# Mechanisms of Goal Commitment and Pursuit

Eleanor Holton

New College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Trinity 2024

# Abstract

In natural environments, most rewards follow a period of pursuit. This requires the ability to plan over multiple steps, as well as the need for commitment to chosen goals. This thesis examines these two components of goal-directed behaviour: planning and commitment.

In the first half of the thesis, I investigate the psychological and neural mechanisms supporting commitment to selected goals. This is addressed using a combination of behavioural modelling, functional magnetic imaging (fMRI), and a study with lesion patients. I propose that commitment is supported by attentional mechanisms which prioritize completion of the chosen goal at the expense of better alternatives. Among healthy people, differences in goal commitment and goal-directed attention are predicted by sustained activity in the ventromedial prefrontal cortex (vmPFC). Damage to the same neural region reduces commitment to goals, which leads to a performance advantage in settings where people tend to over-persist. Nevertheless, it is discussed how seemingly irrational levels of commitment to goals will be beneficial in many environments.

In the latter half of the thesis, I turn to questions of how planning toward goals is affected by medial prefrontal damage, in a population of lesion patients. Two pre-existing paradigms in computational neuroscience are used to dissociate how damage affects the different cognitive components contributing to planning. I find that lesions to medial prefrontal areas impair planning in the more complex setting (the ‘four-in-a-row’ task), but not in a simpler setting (the ‘two-step’ task). Further investigation of the underlying cognitive components suggests that damage affects the capacity to select all the relevant information for planning in complex environments.

This work was funded by a 4-year studentship from the Wellcome Trust.

# Acknowledgements

Starting a PhD during a global pandemic is an odd experience, and I am hugely grateful to many people for making it so worthwhile despite this strange beginning. I would like to thank Nils Kolling for keeping science exciting and collaborative even over zoom in the middle of lock-down. A big thank you also to Chris Summerfield for welcoming me into the community of his lab and for all the insightful discussions, sage advice, and entertaining pub trips since then. Thank you to Sanjay Manohar for giving me the opportunity to work with patients and for fun discussions about the vmPFC. Finally, I am enormously thankful to Jill O'Reilly for bringing such clarity and imagination to every idea, analysis and challenge throughout (as well as encouraging the never-ending underwater metaphors).

It is hard to over-state my thanks to Jan Grohn, for passing on so much over the last four years, never failing to provide the solutions to seemingly unsolvable problems, and at least pretending to be interested in all my far-fetched interpretations. A big thank you to Yael Niv and her lab for an extremely fun and energising research visit to Princeton. To Irene Echeverria-Altuna for being a wonderful, if fated, podcast co-host (and much more). To Pablo Fernandez Velasco for being a continual source of intelligent feedback and uplifting enthusiasm. I also greatly benefited from discussions with Laurence Hunt, Lilian Weber, Jess Thompson, Miriam Klein-Flugge, Lisa Spiering, and Matthew Rushworth and his lab throughout this PhD. I also wish to thank the Wellcome Trust and the Goodger and Schorstein scholarship for generous funding, and New College for helping in numerous ways over the years. Thank you to Ulrik Lyngs for providing the thesis template (Lyngs 2019).

The greatest thank you to my dad (who intentionally sparked curiousity in intentions), to my mum, Natalie, and Joe - for endless encouragement and insight in so many forms.

Eleanor Holton
New College, Oxford
1 April 2024

# Contents

**Appendices**

# Long Abstract

**Chapter 1** introduces conceptual questions motivating the thesis. Goal-directed behaviour in psychology is classically associated with its flexibility in response to environmental change. In contrast, goal-directed behaviour in disciplines like philosophy and behavioural economics is associated with inflexibility in the face of new information. Specifically, people often show over-commitment to chosen goals when better alternatives emerge. To understand this tension, different mechanistic accounts of goal commitment are explored, as well as reasons why goal perseveration biases might be beneficial in ecological settings. It is proposed that setting a goal changes how subsequent information is processed to prioritise goal-relevant material at the cost of excluding good alternative options. An attentional account predicts that over-persistence with a goal manifests more in reduced flexibility to pursue alternative goals than continued pursuit of a failing goal.

**Chapter 2** introduces neuroscientific background for the thesis, with a focus on the role of ventromedial prefrontal cortex (vmPFC) in goal-driven behaviours given its relevance for later chapters. Three possible contributions of vmPFC to goal pursuit are discussed, including (i) supporting all value-guided choice, (ii) supporting choice reliant on model-based inference, and (iii) supporting the selection of relevant over irrelevant information in the light of the current goal or context. Different types of scenarios are discussed where this capacity for context-dependent information processing will be critical. I contrast the contributions of vmPFC in maintaining selected goals to areas like dorsal anterior cingulate cortex (dACC), which seem critical for longer-term choices to engage or disengage with goals.

**Chapter 3** sets up a new cognitive paradigm for investigating incremental pursuit of goals. The task involves sequential decisions between persisting with a current goal, and abandoning in favour of alternative goals that could promise greater long-term reward. The task design enables the quantification of different forms of pressure on peoples' tendency to persist, such as the rate of progress, or the impact from alternative options at different points in pursuit. Attentional biases are also investigated outside the decision context in an interleaved spatial working memory task. People show over-commitment to current goals. Individuals with stronger goal commitment showed higher goal-directed attention in the interleaved

task between decisions. Increasing goal-directed attention is reflected in the decision process itself: while pursuing a goal, people lose their sensitivity to valuable alternative goals, while remaining more sensitive to changes in the current goal.

**Chapter 4** investigates neural areas supporting goal pursuit in healthy individuals using fMRI. A network of medial prefrontal regions centred on vmPFC shows sustained goal-related activity between decisions. Individual differences in behaviour (both goal-oriented attention and decisions to persist) are predicted by baseline activity in vmPFC. These findings offer a potential neural mechanism underlying goal commitment. Specifically, sustained representations of the current goal in vmPFC bias attention and subsequent choices toward current goals. Value-related activity at the time of decision is also modulated by goal pursuit. Mirroring the behavioural finding that people lose sensitivity to alternative value as they near the goal state, ventral striatum also loses sensitivity to the value of alternative goals (but not the current goal) over the course of goal pursuit.

**Chapter 5** directly builds on the previous chapter by asking whether vmPFC is causally involved in commitment to an existing goal in the face of alternatives. The same paradigm is tested in a population of individuals with brain lesions and age-matched healthy controls. Individuals with damage to the same region of vmPFC are less likely to persist with a chosen goal. Importantly, the relationship between lower goal commitment and damage to this region of vmPFC is not explained by higher stochasticity. These patients choose to abandon their current goal at appropriate times, thereby performing better than patients with damage elsewhere who tend to over-persist with goals.

**Chapter 6** investigates a different element of goal pursuit, namely the ability to plan over multiple steps. The cognitive components of planning are investigated in a population of medial prefrontal lesion patients using two pre-existing paradigms: the two-step task and four-in-a-row task. While the two-step task involves planning over a small horizon of binary choices, the four-in-a-row task presents subjects with a more naturalistic planning framework consisting of long sequences of choices in a complex state space. Medial PFC damage selectively impaired performance in the four-in-a-row paradigm. Using a computational model to investigate the contribution of different cognitive components of planning, mPFC patients were more likely to overlook relevant information on any trial. This was contrasted with relatively preserved capacity to simulate future states, or to recognise good heuristics for choice. These findings are linked to a general theory of mPFC involvement in selecting goal-relevant information and guiding goal-oriented attention.

**Chapter 7** discusses further implications and questions arising from earlier chapters. The relationships between goals, behavioural inflexibility, and state representations are discussed. Different interpretations are explored for the contributions of neural areas to computations supporting goal pursuit.

# 1

# Introduction to cognitive questions

> *Swerve me? The path to my fixed purpose is laid with iron rails, whereon my soul is grooved to run. Over unsounded gorges, through the rifled hearts of mountains, under torrents beds, unerringly I rush!*
>
> —Herman Melville, Moby-Dick

Captain Ahab's 'fixed purpose' is his doomed pursuit of the great whale Moby Dick, which, despite the desperate pleas of his crew, leads him to a watery ruination. As our narrator Ishmael woefully observes, 'There is no folly of the beasts of the earth which is not infinitely outdone by the madness of men.' But what exactly are these inescapable grooves directing Ahab's soul?

Humans have the extraordinary capacity to be captivated by almost any goal. From the fictional Ahab's fixation on catching Moby Dick, to the marathon runner's determination to finish through harrowing pain, there seem to be no constraints on the goals that drive us (although there are many constraints on the goals we achieve). I take goals to refer to prospective outcomes that individuals pursue intentionally (Kruglanski 1996; De Martino and Cortese 2023). Loss of the motivation to pursue goals, such as in clinical apathy, has devastating effects on quality of life (Heron et al. 2019).

