, J. L. (1995). Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *The Journal of Comparative Neurology*, *363*(4), 615–641. doi:10.1002/cne.903630408

Carmichael, S. T., & Price, J. L. (1996). Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *The Journal of Comparative Neurology*, *371*(2), 179–207. doi:10.1002/(SICI)1096-9861(19960722)371:2&lt;179::AID-CNE1&gt;3.0.CO;2-#

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1273796

Chau, B. K. H., Kolling, N., Hunt, L. T., Walton, M. E., & Rushworth, M. F. S. (2014). A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nature Neuroscience*, *17*(3), 463–70. doi:10.1038/nn.3649

Chevrier, A., & Schachar, R. J. (2010). Error detection in the stop signal task. *NeuroImage*, *53*(2), 664–73. doi:10.1016/j.neuroimage.2010.06.056

Christopoulos, G. I., Tobler, P. N., Bossaerts, P., Dolan, R. J., & Schultz, W. (2009). Neural correlates of value, risk, and risk aversion contributing to decision making under risk. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(40), 12574–83. doi:10.1523/JNEUROSCI.2614-09.2009

Clark, L., & Limbrick-Oldfield, E. H. (2013). Disordered gambling: a behavioral addiction. *Current Opinion in Neurobiology*, *23*(4), 655–9. doi:10.1016/j.conb.2013.01.004

Cohen, J. D., Botvinick, M., & Carter, C. S. (2000). Anterior cingulate and prefrontal cortex: who's in control? *Nature Neuroscience*, *3*(5), 421–3. doi:10.1038/74783

Cohen, M. X., & van Gaal, S. (2013). Dynamic interactions between large-scale brain networks predict behavioral adaptation after perceptual errors. *Cerebral Cortex (New York, N.Y. : 1991)*, *23*(5), 1061–72. doi:10.1093/cercor/bhs069

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews. Neuroscience*, *3*(3), 201–15. doi:10.1038/nrn755

Crone, E. A., Wendelken, C., Donohue, S. E., & Bunge, S. A. (2006). Neural evidence for dissociable components of task-switching. *Cerebral Cortex (New York, N.Y. : 1991)*, *16*(4), 475–86. doi:10.1093/cercor/bhi127

Croxson, P. L., Walton, M. E., O'Reilly, J. X., Behrens, T. E. J., & Rushworth, M. F. S. (2009). Effort-based cost-benefit valuation and the human brain. *The Journal of*

*Neuroscience : The Official Journal of the Society for Neuroscience*, *29*, 4531–4541. doi:10.1523/JNEUROSCI.4515-08.2009

Damásio, A. R. (1994). *Descartes' error: emotion, reason, and the human brain*. Quill. Retrieved from http://books.google.co.uk/books/about/Descartes_error.html?id=MocSdi6LXCkC&pgis=1

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–15. doi:10.1016/j.neuron.2011.02.027

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–9. doi:10.1038/nature04766

Dayan, E., Censor, N., Buch, E. R., Sandrini, M., & Cohen, L. G. (2013). Noninvasive brain stimulation: from physiology to network dynamics and back. *Nature Neuroscience*, *16*(7), 838–44. doi:10.1038/nn.3422

De Bruijn, E. R. A., Schubotz, R. I., & Ullsperger, M. (2007). An event-related potential study on the observation of erroneous everyday actions. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 278–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18189001

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), 105–10. doi:10.1038/nn.3279

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science (New York, N.Y.)*, *313*(5787), 684–7. doi:10.1126/science.1128356

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage*, *19*(2 Pt 1), 430–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12814592

Di Pellegrino, G., Ciaramelli, E., & Làdavas, E. (2007). The regulation of cognitive control following rostral anterior cingulate cortex lesion in humans. *Journal of Cognitive Neuroscience*, *19*(2), 275–86. doi:10.1162/jocn.2007.19.2.275

Donoso, M., Collins, A. G. E., & Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. *Science (New York, N.Y.)*, *344*(6191), 1481–6. doi:10.1126/science.1252254

Dorris, M. C., & Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*, *44*(2), 365–78. doi:10.1016/j.neuron.2004.09.009

Dove, A., Pollmann, S., Schubert, T., Wiggins, C. J., & von Cramon, D. Y. (2000). Prefrontal cortex activation in task switching: an event-related fMRI study. *Brain Research. Cognitive Brain Research*, *9*(1), 103–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10666562