This thesis is not concerned with the question of how we choose goals in the first place. Rather, it concerns the cognitive and neural mechanisms which support continued pursuit of goals over time. I will argue that the selection of a goal leads to pervasive changes in processing: information relevant to the goal is prioritised in attention over irrelevant information. I propose that over-commitment to a chosen goal partly stems from this attentional shift whereby good alternative goals are excluded from the set of options considered. Finally, I will argue that these changes depend on sustained activity oriented towards the goal in a particular brain area (ventromedial prefrontal cortex; vmPFC) – the background for which I introduce in **chapter 2**.

In this chapter, I set up some of the cognitive questions motivating the thesis. I start with a simple question – what makes pursuit of goals different from other behaviour? I begin by contrasting traditional accounts of goal-directed behaviour in behavioural psychology which emphasise flexibility, with those in philosophy, and behavioural economics, which emphasise inflexibility. I go on to explore the reasons why we show a certain form of inflexibility in goal pursuit - over-commitment to goals. I then explore various ways in which this over-commitment could be understood in current accounts of decision-making. Finally, I consider how goals are re-evaluated when they fail.

## 1.1 The flexibility of goal-directed action

What makes pursuit of goals different from other behaviour? A good place to start is by comparing intentional pursuit of goals to habitual reward-driven behaviour. Rewards in psychology simply refer to positive outcomes, whether that be food, or money, or particular internal states such as pleasure (Schultz et al. 1997). As I absent-mindedly reach for another biscuit while writing this thesis, I perform an action to attain a reward. However, it is not an action to reach a goal of biscuit eating. In fact, I had set myself the opposite goal of biscuit avoiding ten minutes earlier. What makes pursuit of goals distinct from this habitual behaviour to attain rewards? One answer to this is flexibility.

Animals learn over time to perform actions for rewards, a process known as operant conditioning. Thorndike's 'Law of Effect' describes how this learning takes place: actions or behaviours leading to rewarding outcomes are more likely to be repeated, whereas actions leading to no reward or punishment are weakened over time (Thorndike 1898). Crucially, learning action only through direct reinforcement is highly inflexible. If the environment changes, or the outcomes are no longer rewarding, an agent who only learns through reinforcement will need direct experience to relearn actions.

Yet animals are capable of flexibly behaving in new ways to attain rewards. Nothing demonstrates this better than the experiments of Edward Tolman (Tolman and Honzik 1930; Tolman et al. 1946). In classic studies, rats were trained to find food in a maze (Fig.1.1). After the animals learned to navigate the maze, the original path to the food reward was blocked, while 18 new (unexperienced) paths were presented – one of which pointed directly at the location which previously held reward. Despite never having taken this route before, over a third of animals immediately selected the path which led to the previous goal location (Tolman et al. 1946). In other words, they were able to flexibly adapt action beyond the scope of direct experience in order to reach the goal. This is an example of flexibility arising from the use of a 'cognitive map' – a mental model of the environment allowing behaviour to be adjusted if elements in the environment change.

Tolman's rats were able to re-plan to reach the same goal when the environment changed. Another marker of goal-directed flexibility is the capacity to stop pursuing goals if they are no longer valuable. In devaluation paradigms, subjects learn to associate actions with rewarding outcomes, such as pressing a lever for food. After learning, an outcome is then 'devalued', for example through satiation of the food reward. The critical test is whether the subject will continue to perform the action when its outcome becomes undesirable (Fig.1.2; Adams and Dickinson (1981)). Animals only learning through direct reinforcement will persist with the action,

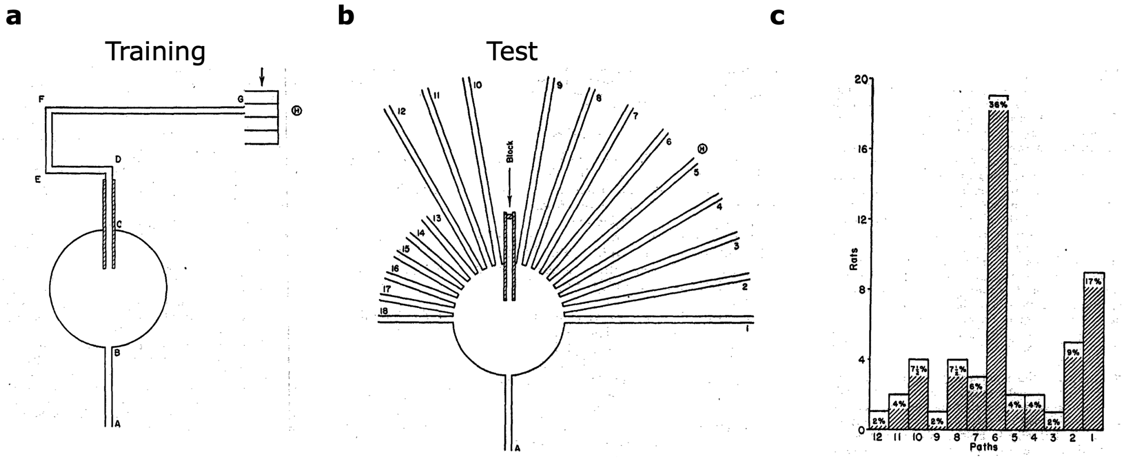


**Figure 1.1:** Tolman and cognitive maps. Reprinted from Tolman et al. (1946). **(a)** Training maze. Rats learned to navigate the maze to reach a food box (marked by the arrow). **(b)** Testing maze. After four days of training on the maze shown in (a), the original pathway was blocked and the animals rats were offered 18 unexperienced paths. One path directly pointed to the previous goal location. **(c)** 36 percent of rats were able to select the path which most directly led to the former goal location of all the offered pathways. This is taken as a classic demonstration of the concept of cognitive maps, namely mental models of the environment which can be flexibly utilised in the service of pursuing goals outside of direct experience.

requiring experience with the devalued outcome to re-learn the action is no longer appropriate. On the other hand, goal-directed animals – sensitive to where they are going – could 'look-ahead' at the outcome through a mental map, and know the action is no longer worth performing (Balleine and Dickinson 1998; Balleine 2005; Dickinson and Balleine 2002; Killcross and Coutureau 2003). This second kind of animal resembles what later became called 'model-based' agents – agents who can make predictions using an explicit model of the environment to select actions best directed to reach desirable goals (Doya et al. 2002; Daw et al. 2005; although see Akam et al. 2015; Russek et al. 2017; Deserno et al. 2015, for alternative accounts of 'model-based' behaviour arising from forms of 'model-free' learning). This is contrasted with the 'model-free' agents described earlier, who are limited by being reactive to direct experience (Doya et al. 2002).

Notably, this goal-directed flexibility comes at a cost. After extensive training in a particular task, animal behaviour transitions from being goal-directed (sensitive to devaluation) to being habitual (insensitive to devaluation; see Fig.1.2c)

(Dickinson 1985; Dickinson and Balleine 2002). One explanation for this is that goal-directed behaviour is slow and costly, requiring animals to utilise model-based prediction. If the environment stays stable for long enough, goal-directed behaviour can revert to habit, which is efficient but inflexible (Keramati et al. 2011; Gershman et al. 2015).

So one answer to the question of how goal-directed behaviour can be identified is through its flexibility. In contrast to habitual reinforced action, behaviour which is in pursuit of a goal must be sensitive to the goal itself (is the goal still desirable?) and the structure of the environment (how do I reach the goal?). I will come back to ideas about model-based and model-free behaviour in **chapter 2**, but for now let's park this topic and turn to a completely different perspective on goals: one which emphasises inflexibility.

## 1.2 The inflexibility of goal-directed action

So far, we have painted a picture of goal-directed behaviour as flexible: adaptable to changes in the environment, as well as changes in the desirability of the outcome itself. However, a contrasting perspective on the character of goal-directed behaviour emerges from other disciplines.

A central concept in philosophy of mind is the notion of intentions: action-guiding mental states which are created by decisions (Bratman 1987; Searle 1980). Intentions have similarities to goals in the sense that they lead to intentional action towards prospective outcomes. However they have two key features which contrast to the goal-directed behaviour discussed earlier. First, intentions are “controlling” in the sense that, similar to habit, they tend to guide behaviour automatically unless interrupted (Bratman 1987). Second, intentions are “stable” in the sense that, unless explicitly revised, they will persist over time. This stability can be understood as a shift in the threshold for re-evaluation – a higher threshold of contradictory evidence may be required to trigger re-evaluation of an intention than the evidence required to form an intention in the first place.

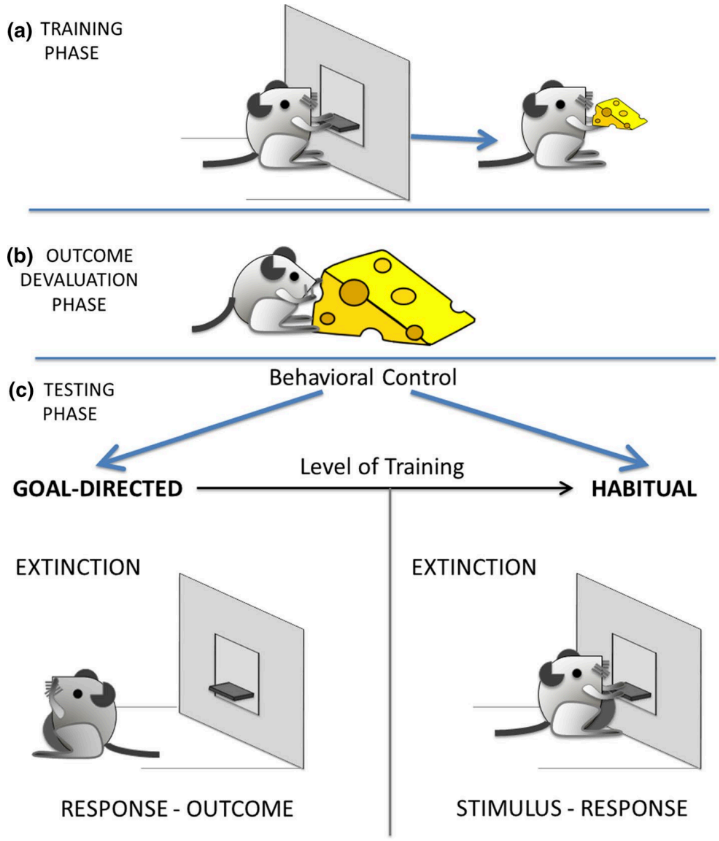


**Figure 1.2:** Measuring goal-directed behaviour in reinforcer devaluation paradigms. Reprinted with permission from Gür et al. (2018). **(a)** In the training phase, animals learn to associate an action (e.g. lever press) with a reward (e.g. food). **(b)** In the devaluation phase, the reward is de-valued for example through satiation. **(c)** In the test phase, the animal is presented with the previously reinforced action. Goal-directed agents will no longer perform the action for the devalued outcome. Habitual agents will continue to perform the action despite the outcome no longer being desirable.

The theory of intentions suggests that a psychological shift occurs after an individual has chosen to pursue a goal, which reduces further deliberation and flexibility. Previously, I contrasted goal-directed behaviour with habit (Balleine and Dickinson 1998). Goal-directed action is flexible and sensitive to changes in the environment; habit-based action is automatic and associated with over-persistence in the face of change. However the philosophical account I just introduced seems to propose that goal-directed behaviour can show these markers of habit – automatic action and over-persistence in the face of change. Where does this idea that goal pursuit reduces flexibility fit into the equation?

The rest of this chapter addresses this question in three parts. First, do people show this behaviour? If so, why would this behaviour emerge? And finally, how could this psychological shift best be reflected in theories of decision-making?

## 1.3 Evidence for intentions

Do people show inflexibility when it comes to re-considering goals? Economists define rationality in terms of making choices that maximise future expected reward (Samuelson 1938; Von Neumann and Morgenstern 1944). A well-established deviation from this in behavioural economics is the sunk cost fallacy: the 'irrational escalation in commitment' following the selection of a particular goal (Arkes and Blumer 1985; Staw 1976). This usually refers to the reluctance to abandon goals after investing effort, time or money, in cases when it would be rational to do so. Similarly observed phenomena emerge in many other disciplines. A closely related idea in the field of aerospace is referred to as 'plan continuation bias', defined in a NASA review as the 'deep-rooted tendency of individuals to continue their original plan of action even when changing circumstances require a new plan' (Berman and Dismukes 2006; Orasanu et al. 1998; Orasanu et al. 2001). It is also possible that sunk cost biases emerge in other animals – escalations in commitment as a function of investing time have been observed in rodents too (Sweis et al. 2018).

In the case of sunk cost biases, irrational levels of commitment are measured after time or effort has already been invested in the goals. However more recent studies have attempted to step away from sunk-cost scenarios by examining commitment before any 'investment' in the goal has even been made. This is the idea that the mere decision to pursue a goal spontaneously induces commitment toward it (Cheng et al. 2023; Chu and Schulz 2022). Cheng and colleagues presented subjects with the task of navigating an avatar towards one of two equidistant and equally rewarding destinations on a grid (Fig.1.3). In some games, the location of the participant's avatar was disrupted to make the non-goal destination a closer target. Participants tended to persist with the original goal despite the alternative

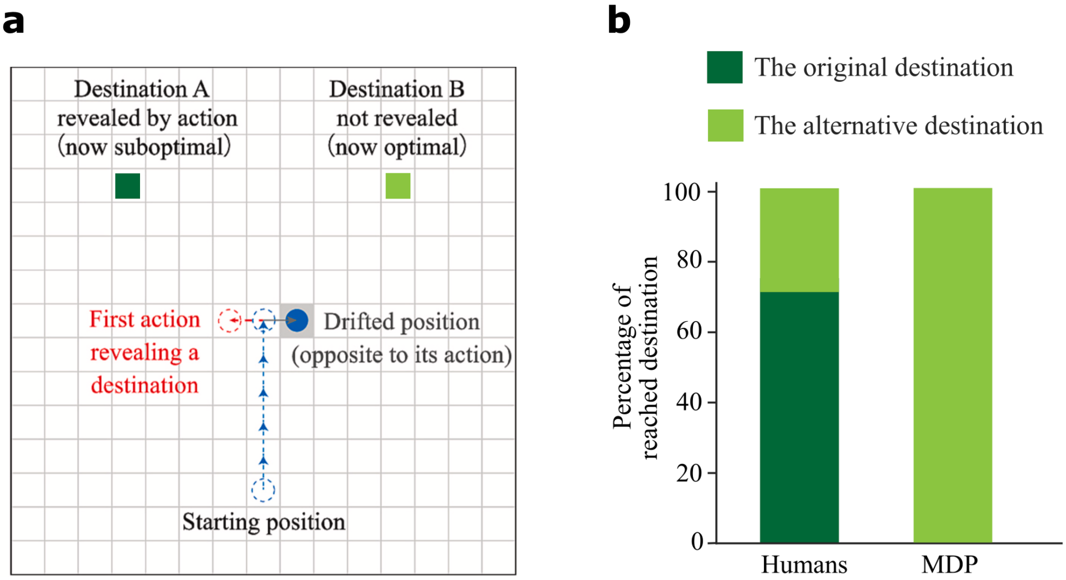


**Figure 1.3:** Spontaneous commitment to goals. Reprinted with permission from Cheng et al. (2023). **(a)** Participants navigated a grid world to reach one of two goals of equivalent distance from their starting point and equal value. Actions could be randomly disrupted to move the avatar to a nearby location on any trial. However, on critical test trials, the disruption changed the location of the avatar to favour the alternative destination. Specifically, on certain trials when participants revealed their intended goal (i.e. moved toward destination A), the location of their avatar was perturbed to be closer to the alternative goal. **(b)** People showed substantial over-commitment to their original goal in the cases of deliberate disruption shown in (a), despite the fact that a rational agent (modeled using a markov decision process) would always switch goals in light of their new location.

goal being the closer and thus more rational choice, and despite having made no progress towards the intended goal.

In another study, participants chose between two goals they would later have to pursue (Chu and Schulz 2022). In one condition, the goals were presented alongside their costs, while in the other condition the costs were only revealed after choosing the goal. For adults and children, making an initial choice led to irrational levels of persistence with the same goal after the costs were revealed, despite the fact there was no penalization for changing one's mind. In these examples, selection of a goal biases subsequent choices towards persisting with the goal even when no time or effort has yet been invested. In other words, goals spontaneously induce commitment.

### 1.3.1 Goal selection and goal pursuit

Between weighing up options, and selecting a goal, what exactly has changed? In the field of social psychology, there has long been interest in the idea of discrete decision-making phases corresponding to (i) evaluation of possible goals, followed by (ii) pursuit of the chosen goal (Klinger 1975; Heckhausen and Gollwitzer 1987; Achtziger and Gollwitzer 2008). In this framework, an initial phase of goal selection might consist of careful evaluation of alternatives in the rational style proposed by economists. However, once a goal is selected, the theory predicts that processing resources are allocated towards implementing the chosen goal rather than deliberating further. This idea has been developed more recently into a theory of goal-driven cognition in neuroscience (O'Reilly et al. 2014; OReilly 2020). Before delving into possible mechanisms, it is worth taking a moment to consider why commitment might be beneficial.

## 1.4 Benefits of goal commitment

Why would goal commitment help? There is a long tradition of understanding irrational economic biases in the context of adaptive pressures (Simon 1955; Sims 2003). There are roughly two ways of approaching this (Gigerenzer and Gaissmaier 2011).

One approach focusses on the role that evolution – the joint constraints of our cognitive abilities and ecological environments – has played in shaping hardwired heuristics and biases (Simon 1957; Gigerenzer and Gaissmaier 2011; Giguère and Love 2013). As Herbert Simon described, "human rational behaviour is shaped by a scissors whose blades are the structure of task environments and the computational capabilities of the actor". This approach, first coined 'bounded rationality', presents these biases as built-in mechanisms of computation honed by evolution.