Dulawa, S. C., Grandy, D. K., Low, M. J., Paulus, M. P., & Geyer, M. A. (1999). Dopamine D4 receptor-knock-out mice exhibit reduced exploration of novel stimuli. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(21), 9550–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10531457

Dum, R. P., & Strick, P. L. (1996). Spinal cord terminations of the medial wall motor areas in macaque monkeys. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *16*(20), 6513–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8815929

Fellows, L. K. (2006). Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. *Brain : A Journal of Neurology*, *129*(Pt 4), 944–52. doi:10.1093/brain/awl017

Fellows, L. K., & Farah, M. J. (2005). Is anterior cingulate cortex necessary for cognitive control? *Brain : A Journal of Neurology*, *128*(Pt 4), 788–96. doi:10.1093/brain/awh405

Fellows, L. K., & Farah, M. J. (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cerebral Cortex (New York, N.Y. : 1991)*, *17*(11), 2669–74. doi:10.1093/cercor/bhl176

Fiehler, K., Ullsperger, M., & von Cramon, D. Y. (2004). Neural correlates of error detection and error correction: is there a common neuroanatomical substrate? *The European Journal of Neuroscience*, *19*(11), 3081–7. doi:10.1111/j.0953-816X.2004.03414.x

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–55. doi:10.1016/j.neuron.2013.07.006

FitzGerald, T. H. B., Seymour, B., & Dolan, R. J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(26), 8388–95. doi:10.1523/JNEUROSCI.0717-09.2009

Freidin, E., & Kacelnik, A. (2011). Rational choice, context dependence, and the value of information in European starlings (Sturnus vulgaris). *Science (New York, N.Y.)*, *334*(6058), 1000–2. doi:10.1126/science.1209626

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, *6*(3), 218–29. doi:10.1006/nimg.1997.0291

Gan, J. O., Walton, M. E., & Phillips, P. E. M. (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nature Neuroscience*, *13*(1), 25–7. doi:10.1038/nn.2460

Gehring, W., Coles, M., Meyer, D., & Donchin, E. (1990). The error-related negativity: an event-related brain potential accompanying errors. *Psychophysiology*. Retrieved from http://scholar.google.com/scholar?cluster=4425466584515125822&hl=en&oi=scholarr#0

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A NEURAL SYSTEM FOR ERROR DETECTION AND COMPENSATION. *Psychological Science*, *4*(6), 385–390. doi:10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science (New York, N.Y.)*, *295*(5563), 2279–82. doi:10.1126/science.1066893

Gemba, H., Sasaki, K., & Brooks, V. B. (1986). “Error” potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience Letters*, *70*(2), 223–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3774229

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2008). Encoding problem-solving strategies in prefrontal cortex: activity during strategic errors. *The European Journal of Neuroscience*, *27*(4), 984–90. doi:10.1111/j.1460-9568.2008.06048.x

Gläscher, J., Daw, N., Dayan, P., & O’Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–95. doi:10.1016/j.neuron.2010.04.016

Gläscher, J., Hampton, A. N., & O’Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(2), 483–95. doi:10.1093/cercor/bhn098

Glimcher, P. W. (2002). Decisions, decisions, decisions: Choosing a biological science of choice. *Neuron*. doi:10.1016/S0896-6273(02)00962-5

Glimcher, P. W. (2004). *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (p. 375). MIT Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=BlxVBbJAUdoC&pgis=1

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–74. doi:10.1146/annurev.neuro.29.051605.113038

Hanks, T. D., Ditterich, J., & Shadlen, M. N. (2006). Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nature Neuroscience*, *9*(5), 682–9. doi:10.1038/nn1683

Hare, T. a, Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science (New York, N.Y.)*, *324*(5927), 646–8. doi:10.1126/science.1168450

Hare, T. a, O'Doherty, J., Camerer, C. F., Schultz, W., & Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(22), 5623–30. doi:10.1523/JNEUROSCI.1309-08.2008

Hare, T. A., Camerer, C. F., Knoepfle, D. T., & Rangel, A. (2010). Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*, 583–590. doi:10.1523/JNEUROSCI.4089-09.2010

Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P., & Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(44), 18120–5. doi:10.1073/pnas.1109322108

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci*, *31*(11), 4178–4187. doi:10.1523/JNEUROSCI.4652-10.2011

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive reward signals in the anterior cingulate cortex. *Science (New York, N.Y.)*, *324*(5929), 948–50. doi:10.1126/science.1168488

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), 933–9. doi:10.1038/nn.2856

Hayden, B. Y., & Platt, M. L. (2010). Neurons in anterior cingulate cortex multiplex information about reward and action. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(9), 3339–46. doi:10.1523/JNEUROSCI.4874-09.2010

Henri-Bhargava, A., Simioni, A., & Fellows, L. K. (2012). Ventromedial frontal lobe damage disrupts the accuracy, but not the speed, of value-based preference judgments. *Neuropsychologia*, *50*(7), 1536–42. doi:10.1016/j.neuropsychologia.2012.03.006

Hikosaka, O., & Isoda, M. (2010). Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. *Trends in Cognitive Sciences*, *14*(4), 154–61. doi:10.1016/j.tics.2010.01.006

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12374324

Holroyd, C. B., & Coles, M. G. H. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex; a Journal Devoted to the*

*Study of the Nervous System and Behavior*, *44*(5), 548–59. doi:10.1016/j.cortex.2007.08.013

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science (New York, N.Y.)*, *310*(5754), 1680–3. doi:10.1126/science.1115327

Hsu, M., Krajbich, I., Zhao, C., & Camerer, C. F. (2009). Neural response to reward anticipation under risk is nonlinear in probabilities. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(7), 2231–7. doi:10.1523/JNEUROSCI.5296-08.2009

Huettel, S. a, Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, *49*(5), 765–75. doi:10.1016/j.neuron.2006.01.024

Hunt, L. T. (2014). What are the neural origins of choice variability? *Trends in Cognitive Sciences*, *18*(5), 222–4. doi:10.1016/j.tics.2014.01.004

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*. doi:10.1038/nn.3017

Hunt, L. T., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2013). Trial-type dependent frames of reference for value comparison. *PLoS Computational Biology*, *9*(9), e1003225. doi:10.1371/journal.pcbi.1003225

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science (New York, N.Y.)*, *302*(5642), 120–2. doi:10.1126/science.1087847

Izquierdo, a. (2004). Bilateral Orbital Prefrontal Cortex Lesions in Rhesus Monkeys Disrupt Choices Guided by Both Reward Value and Reward Contingency. *Journal of Neuroscience*, *24*(34), 7540–7548. doi:10.1523/JNEUROSCI.1921-04.2004

Jessup, R. K., Busemeyer, J. R., & Brown, J. W. (2010). Error effects in anterior cingulate cortex reverse when error likelihood is high. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(9), 3467–72. doi:10.1523/JNEUROSCI.4130-09.2010

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., & Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(23), 7489–96. doi:10.1523/JNEUROSCI.0349-09.2009

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science (New York, N.Y.)*, *338*(6109), 953–6. doi:10.1126/science.1227489

Kacelnik, a, & Bateson, M. (1997). Risk-sensitivity: crossroads for theories of decision-making. *Trends in Cognitive Sciences*, *1*(8), 304–9. doi:10.1016/S1364-6613(97)01093-0

Kahneman, D., & Tversky, A. (2000). *Choices, values, and frames.*

Kang, M. J., Rangel, A., Camus, M., & Camerer, C. F. (2011). Hypothetical and real choice differentially activate common valuation areas. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*, 461–468. doi:10.1523/JNEUROSCI.1583-10.2011

Kelly, R. E., Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murphy, C. F., Morimoto, S. S., … Hoptman, M. J. (2010). Visual inspection of independent components: defining a procedure for artifact removal from fMRI data. *Journal of Neuroscience Methods*, *189*(2), 233–45. doi:10.1016/j.jneumeth.2010.03.028

Kennerley, S. W., Behrens, T. E. J., & Wallis, J. D. (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience*, *14*(12), 1581–9. doi:10.1038/nn.2961

Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J., & Rushworth, M. F. (2006). Optimal decision making and the anterior cingulate cortex. *Nat Neurosci*, *9*, 940–947. doi:10.1038/nn1724

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., & Rushworth, M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience*, *9*(7), 940–7. doi:10.1038/nn1724