A subtly different approach considers the use of heuristics itself as a meta-decision, rather than an inflexible product of evolution (Thorngate 1980; Gershman et al. 2015; Shah and Oppenheimer 2008; Payne et al. 1993; Lee and Daunizeau 2021). Considering that deliberation is computationally expensive, whether and how to embark on the process poses a problem at the meta level. The choice of a heuristic strategy in these cases may reflect sensitivity to the trade-offs: lower accuracy but savings on effort. In support of this, various psychological paradigms have found that peoples' choice of 'simpler' heuristics take into account factors affecting the cost/benefit trade-off of the deliberation such as time pressure (Keramati et al. 2016), or the extent to which further deliberation could actually improve their decision (Russek et al. 2022; Callaway et al. 2022).

As Gigerenzer points out, the 'accuracy-effort' trade-off is not always an appropriate description because using heuristics does not always entail sacrifices in accuracy (Gigerenzer and Todd 1999; Gigerenzer and Gaissmaier 2011). A critical idea here is that simple heuristics often out-perform computationally expensive models in complex environments (for example, Czerlinski et al. 1999; Einhorn and Hogarth 1975). This is particularly so in environments with high levels of uncertainty or when relevant information in unknown (Gigerenzer and Gaissmaier 2011). This idea of 'less-is-more' in computation brings the debate full circle back to bounded rationality – sometimes simpler heuristics give rise to better results in the complex environments we evolved in (Beer 1995; Gigerenzer and Todd 1999; Brooks 1991).

With these approaches in mind, I briefly touch on some considerations for biases towards goal commitment in this context.

### 1.4.1 Resource rational explanations

From both the bounded rationality perspective (hardwired biases), and the accuracy-effort trade-off perspective, there are reasons why commitment biases could result from effective use of limited resources (Gershman et al. 2015). In real world environments, there is no limit to the number of alternative goals available for deliberation (Kearns et al. 2002). Given the computational costs associated with deliberation – particularly requiring planning across time – using resources to pursue a selected goal may be adaptive even if good alternative options are sacrificed. It is unclear if commitment biases themselves reflect a meta-decision to trade off accuracy and effort (although for evidence in a similar setting see Lee and Daunizeau (2021)). If commitment biases reflected a meta-decision, we might expect to observe lower biases when there is more time for deliberation, or when goal implementation is inexpensive.

### 1.4.2 Structuring behaviour under uncertainty

Even without limits on cognitive capacity, intentions in philosophy have been argued to play a necessary role in structuring behaviour (Bratman 1987). In the parable of Buridan's Ass, a donkey placed between two equally valuable bales of hay perishes out of indecision. A related phenomenon plagues the lives of many PhD students, who over the course of their PhD discover successively better project ideas only to find at the end of their program that not one project is complete. In this example, rather than oscillate between two options of equal value, the individual shifts successively to new options of higher value, never completing a goal as a result. In noisy and uncertain environments, we can see how this problem will be exacerbated: agents may switch constantly on the basis of noisy estimates which fail to reflect accurate evaluations. How could goals help in these cases? The stability of goal commitment could bring stability of behaviour in worlds where future options are uncertain and evidence is noisy. This may be a case of the 'less-is-more' genre of heuristic – where using a simple rule of thumb favouring goal persistence outperforms more complex computations in ambiguous and uncertain ecological settings.

### 1.4.3 Self-control and resisting temptation

A different angle of philosophical interest in intentions originates from the question of how we exert self-control, specifically in terms of resisting temptation (Bratman 2014; Zelazo et al. 2023). In classic philosophical accounts, self-control involves actively exerting effort to avoid temptation. However, an alternative view is that it involves a capacity to prevent redeliberation in tempting scenarios, by setting intentions. To illustrate this, image you are trying to avoid the temptation of going on social media before bed. One picture of self-control involves heroically abstaining from picking up the phone as it lies enticingly on the bedside table. Another picture involves using pre-emptive action to remove it from the set of options, for example by leaving the phone downstairs. Intentions can be considered the psychological version of the second solution: form the intention to read a book instead and the phone is removed from your set of options.

## 1.5 Mechanisms of goal commitment

So far, I have discussed some empirical work suggesting goal selection induces commitment, and proposed some reasons for why mechanisms of commitment could be helpful. But what psychological mechanisms could realistically support this post-decision shift, while also being favourable from an ecological perspective? Various changes in computation could give rise to a behavioural bias to persist with a goal.

Perhaps after a selecting a goal, decisions continue to proceed in the same way, but the goal option is simply more likely to be preferred. This might manifest in an additive bias on options associated with the current goal compared to any alternatives.

Alternatively, perhaps there is a radical change after selecting a goal where decision-making systems are allocated to implementing the goal. This could manifest as a total re-framing of value to express the usefulness of options with respect to achieving the selected goal. In this case, commitment biases would emerge simply because decision systems are dedicated to achieving the goal, but not re-evaluating the goal itself.

Finally, perhaps the set of options being evaluated changes after goal selection. Individuals might continue to evaluate the chosen goal, but filter information which is irrelevant for goal pursuit. If this were the case, we might expect to see continued sensitivity to the worth of the chosen goal, but reduced sensitivity to alternative options. In the next section I delve into how these three hypotheses manifest in existing frameworks in the literature, and argue that this third option has the most potential as a realistic psychological mechanism.

### 1.5.1 Choice-induced preference change

The first possibility is that goal commitment stems from a post-choice change in preference: simply selecting an option increases the preference for it. In the past, various studies have suggested that the value associated with individual items increases after being chosen (Brehm 1956; Ariely and Norton 2008; Sharot et al. 2009; Sharot et al. 2010). It has been argued that monkeys show this bias too (Egan et al. 2007; although see Chen 2008; Chen and Risen 2009). Classically, the bias was explained in terms of cognitive dissonance (Festinger 1957): having made a choice, agents are motivated to reduce the 'dissonance' associated with the regret of failing to choose the alternative option (although see Lee and Daunizeau 2020, for an alternative explanation of the bias). However, importantly in the case of goal pursuit, a preference for always choosing the same option which extends beyond the boundaries of a goal could not explain goal commitment – once a goal is completed, continuing to over-value previously selected options could hinder pursuit of future goals. More importantly, simply increasing the preference for a current goal does not help with questions of resource allocation. If all possible options continue to be evaluated at every step, fewer resources are available for implementing a goal.

### 1.5.2 Goal-centric accounts

A more radical proposal comes from recent 'goal-driven cognition' theories that situate goals as the driving force in decision-making (Ringstrom 2022; Amir et al. 2024; De Martino and Cortese 2023; Molinaro and Collins 2023; Juechems and Summerfield 2019). In the computational setting of reinforcement learning (RL), agents try to maximise an abstract reward signal, where the concept of reward is kept explicitly vague. Goal-centric accounts propose that the concept of 'reward' itself should be quantified in terms of how much closer options bring us to particular goals or set-points. This could include homeostatic goals, such as avoiding hunger or thirst (Keramati and Gutkin 2014; Juechems and Summerfield 2019), as well as more abstract goals such as fulfilling an intention (Amir et al. 2024). This theory predicts that setting a goal fundamentally changes the meaning of subsequent value: options are valuable insofar as they help to achieve the goal. A similar idea has motivated computational accounts of goal-conditioned reinforcement learning (Kaelbling 1993; Eysenbach et al. 2022; Liu et al. 2022; Plappert et al. 2018). Some authors have focussed on developing RL algorithms that determine their own reward signal based on setting goals (rather than maximising rewards emitted by the environment) which could mimic more naturalistic accounts of motivation in complex environments (Ringstrom 2022). A full review of this computational approach is beyond the scope of this thesis. However, this more radical interpretation of 'goal selection' could also lead to commitment biases: once a goal is set, the goal itself may be the driver of subsequent reward.

### 1.5.3 Change in state representation

The previous section introduced the idea that goal selection could fundamentally change what value is to agents. At its most extreme, value is identified with progress towards the selected goal. However, if all reward signals are now in the service of the goal, how do we escape bad goals? The notion that goals are never re-evaluated once selected is neither practically nor observationally true. A realistic framework for understanding biological agents should be able to deliver prioritization of goal pursuit while retaining the necessary degree of flexibility. The final account I consider proposes that goal selection involves a change in the information represented. In particular, dimensions of the environment that are relevant for achieving the current goal are selected over irrelevant dimensions.

One way of understanding this is through the concept of 'state representation' in the field of reinforcement learning (Wilson et al. 2014). The notion of a state describes the collection of information relevant for an agent to make a decision to maximise reward (see Fig.1.4) (Sutton and Barto 1998). A key challenge for both biological and artificial learning systems is to isolate the elements of complex environments which are predictive of reward in the current task, and which should therefore be included when representing the current state (Wilson and Niv 2011; Bengio et al. 2014).