Kepecs, A., Uchida, N., Zariwala, H. A., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*(7210), 227–31. doi:10.1038/nature07200

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science (New York, N.Y.)*, *303*(5660), 1023–6. doi:10.1126/science.1089910

Klavir, O., Genud-Gabai, R., & Paz, R. (2013). Functional connectivity between amygdala and cingulate cortex for adaptive aversive learning. *Neuron*, *80*(5), 1290–300. doi:10.1016/j.neuron.2013.09.035

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. *Nature*, *399*(6732), 148–51. doi:10.1038/20178

Koechlin, E., Corrado, G., Pietrini, P., & Grafman, J. (2000). Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(13), 7651–6. doi:10.1073/pnas.130177397

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science*. doi:10.1126/science.1216930

Kolling, N., Wittmann, M., & Rushworth, M. F. S. (2014). Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron*, *81*(5), 1190–1202.

Lau, H. C., Rogers, R. D., Ramnani, N., & Passingham, R. E. (2004). Willed action and attention to the selection of action. *NeuroImage*, *21*(4), 1407–15. doi:10.1016/j.neuroimage.2003.10.034

Lebreton, M., Jorge, S., Michel, V., Thirion, B., & Pessiglione, M. (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*, *64*(3), 431–9. doi:10.1016/j.neuron.2009.09.040

Levy, R., & Goldman-Rakic, P. S. (1999). Association of storage and processing functions in the dorsolateral prefrontal cortex of the nonhuman primate. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*(12), 5149–58. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10366648

Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*, 13214–13223. doi:10.1523/JNEUROSCI.1246-11.2011

Liston, C., Matalon, S., Hare, T. A., Davidson, M. C., & Casey, B. J. (2006). Anterior cingulate and posterior parietal cortices are sensitive to dissociable forms of conflict in a task-switching paradigm. *Neuron*, *50*(4), 643–53. doi:10.1016/j.neuron.2006.04.015

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–78. doi:10.1038/nature06976

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–7. doi:10.1038/35084005

Lu, M. T., Preston, J. B., & Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *The Journal of Comparative Neurology*, *341*(3), 375–92. doi:10.1002/cne.903410308

Luk, C.-H., & Wallis, J. D. (2013). Choice coding in frontal cortex during stimulus-guided or action-guided decision-making. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(5), 1864–71. doi:10.1523/JNEUROSCI.4920-12.2013

Luppino, G., Matelli, M., Camarda, R. M., Gallese, V., & Rizzolatti, G. (1991). Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: an intracortical microstimulation study in the macaque monkey. *The Journal of Comparative Neurology*, *311*(4), 463–82. doi:10.1002/cne.903110403

Magno, E., Foxe, J. J., Molholm, S., Robertson, I. H., & Garavan, H. (2006). The anterior cingulate and error avoidance. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *26*(18), 4769–73. doi:10.1523/JNEUROSCI.0369-06.2006

Mansouri, F. A., Buckley, M. J., & Tanaka, K. (2007). Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science (New York, N.Y.)*, *318*(5852), 987–90. doi:10.1126/science.1146384

Matsumoto, K., Suzuki, W., & Tanaka, K. (2003). Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science (New York, N.Y.)*, *301*(5630), 229–32. doi:10.1126/science.1084204

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, *10*(5), 647–56. doi:10.1038/nn1890

McNamara, J. M., & Houston, A. I. (1992). Risk-sensitive foraging: A review of the theory. *Bulletin of Mathematical Biology*, *54*, 355–378. doi:10.1007/BF02464838

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nat Neurosci*, *16*(4), 479–485. doi:10.1038/nn.3337

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–98. doi:10.1162/jocn.1997.9.6.788

Morecraft, R. J., & Van Hoesen, G. W. (1993). Frontal granular cortex input to the cingulate (M3), supplementary (M2) and primary (M1) motor cortices in the rhesus monkey. *The Journal of Comparative Neurology*, *337*(4), 669–89. doi:10.1002/cne.903370411

Neubert, F. X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. S. (2014). Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron*, *81*, 700–713. doi:10.1016/j.neuron.2013.11.012

Nicolle, A., Klein-Flügge, M. C., Hunt, L. T., Vlaev, I., Dolan, R. J., & Behrens, T. E. J. (2012). An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron*, *75*(6), 1114–21. doi:10.1016/j.neuron.2012.07.023