How do biological agents select the relevant state information? Mechanisms of selective attention allow animals to filter irrelevant sensory input to enhance relevant information (Treisman 1969; Duncan 1984). Attention can be captured both by salient stimuli in the environment or deployed 'top-down' toward current goals (Corbetta and Shulman 2002). Within the learning literature, selective attention has been proposed as a critical means of reducing high dimensional input, by isolating the subset of dimensions important for value ('state representation'), while compressing irrelevant dimensions (Mackintosh 1975; Jones and Canas 2010; Leong et al. 2017). At the centre of this insight is the conceptualization of visuospatial attention as feature weights associated with the allocation of processing resources (Desimone and Duncan 1995). This insight has also been used in models of category learning, where category-irrelevant features are 'compressed' (ignored in attention) to prevent disruption from irrelevant information which could interfere with the correct answer (Love et al. 2004).

Could commitment biases emerge from changes in how information is represented – specifically, selecting information relevant for the current goal at the expense of alternative goals? This theory makes specific predictions about the character of commitment biases: namely there will be greater flexibility if the current goal changes (because goal-relevant information will continue to be represented), but attenuated flexibility in response to goal-irrelevant information (for example, information about good alternative options which is filtered out). There is at least one study in the literature that supports this prediction. As described earlier, individuals show strong goal perseveration when encountering high costs for the current goal relative to other options, after selecting a goal (Chu and Schulz 2022). In this case, the offer of a less effortful alternative did not weaken commitment to the initial goal. However, when the goal itself was devalued (i.e. the outcome of the chosen goal was deemed no longer valuable), participants did not persevere with the goal. While there are various explanations for the effect, the result is consistent with the idea that decisions to abandon a goal may be more flexible in the face of devaluation of the current course of action than offers of better alternatives.

This mechanism of commitment has a few advantages over both previous accounts. First, it fits with a 'resource rational' perspective in the sense that by selectively processing goal-relevant information, resources are primarily allocated to goal pursuit. Second, it enables flexibility in two ways. Selective attention can be graded, such that remarkably salient or valuable alternative options might still capture attention (Lavie 2005). Moreover, continuous evaluation of the goal itself could be possible, even if alternative goals are not being evaluated. In the final section, I discuss how concepts from foraging paradigms could help explain how higher level monitoring of the goal might be performed alongside goal pursuit.

Incorporating attention into models of decision-making has proven powerful in helping to predict observed decision biases elsewhere (Gluth et al. 2020a; Talluri et al. 2018; Hunt et al. 2018). Within the drift diffusion framework of models, decisions are the result of accumulating evidence over time until a predetermined threshold for evidence is met, and a decision is made (Shadlen and Shohamy 2016). The biasing impact of selective attention can be captured as a higher drift rate (faster evidence accumulation) on options currently attended, increasing choices for these options (Krajbich et al. 2010; Krajbich 2019). An example in a similar setting comes from two studies looking at the confirmation bias – reluctance to revise a previous judgement. Both in monkeys and humans, the bias is linked to greater attentional allocation toward belief confirmatory evidence (Hunt et al. 2018; Talluri et al. 2018).

### 1.5.4 Summary of goal commitment

In the previous sections, I discussed three different ways in which goal commitment could be incorporated into our understanding of how decisions are made. First, goal commitment could arise from a simple bias towards preferring options associated with the current goal. By itself, this account does not speak to considerations

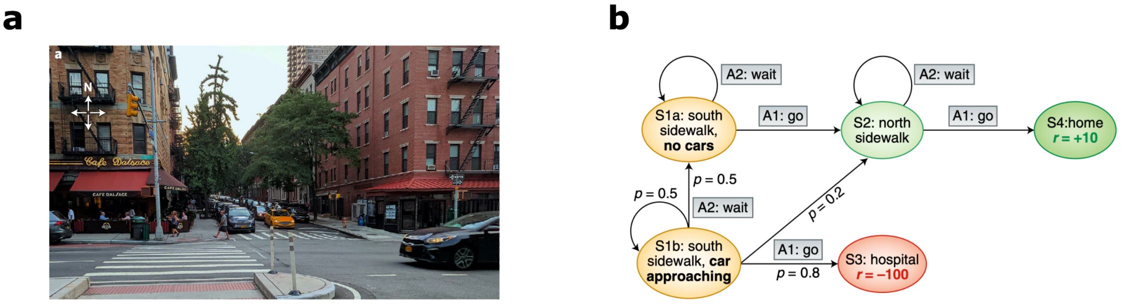


**Figure 1.4:** Task-state representations. Reprinted with permission from Niv (2019). **(a)** Given the goal of crossing a busy New York road, what information should an agent include in a representation of the task? **(b)** An illustration of an appropriate state represention which includes all the necessary information to cross the road, from a reinforcement learning perspective. This includes information about the set of relevant states, transitions between states, available actions, and rewards and punishments associated with each state.

of how commitment could help allocate resources towards goal pursuit. Second, goal commitment could reflect a total overhaul of how we understand reward-driven cognition, such that future rewards are defined in terms of how much closer they bring an agent towards a goal. The problem with this account is that it fails to provide an adequate explanation for how agents could escape bad goals. Third, I considered goal commitment arising from shifts in selective attention, where goal-relevant information is selectively processed at the cost of goal-irrelevant information. In the final two sections, I explore how this third account could merge with considerations about goal pursuit in real-world environments.

## 1.6 Goals in natural environments

At the beginning of this chapter, I introduced the idea that goal-directed behaviour partly rests on the capacity to flexibly plan toward goals. This was discussed with the example of Tolman's rats, who were able to run straight to the goal location without ever previously taking that route. In this section, I want to consider how the ability to use a cognitive model for this kind of goal-directed behaviour will interact with the ability to select relevant information.

In computational terms, using a model of the world to guide action involves the agent simulating possible futures. In psychology and computer science, this process of imagining future trajectories has been characterised as a decision tree, where each decision is a branching point leading to alternative futures (Newell et al. 1959; Huys et al. 2012). Within this space, the 'depth' of the tree corresponds to how far in the future planning extends.

In the real world, the depth of a decision tree is enormous – the number of possible actions we could take at any branch leads to exponential growth of options, and makes full simulation intractable. In order to effectively plan in spaces this large, both people and algorithms must rely on forms of heuristics (Pearl 1988; Sutton and Barto 1998; Koller and Friedman 2009; Russell and Norvig 2016; Gershman et al. 2015). These heuristics include guidelines for selecting which branches to explore (Pearl 1988; Dechter and Pearl 1985; Huys et al. 2012), as well as simplifying the value representations at each node (Sutton and Barto 1998; van Opheusden et al. 2023). An example of the first is a class of search algorithms known as 'best-first search', which guides exploration of the decision-tree by exploring the most promising nodes first (Dechter and Pearl 1985). Another example comes from empirical studies showing that people stop exploring particular branches after encountering losses, a phenomenon known as 'pruning' (Huys et al. 2012).

It is clear how in these complex environments, being able to select relevant dimensions and filter out irrelevant information will be even more critical for goal pursuit. From this perspective, the very same mechanisms which give rise to goal commitment biases might play a critical role in enabling planning in complex environments – specifically, by selecting appropriate dimensions and compressing irrelevant information. In the final chapter of this thesis, I will return to questions of how mechanisms of selection could interact with goal-directed planning in more complex environments.

## 1.7 Escaping goals

Attending to only goal relevant information means that alternative options are more likely to be missed. So the question arises: when we don't know what alternatives we are missing, how do we know whether to be happy with the current goal? In this final section, I consider how studies of natural decisions in foraging paradigms can help provide an answer to this.

In ‘foraging’ approaches, decision-making is investigated in settings that mimic ecological choice problems (Stephens and Krebs 1986). One of the important ideas that has emerged from the foraging literature is that the majority of decisions made in natural environments are not between multiple options at once, but about evaluating options in isolation (Cisek 2012; Kolling et al. 2012; Hayden and Moreno-Bote 2018; Hayden 2018). In classic economic choice paradigms, options are often presented as binary choices between two goods. In contrast, reward opportunities in natural environments tend to be encountered one at a time, in sequence as the animal moves through the world. Often these cases involve much more explicit ignorance about alternative options than the situations where alternatives are merely not attended to. This includes the type of decisions made by ‘foragers’ about whether to stay in a current patch or search for better alternatives as well the type of decisions made by ‘predators’ to pursue a particular prey or wait for better opportunities to come along. At the heart of it, both these ‘stay/switch’ and ‘accept/reject’ decisions are choices concerning whether to continue with the current default or initiate a change (Hayden 2018; Kolling and O’Reilly 2018).