Niv, Y. (2007). Cost, benefit, tonic, phasic: what do response rates tell us about dopamine and motivation? *Annals of the New York Academy of Sciences*, *1104*, 357–76. doi:10.1196/annals.1390.018

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*, *191*(3), 507–20. doi:10.1007/s00213-006-0502-4

Noonan, M. P., Kolling, N., Walton, M. E., & Rushworth, M. F. S. (2012). Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *European Journal of Neuroscience*, *35*, 997–1010. doi:10.1111/j.1460-9568.2012.08023.x

Noonan, M. P., Mars, R. B., & Rushworth, M. F. S. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. *The Journal of Neuroscience : The Official*

*Journal of the Society for Neuroscience*, *31*(40), 14399–412. doi:10.1523/JNEUROSCI.6456-10.2011

Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., & Rushworth, M. F. S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 20547–20552. doi:10.1073/pnas.1012246107

O'Doherty, J., Critchley, H., Deichmann, R., & Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(21), 7931–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12944524

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), E3660–9. doi:10.1073/pnas.1305373110

Ongür, D., & Price, J. L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cerebral Cortex (New York, N.Y. : 1991)*, *10*(3), 206–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10731217

Padoa-Schioppa, C. (2007). Orbitofrontal cortex and the computation of economic value. In *Annals of the New York Academy of Sciences* (Vol. 1121, pp. 232–253). doi:10.1196/annals.1401.011

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*, 14004–14014. doi:10.1523/JNEUROSCI.3751-09.2009

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annual Review of Neuroscience*, *34*, 333–59. doi:10.1146/annurev-neuro-061010-113648

Padoa-Schioppa, C. (2013). Neuronal origins of choice variability in economic decisions. *Neuron*, *80*(5), 1322–36. doi:10.1016/j.neuron.2013.09.013

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*, 223–226. doi:10.1038/nature04676

Padoa-Schioppa, C., & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*, 95–102. doi:10.1038/nn2020

Padoa-Schioppa, C., & Cai, X. (2011). The orbitofrontal cortex and the computation of subjective value: Consolidated concepts and new perspectives. *Annals of the New York Academy of Sciences*, *1239*, 130–137. doi:10.1111/j.1749-6632.2011.06262.x

Parvizi, J., Rangarajan, V., Shirer, W. R., Desai, N., & Greicius, M. D. (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron*, *80*(6), 1359–67. doi:10.1016/j.neuron.2013.10.057

Parvizi, J., Van Hoesen, G. W., Buckwalter, J., & Damasio, A. (2006). Neural connections of the posteromedial cortex in the macaque. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 1563–1568. doi:10.1073/pnas.0507729103

Passingham, D., & Sakai, K. (2004). The prefrontal cortex and working memory: physiology and brain imaging. *Current Opinion in Neurobiology*, *14*(2), 163–8. doi:10.1016/j.conb.2004.03.003

Paulsen, D. J., Platt, M. L., Huettel, S. A., & Brannon, E. M. (2012). From Risk-Seeking to Risk-Averse: The Development of Economic Risk Preference from Childhood to Adulthood. *Frontiers in Psychology*. doi:10.3389/fpsyg.2012.00313

Pearson, J. M., Heilbronner, S. R., Barack, D. L., Hayden, B. Y., & Platt, M. L. (2011). Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn Sci*, *15*(4), 143–151. doi:10.1016/j.tics.2011.02.002

Pearson, J. M., & Platt, M. L. (2013). Change detection, multiple controllers, and dynamic environments: insights from the brain. *Journal of the Experimental Analysis of Behavior*, *99*(1), 74–84. doi:10.1002/jeab.5

Petrides, M., & Pandya, D. N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(43), 11573–86. doi:10.1523/JNEUROSCI.2419-07.2007

Philiastides, M. G., Biele, G., & Heekeren, H. R. (2010). A mechanistic account of value computation in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(20), 9430–5. doi:10.1073/pnas.1001732107

Phillips, P. E. M., Robinson, D. L., Stuber, G. D., Carelli, R. M., & Wightman, R. M. (2003). Real-time measurements of phasic changes in extracellular dopamine concentration in freely moving rats by fast-scan cyclic voltammetry. *Methods in Molecular Medicine*, *79*, 443–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12506716