At its limit case, the value of alternatives are completely unknown, and the decision must be performed by comparing the current value of the patch with some long-run estimate of average experience (Stephens and Krebs 1986). An example of this in the case of ‘patch-leaving’ decisions is the marginal value theorem (MVT). Decisions about patch-leaving concern what point to leave a depleting resource (such as a patch of food) and move on to another (see Fig.1.5). MVT predicts that animals should switch to a different patch as soon as the rewards in the current patch dip below the average reward rate across patches (taking into account the costs of traveling) (Charnov 1976). Indeed there is good evidence that animals perform this type of ecological decision near optimally (Wajnberg et al. 2006; Blanchard and Hayden 2015).

Recently, it has been proposed that even when multiple options are presented simultaneously, the brain evaluates them sequentially in the style of foraging choices (Hayden 2018; Hunt et al. 2018; Hunt 2021; Hayden 2018). This has been linked to the way in which primates forage through visual attention (Hunt 2021): anthropoid primates tend to fixate sequentially on options in the environment, selecting or rejecting possible courses of action (Coe 1984).

How is this relevant for theories of goal commitment? The foraging framework predicts that specialised neural structures have evolved to support decisions to disengage from a default option. In the context of goal pursuit, this could help explain how the brain performs continuous evaluation of the goal itself, regardless of whether there is input about the value of alternative options. This gives us a mechanism by which agents can escape from bad goals, while still preferentially committing resources to the current goal.

## 1.8 Summary and Research Questions

In this chapter, I began by asking what makes goal-directed behaviour special from any actions to attain reward. One answer to this was ‘flexibility’, in the sense that goal-directed action (as opposed to habit) is sensitive to changes in the environment and the desirability of the goal. In contrast, there are features of goal-directed behaviour which are ‘inflexible’ – namely that people systematically over-persist with goals in the face of better options.

One possible explanation for this over-persistence is that goal pursuit constrains processing to information relevant for the current goal. From this perspective, I discussed how the same mechanisms underpinning goal commitment might be particularly important for planning goals in complex environments with large state spaces. Finally, I considered how valuation of the current goal could be carried out, even without information about alternative goals (such as in foraging paradigms).

These thoughts lead us to the following research questions, which motivate the main chapters of this thesis:

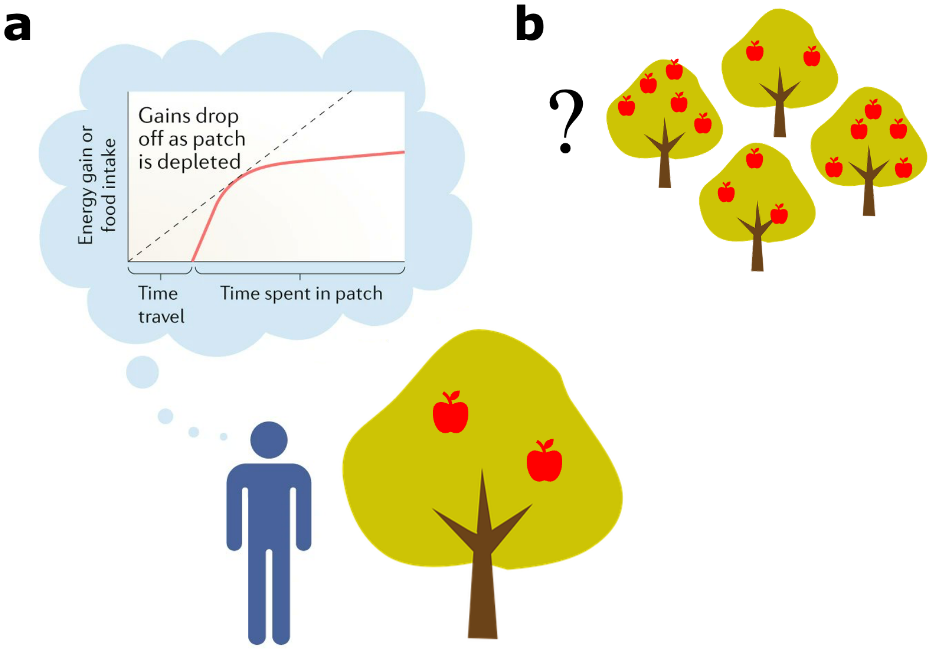


**Figure 1.5:** Patch-leaving tasks. **a**(left) is reprinted with permission from Mobbs et al. (2018). **(a)** In patch-leaving tasks, subjects make decisions about whether to continue exploiting rewards in the current environment (e.g. picking apples), or move to alternative patches. Since resources deplete over time, the decision concerns the question of when to abandon the current patch at the cost of having to travel to a new patch. **(b)** Often in patch-leaving paradigms, alternatives options are not explicitly presented, but decisions to switch patches must be based on other factors like long-run average reward rate. As described in the main text, the marginal value theorem provides a computational framework for how to make optimal abandonment choices despite the lack of explicit knowledge about alternatives.

1. How does goal commitment relate to changes in information processing such as in attention? (**chapter 3**)
2. Are commitment biases more flexible in the face of information about the current goal compared to information about alternative goals? (**chapter 3**)
3. What neural mechanisms support commitment to a goal, and disengagement from goals which are no longer worthwhile? (**chapters 4-5**)
4. What is the relationship between attention and planning towards goals in complex environments? (**chapter 6**)

Exploring decision-making through the lens of goals will be key to understanding how seemingly irrational behavioural biases arise in the context of ecological behaviour. In the next chapter, I return to some of the neuroscience background that will be needed to investigate these ideas.

# 2

# Medial PFC and goal pursuit

In this thesis I include a number of lesion studies focussing on ventral areas of the medial prefrontal cortex (vmPFC; see Fig.2.1). Our focus on this region arose from the results of the first fMRI experiment, in **chapter 4**. As background for the upcoming studies on these lesion patients, here I pose the question of what role this region might be playing in supporting goal-directed behaviour, with reference to some of the ideas developed in the previous chapter.

Lesions to prefrontal cortex have life-altering effects for patients, often leading to devastating consequences for employment, relationships and general lifestyle (Harlow 1868; Schneider and Koenigs 2017). Despite this, pin-pointing the precise cognitive deficits causing these real-world problems has proved an arduous research task (Eslinger and Damasio 1985; Shallice and Burgess 1991; Harlow 1868; Burgess et al. 2000; Tranel et al. 2007). For many years, researchers were limited by the paradoxical finding that catastrophic impairments in navigating the real world often contrasted with perfect performance on classic neuropsychological tests (Burgess et al. 2000).

Some of the earliest observations about the effects of prefrontal cortex damage concerned difficulties with goal-driven behaviour in the real world. In 1868, John

Harlow wrote of Phineas Gage, the first documented patient with prefrontal damage, “he is fitful. . . devising many plans of future operations. . . which are no sooner arranged than they are abandoned.” Similarly, an early example of a cognitive test where prefrontal lesion patients did show impairments is the Multiple Errands Task (Shallice and Burgess 1991; Tranel et al. 2007). This task involves planning and carrying out a series of naturalistic tasks (e.g. completing a shopping list) while abiding by various rules. Damage to vmPFC leads to difficulties structuring behaviour across the various tasks (Tranel et al. 2007).

What could vmPFC be doing to support goal-driven behaviour? This chapter is structured by three possible answers to this question. First, vmPFC damage could be disrupting all value-guided decision. Second, perhaps vmPFC is particularly necessary for constructing value when these decisions require the use of inferred knowledge (or a ‘map’) about the relationship between events. Third, perhaps vmPFC is critical when value is context-dependent - supporting the selection and representation of goal-relevant over goal-irrelevant variables to guide choice. The second answer can be considered the contribution of ‘using a map’ while the third answer can be considered the contribution of ‘selecting the appropriate map’.

Before addressing these three elements, I briefly introduce the anatomy of vmPFC. In the context of thinking about vmPFC’s role in goal-directed behaviour, I make comparisons with various other regions in this chapter, especially hippocampus. I also end with some thoughts about the role of dACC in supporting disengagement and selection of goals, on longer time scales.