Picard, N., & Strick, P. L. (1991). Motor areas of the medial wall: a review of their location and functional activation. *Cerebral Cortex*, *6*(3), 342–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8670662

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*, 9984–9988. doi:10.1523/JNEUROSCI.2131-07.2007

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature*, *400*(6741), 233–8. doi:10.1038/22268

Platt, M. L., & Huettel, S. A. (2008). Risky business: the neuroeconomics of decision making under uncertainty. *Nat Neurosci*, *11*(4), 398–403. doi:10.1038/nn2062

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*(2), 59–63. doi:10.1016/j.tics.2005.12.004

Pompilio, L., Kacelnik, A., & Behmer, S. T. (2006). State-dependent learned valuation drives choice in an invertebrate. *Science (New York, N.Y.)*, *311*(5767), 1613–5. doi:10.1126/science.1123924

Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*(3), 381–90. doi:10.1016/j.neuron.2006.06.024

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.-L., & Dreher, J.-C. (2010). Separate valuation subsystems for delay and effort decision costs. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(42), 14080–90. doi:10.1523/JNEUROSCI.2752-10.2010

Procyk, E., Tanaka, Y. L., & Joseph, J. P. (2000). Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nature Neuroscience*, *3*(5), 502–8. doi:10.1038/74880

Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*(2), 314–25. doi:10.1016/j.neuron.2007.11.031

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews. Neuroscience*, *9*, 545–556. doi:10.1038/nrn2357

Rangel, A., & Hare, T. (2010). Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology*. doi:10.1016/j.conb.2010.03.001

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922. doi:10.1162/neco.2008.12-06-420

Real, L., & Caraco, T. (1986). Risk and Foraging in Stochastic Environments. *Annual Review of Ecology and Systematics*, *17*(1), 371–390. doi:10.1146/annurev.es.17.110186.002103

Remondes, M., & Wilson, M. A. (2013). Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. *Neuron*, *80*(5), 1277–89. doi:10.1016/j.neuron.2013.08.037

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science (New York, N.Y.)*, *306*(5695), 443–7. doi:10.1126/science.1100301

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2008). Frontal cortex subregions play distinct roles in choices between actions and stimuli. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*, 13775–13785. doi:10.1523/JNEUROSCI.3541-08.2008

Rudebeck, P. H., Buckley, M. J., Walton, M. E., & Rushworth, M. F. S. (2006). A role for the macaque anterior cingulate gyrus in social valuation. *Science (New York, N.Y.)*, *313*(5791), 1310–2. doi:10.1126/science.1128197

Rudebeck, P. H., Mitz, A. R., Chacko, R. V, & Murray, E. a. (2013). Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron*, *80*(6), 1519–31. doi:10.1016/j.neuron.2013.09.036

Rudebeck, P. H., & Murray, E. A. (2011). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(29), 10569–78. doi:10.1523/JNEUROSCI.0091-11.2011

Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., & Murray, E. a. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nature Neuroscience*, *16*(8), 1140–5. doi:10.1038/nn.3440

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., & Rushworth, M. F. S. (2006). Separate neural pathways process different decision costs. *Nature Neuroscience*, *9*, 1161–1168. doi:10.1038/nn1756

Rushworth, M. F., Kolling, N., Sallet, J., & Mars, R. B. (2012). Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Current Opinion in Neurobiology*. doi:10.1016/j.conb.2012.04.011

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*, 389–397. doi:10.1038/nn2066

Rushworth, M. F. S., Behrens, T. E. J., Rudebeck, P. H., & Walton, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in Cognitive Sciences*. doi:10.1016/j.tics.2007.01.004

Rushworth, M. F. S., Hadland, K. a, Gaffan, D., & Passingham, R. E. (2003). The effect of cingulate cortex lesions on task switching and working memory. *Journal of Cognitive Neuroscience*, *15*(3), 338–53. doi:10.1162/089892903321593072

Rushworth, M. F. S., Hadland, K. A., Paus, T., & Sipila, P. K. (2002). Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. *Journal of Neurophysiology*, *87*(5), 2577–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11976394

Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–69. doi:10.1016/j.neuron.2011.05.014

Rushworth, M. F. S., Walton, M. E., Kennerley, S. W., & Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends in Cognitive Sciences*, *8*(9), 410–7. doi:10.1016/j.tics.2004.07.009