## 2.1 Anatomy of vmPFC

The vmPFC is not a region with clearly defined anatomical borders, but includes a set of regions in the lower orbital and medial prefrontal cortex, with connections to sensory and limbic areas (Mackey and Petrides 2014; Ongür et al. 2003; Ongür and Price 2000; Wise 2008; Carmichael and Price 1995a; Carmichael and Price 1995b). The broader region can include parts of Brodmann areas 14, 11m, 10m

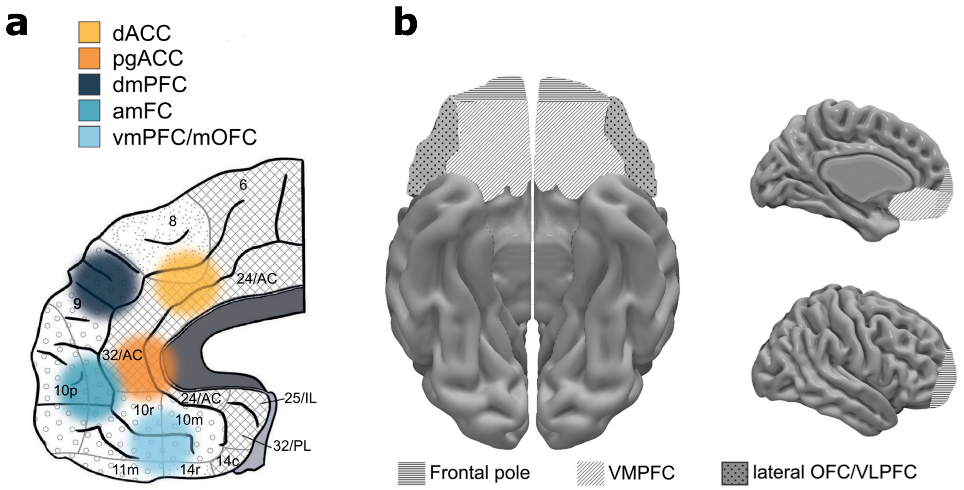


**Figure 2.1:** Anatomy of medial PFC. **(a)** Areas of medial prefrontal cortex in humans. Reprinted with permission from Klein-Flügge et al. (2022). The authors suggest anterior medial prefrontal cortex (amFC) shows greater activity during tasks investigating relational knowledge, while more ventral areas (vmPFC/mOFC) are implicated in value-guided choice. However the authors note that across many fMRI studies of decision making, activity is found in both locations or at the border (Bartra et al. 2013). **(b)** Lower PFC areas, reprinted with permission from Yu et al. (2020). Damage to frontal pole is often observed in patients with vmPFC damage. Note that based on connectivity, lateral areas belong to a different network (Carmichael and Price 1996), and have been linked to diverging functional impairments, for example the ability to attribute reward to its cause (credit assignment); see Noonan et al. (2010); Walton et al. (2011); Noonan et al. (2012); Noonan et al. (2017); Rudebeck et al. (2017).

(medial frontal pole), sometimes also including regions in 25, 32p (Fig.2.1). Spatial precision becomes a substantial challenge in studies of human lesion damage. Some studies refer to the total area including vmPFC and medial OFC as ‘ventromedial frontal lobes’ (VMF) (Yu et al. 2020; Camille et al. 2011)).

Within this large region, there are various proposals for functionally specialised subregions, some of which are discussed in this chapter (see also legend for Fig.2.1). However, given that our understanding of the anatomical divisions are still evolving (Glasser et al. 2016), and that lesion data has significant spatial limitations (Yu et al. 2020), I will avoid making claims resting on precise anatomical specificity within the ventral medial frontal lobes.

## 2.2 Value-guided decisions

What could be affecting goal-directed behaviour in patients with damage to vmPFC? One answer to this is that vmPFC patients find it generally difficult to make decisions guided by value. In the context of economic choice, the term value refers to the worth of options, as revealed by people's preferences (Jevons 1866; Samuelson 1938). It is well established that vmPFC lesion damage results in people showing less consistent preferences (Fellows and Farah 2007).

There is a long literature implicating the vmPFC in subjective value generally (Bartra et al. 2013; Klein-Flügge et al. 2019). Activity in vmPFC varies with value across different categories of reward (Bartra et al. 2013; Levy and Glimcher 2011; Chib et al. 2009; Lopez-Persem et al. 2020), and also incorporates the uncertainty of options in the way that economic expected utility signals should (De Martino et al. 2013). Even when items are presented on their own outside of choice scenarios, vmPFC activity varies with their value (Lebreton et al. 2009). Since these correlations are not limited to any single domain of reward (e.g. 'food' or 'aesthetic beauty') the theory was developed that vmPFC transforms value into a 'common currency' to guide a unitary behavioural output in the face of multifarious reward currencies (Levy and Glimcher 2012). Moreover, activity correlates specifically with subjective value – the preferences specific to the individual (Bartra et al. 2013; Lopez-Persem et al. 2020).

There has been great interest in understanding how neural activity in vmPFC could support the process of option selection. Comparison between binary options can be captured by models of mutual inhibition (Wang 2002; Hunt et al. 2021). Separate populations of neurons representing each option compete for choice. Recurrent excitation within populations and inhibition between populations produce dynamics which amplify the difference between the options (Strait et al. 2014). These dynamics can explain the trade-mark post-decision signal observed in fMRI studies of vmPFC, where activity often reflects the overall value difference between chosen and unchosen options (Boorman et al. 2013; Hunt et al. 2012; Hare et al. 2011).

However, if there is one conclusion to be drawn from previous decades of decision neuroscience, it is that no single brain area is uniquely responsible for all forms of decision (Rushworth et al. 2012; Hayden 2018; Kolling and O'Reilly 2018). Correlates of value have been identified in multiple other neural areas including dorsal anterior cingulate cortex (dACC), dorsolateral prefrontal cortex (dlPFC), and striatum, alongside similar dynamics of mutual inhibition (Rushworth et al. 2012; Hunt et al. 2015; Strait et al. 2015). These findings point towards the conclusion that choice emerges from the coordination of computations performed across multiple distributed areas (Rushworth et al. 2012; Cisek 2012; Hunt et al. 2015). So the question becomes, what does the vmPFC contribute to these processes, and when does it become necessary for effective choice?

Damage to vmPFC affects how people make decisions, but not always in predictable ways (Fellows 2018; Burgess et al. 2000; Pelletier and Fellows 2021; Yu et al. 2020; Tranel et al. 2007). Generally, choices become less consistent after vmPFC damage (Fellows and Farah 2007; Camille et al. 2011). One of the prevalent themes is that vmPFC damage affects decisions in more complex settings, for example when the demands of the task change, or where choices require integrating information across multiple attributes (Fellows 2018). For example, in various settings, vmPFC damage leads to differences in how attributes of items are selected or weighed up for choice compared to control groups (Vaidya et al. 2018; Pelletier and Fellows 2021; Bowren et al. 2018).

## 2.3 Cognitive maps

Could these difficulties shown by vmPFC patients in more complex settings reflect problems constructing value using a model of the environment? In the context of later discussions in this thesis, it is worth making a distinction at this point between using a model to simulate transitions in the environment, and selecting the appropriate model on the basis of context. To think about why these are different, consider the following story (Fig.2.2):

> Two neighbours are each hosting friends for dinner. They both head out to the supermarket to get ingredients only to realise the supermarket has closed for construction works. The bakery and the cheese shop will be open. However, neither individual has ever navigated to the bakery from the supermarket before.
>
> Person One simply cannot imagine a journey that would take them to the bakery from their current location, or the route from the bakery to the cheese shop. Their only option is to follow their usual route home, sad and frustrated.
>
> Person Two has a different problem. Unlike Person One, they are perfectly capable of determining how to get to the bakery. They know if they turn left and go up the road, theyll be able to find the bakery. From there, they could turn right and go to the cheese shop, but they also know that if they turned left instead they would arrive at the cinema. Person Two comes out of Oppenheimer three hours later. They loved it.

Both of these people have a problem using cognitive maps to achieve the goal– but the problem seems different in the two cases. Person One has a problem with inference, in the sense that they lack knowledge of the transition structures between states, or the ability to simulate these transitions (Fig.2.2b). Person Two understands the transition structure, but they have a problem selecting the relevant information in the context of their goal (Fig.2.2c). They include information about a completely irrelevant state – going to the cinema – when their goal is to buy groceries.

Note that both individuals will make similar mistakes in some settings. Specifically, they might both be likely to choose options which have been highly rewarded in the past but are not relevant for the current goal. However, crucially this would be for different reasons. For Person One, it is due to a reliance on habitual (model-free) action. Person Two, in contrast, has a map of the world, but they fail to constrain the map to the relevant information for achieving the goal. What deficits, if either, do patients with vmPFC show?