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(40), 13525–36. doi:10.1523/JNEUROSCI.1747-10.2010

Sallet, J., Camille, N., & Procyk, E. (2013). Modulation of feedback-related negativity during trial-and-error exploration and encoding of behavioral shifts. *Frontiers in Neuroscience*, *7*, 209. doi:10.3389/fnins.2013.00209

Schoenbaum, G., Chiba, A. A., & Gallagher, M. (1998). Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience*, *1*(2), 155–9. doi:10.1038/407

Schoenbaum, G., & Roesch, M. (2005). Orbitofrontal cortex, associative learning, and expectancies. *Neuron*, *47*(5), 633–6. doi:10.1016/j.neuron.2005.07.018

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9054347

Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci*, *27*(31), 8366–8377. doi:10.1523/JNEUROSCI.2369-07.2007

Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(2), 628–33. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40102&tool=pmcentrez&rendertype=abstract

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–40. doi:10.1016/j.neuron.2013.07.007

Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., … Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*(7410), 218–21. doi:10.1038/nature11239

Shidara, M., & Richmond, B. J. (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science (New York, N.Y.)*, *296*(5573), 1709–11. doi:10.1126/science.1069504

Shidara, M., & Richmond, B. J. (2004). Differential encoding of information about progress through multi-trial reward schedules by three groups of ventral striatal neurons. *Neuroscience Research*, *49*(3), 307–14. doi:10.1016/j.neures.2004.03.008

Shima, K., Isoda, M., Mushiake, H., & Tanji, J. (2007). Categorization of behavioural sequences in the prefrontal cortex. *Nature*, *445*(7125), 315–8. doi:10.1038/nature05470

Shima, K., & Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. *Science (New York, N.Y.)*, *282*(5392), 1335–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9812901

Shoham, S., O'Connor, D. H., & Segev, R. (2006). How silent is the brain: is there a "dark matter" problem in neuroscience? *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *192*(8), 777–84. doi:10.1007/s00359-006-0117-6

Simmons, J. M., Minamimoto, T., Murray, E. A., & Richmond, B. J. (2010). Selective ablations reveal that orbital and lateral prefrontal cortex play different roles in estimating predicted reward value. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(47), 15878–87. doi:10.1523/JNEUROSCI.1802-10.2010

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–55. doi:10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208–19. doi:10.1016/j.neuroimage.2004.07.051

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neuroscience*, *17*(7), 995–1002. doi:10.1038/nn.3740

Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. *Evolutionary Behavioral Ecology* (Vol. 121, p. 247). doi:10.2307/2409049

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic coding for cognitive control in prefrontal cortex. *Neuron*, *78*(2), 364–75. doi:10.1016/j.neuron.2013.01.039

Strait, C. E., Blanchard, T. C., & Hayden, B. Y. (2014). Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex. *Neuron*, *82*(6), 1357–66. doi:10.1016/j.neuron.2014.04.032

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science (New York, N.Y.)*, *304*(5678), 1782–7. doi:10.1126/science.1094765

Swann, N., Tandon, N., Canolty, R., Ellmore, T. M., McEvoy, L. K., Dreyer, S., … Aron, A. R. (2009). Intracranial EEG reveals a time- and frequency-specific role for the right inferior frontal gyrus and primary motor cortex in stopping initiated responses. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(40), 12675–85. doi:10.1523/JNEUROSCI.3359-09.2009

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. a, Haney, R. Z., Calu, D. J., Taylor, A. R., … Schoenbaum, G. (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron*, *62*(2), 269–80. doi:10.1016/j.neuron.2009.03.005

Talmi, D., Fuentemilla, L., Litvak, V., Duzel, E., & Dolan, R. J. (2012). An MEG signature corresponding to an axiomatic model of reward prediction error. *NeuroImage*, *59*(1), 635–45. doi:10.1016/j.neuroimage.2011.06.051

Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2008). Neuronal distortions of reward probability without choice. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(45), 11703–11. doi:10.1523/JNEUROSCI.2870-08.2008

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*(2), 1621–32. doi:10.1152/jn.00745.2006

Toni, I., Ramnani, N., Josephs, O., Ashburner, J., & Passingham, R. E. (2001). Learning arbitrary visuomotor associations: temporal dynamic of brain activity. *NeuroImage*, *14*(5), 1048–57. doi:10.1006/nimg.2001.0894

Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G. D., Bonci, A., de Lecea, L., & Deisseroth, K. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science (New York, N.Y.)*, *324*(5930), 1080–4. doi:10.1126/science.1168878

Tsujimoto, S., Genovesio, A., & Wise, S. P. (2012). Neuronal activity during a cued strategy task: comparison of dorsolateral, orbital, and polar prefrontal cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(32), 11017–31. doi:10.1523/JNEUROSCI.1230-12.2012

Turken, A. U., & Swick, D. (1999). Response selection in the human anterior cingulate cortex. *Nature Neuroscience*, *2*(10), 920–4. doi:10.1038/13224

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. doi:10.1152/physrev.00041.2012

Ullsperger, M., & King, J. A. (2010). Proactive and reactive recruitment of cognitive control: Comment on Hikosaka and Isoda. *Trends in Cognitive Sciences*, *14*(5), 191–2. doi:10.1016/j.tics.2010.02.006

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: a dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, *14*(6), 1387–401. doi:10.1006/nimg.2001.0935

Ullsperger, M., & von Cramon, D. Y. (2004). Neuroimaging of performance monitoring: error detection and beyond. *Cortex*, *40*(4-5), 593–604. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15505969

Van Schie, H. T., Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549–54. doi:10.1038/nn1239

Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F., & Huettel, S. A. (2009). Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron*, *62*(4), 593–602. doi:10.1016/j.neuron.2009.04.007

Venkatraman, V., Rosati, A. G., Taren, A. A., & Huettel, S. A. (2009). Resolving response, decision, and strategic control: evidence for a functional topography in dorsomedial prefrontal cortex. *J Neurosci*, *29*(42), 13158–13164. doi:10.1523/JNEUROSCI.2708-09.2009

Vickery, T. J., Chun, M. M., & Lee, D. (2011). Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*, *72*(1), 166–77. doi:10.1016/j.neuron.2011.08.011

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, *411*(6840), 953–6. doi:10.1038/35082081

Wallis, J. D., & Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. *Current Opinion in Neurobiology*, *20*(2), 191–8. doi:10.1016/j.conb.2010.02.009

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–84. doi:10.1016/j.neubiorev.2012.05.008

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*, *65*(6), 927–39. doi:10.1016/j.neuron.2010.02.027

Walton, M. E., Devlin, J. T., & Rushworth, M. F. S. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, *7*, 1259–1265. doi:10.1038/nn1339

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E. M., & Rushworth, M. F. S. (2006). Weighing up the benefits of work: Behavioral and neural analyses of effort-related decision making. *Neural Networks*, *19*, 1302–1314. doi:10.1016/j.neunet.2006.03.005

Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, *36*(5), 955–68. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12467598

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *32*(22), 7528–37. doi:10.1523/JNEUROSCI.6352-11.2012

Williams, Z. M., Bush, G., Rauch, S. L., Cosgrove, G. R., & Eskandar, E. N. (2004). Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nature Neuroscience*, *7*(12), 1370–5. doi:10.1038/nn1354

Wise, S. P. (2008). Forward frontal fields: phylogeny and fundamental function. *Trends in Neurosciences*, *31*(12), 599–608. doi:10.1016/j.tins.2008.08.008

Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron*, *58*(6), 967–73. doi:10.1016/j.neuron.2008.04.027

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–47. doi:10.1016/j.neuroimage.2003.12.023

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, *14*(6), 1370–86. doi:10.1006/nimg.2001.0931

Wunderlich, K., Dayan, P., & Dolan, R. J. (2012). Mapping value based planning and extensively trained choice in the human brain. *Nature Neuroscience*, *15*(5), 786–91. doi:10.1038/nn.3068

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). *Neural computations underlying action-based decision making in the human brain. Proceedings of the National Academy of Sciences of the United States of America* (Vol. 106, pp. 17199–17204). doi:10.1073/pnas.0901077106

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2010). Economic choices can be made using only stimulus values. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 15005–15010. doi:10.1073/pnas.1002258107

Yoshida, W., & Ishii, S. (2006). Resolution of uncertainty in prefrontal cortex. *Neuron*, *50*(5), 781–9. doi:10.1016/j.neuron.2006.05.006