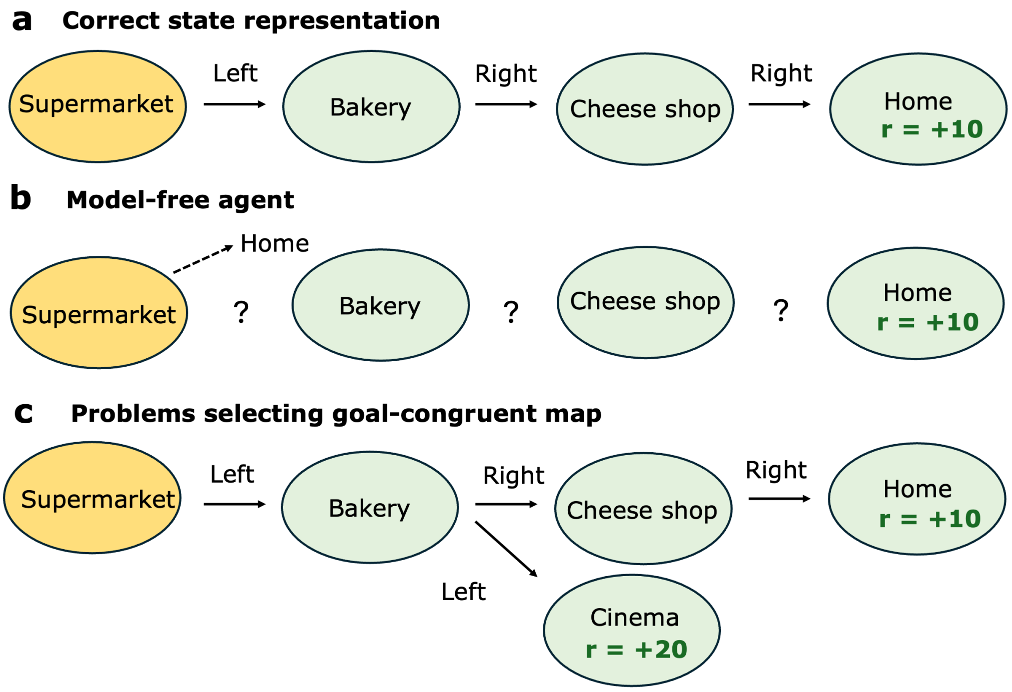


**Figure 2.2:** Depiction of three maps representing the task of buying bread and cheese **(a)** The correct map for achieving this goal. **(b)** A deficient map which does not include relational information. **(c)** A deficient map which fails to constrain options to goal-relevant information. A different perspective on this is that the agent has an incorrect state representation (where the notion of state is taken from reinforcement learning).

### 2.3.1 Using a map

There is some evidence implicating vmPFC damage in both types of deficits. Let's start with the first – planning using a model of the environment. This would be critical for goal-directed behaviour as described at the beginning of **chapter 1**. In particular, being able to use a cognitive map to guide choices allows agents to plan towards goals rather than rely on reinforced action. A classic measure of knowledge about transition structures has been inference – responses which reflect structural knowledge about the world, where the conclusion could not have been drawn from direct experience. This is just as true for inference in non-spatial domains as it is in spatial domains (Behrens et al. 2018).

Some studies have suggested lesions to vmPFC in humans impair the ability to make inferences (Koscik and Tranel 2012; Spalding et al. 2018; Wing et al. 2021). For example, in one task participants learned about multiple pairs of ordered stimuli, represented by abstract images (e.g. A>B and B>C). At the end, the participants were tested on their memory for learned pairs, as well as their ability to perform transitive inference across pairs (e.g. to infer that A>C in the examples above, despite never being told this directly). VmPFC lesion patients showed no difference in memory, but were impaired at making transitive inferences compared to control lesion patients and healthy controls (Koscik and Tranel 2012). This suggests that the patients have difficulty forming a map of the relationship between items.

OFC areas more broadly in rodents seem necessary for decisions based on inferred rather than experienced value (Jones et al. 2012). For example, in one study rodents learned to associate two pairs of cues – A with B, and C with D. Following this, one of the secondary items (B) was paired with a reward. After lesioning OFC with muscimol, rats were presented with each cue. Control animals reponded appetitively both to the experienced cue (B) and inferred cue (A), which both signal reward. OFC-lesioned animals continued to respond to the directly rewarded item (B), but not to the item predicting reward through inference (A).

One general question here concerns how these findings about inference processes relate to anatomical divisions between lateral and medial OFC regions in primates. Lateral prefrontal cortex has been associated with model-based planning in various tasks (Gläscher et al. 2010; Tanji et al. 2007; Fermin et al. 2016; Bartolo and Averbeck 2020; Smittenaar et al. 2013). One study found state prediction errors – i.e. the surprise involved in making an unexpected transition between states – in lateral PFC, suggesting it could be important for tracking one's current position within a map of states (Gläscher et al. 2010). Supporting this idea, disruption to human dlPFC using theta burst transcranial magnetic stimulation reduced model-based planning (Smittenaar et al. 2013). Given that medial and lateral PFC regions belong to different anatomical networks (Carmichael and Price 1996), they are likely to make different contributions to inference and model-based planning. Notably, lesions to lateral OFC in both macaques and humans damage the ability to attribute reward to its cause ('credit assignment'), while lesions to medial vmPFC disrupt value maximisation (Noonan et al. 2010; Walton et al. 2011; Noonan et al. 2017; Rudebeck et al. 2017).

Another consideration is how vmPFC may be interacting with other structures important for inference. One of the earliest proposals for neural ‘cognitive maps’ was the hippocampus (O’Keefe and Nadel 1978; Redish 1999). A variety of cell types–most famously place and grid cells–support spatial mapping in the hippocampus and entorhinal cortex (O’Keefe and Nadel 1978; Hafting et al. 2005; Ekstrom et al. 2003). Hippocampal neural ensembles encode ‘sweeps’ of potential paths in freely moving rats at choice points in a T-shaped maze (Johnson and Redish 2007). These same architectures may support simulation using conceptual (non-spatial) forms of knowledge (Addis and Schacter 2012; Eichenbaum and Cohen 2014). Recently, there has been evidence that damage to hippocampus impairs model based inference in both humans (Vikbladh et al. 2019) and rodents too (Miller et al. 2017).

Many tasks where the outcome of an inference is decoded in vmPFC activity also find hippocampal activity at choice (Barron et al. 2013; Barron et al. 2020; Park et al. 2021). Greater functional coupling between hippocampus and vmPFC predicts better inference based on recalled relationships (Zeithamova et al. 2012). One possibility given these findings is that vmPFC is involved in recruiting computations for the current task, while the inference itself is computed in other areas such as hippocampus.

To summarize, it is possible vmPFC damage impairs goal-directed behaviour partly because it impairs the ability to make model-based inferences. However, the distinct contribution of vmPFC to this process is unclear. In **chapter 6** of this thesis, I investigate whether vmPFC damage affects model-based planning using two previously established tasks from the planning literature (Two-Step and Four-in-a-row). In general, I do not find evidence supporting the disturbance of model-based planning generally, but I do find alterations in behaviour related to selecting relevant information. It is possible that vmPFC impairments in model-based inference concern difficulties selecting and coordinating computations in other regions. With this in mind, in the next section I review the idea that vmPFC is important for selecting the appropriate map.

### 2.3.2 Selecting a map

Whether or not vmPFC is necessary for making all model-based inferences, another idea is that vmPFC is particularly important for selecting which map to use. In other words, these regions might be critical for representing appropriate dimensions for guiding behaviour, while ignoring task-irrelevant dimensions. This brings a different perspective to some of the earlier ideas concerning subjective value. In this context, vmPFC may be critical when decisions require selecting and representing the appropriate pieces of information in context-dependent situations.

It is well established that value correlates in vmPFC shift dramatically depending on context. An example of this is the effect of homeostatic needs on neuronal firing patterns. There is a decrease in vmPFC neurons selective to water when monkeys are no longer thirsty (Bouret and Richmond 2010). In the domain of mood, it has been shown that positive mood manipulations enhance the effect of subsequent rewards on vmPFC activity in humans (Young and Nusslock 2016). In both these cases, changes in underlying behavioural state alter the reward response in vmPFC. However these observations alone could be explained by changes in the hedonic reward properties of water (less pleasurable when no longer thirsty), or of reward (less pleasurable when in low mood).

Stronger evidence for the flexibility of these value signals come from findings demonstrating vmPFC activity dynamically adapts to abstract task rules or behavioural goals (Rudorf and Hare 2014; Grueschow et al. 2015; Frömer et al. 2019; Castegnetti et al. 2021; Trudel et al. 2021). In one recent study, Trudel and colleagues found that uncertainty had opposite effects on vmPFC activity depending on whether the individual was in an explorative state near the beginning of a trial (positive effect of uncertainty), or an exploitative state later on (negative effect of uncertainty). The authors argue this can be explained by an adaptation of vmPFC signal to the individual's behavioural goals – namely whether their goal is to 'explore' for new information, or 'exploit' their existing knowledge (Trudel et al. 2021). During early learning about potential predictors, uncertainty indicates opportunities for new information, leading to positive vmPFC signals. Once valuable predictors have been identified, uncertainty is avoided since it decreases the expected value of the outcome, and therefore is negatively signalled in vmPFC.

This finding chimes with various other fMRI studies showing vmPFC signals do not reflect a context-neutral evaluation, but flexibly change depending on how appropriate the choices are for the current task. Castegnetti and colleagues designed a study where objects were chosen on the basis of their usefulness for achieving different goals (Castegnetti et al. 2021). A metal chair will be more useful when the goal is to anchor a boat, but a wooden chair will be more useful if the goal is to light a fire. The authors found vmPFC encoded the degree of goal-dependent usefulness, rather than a simple preference between objects based on prior subjective valuation. Another striking demonstration of this is the observation that vmPFC activity reverses for the very same options when the instruction is to select the "worst" rather than "best" item, showing clear goal-congruent rather than hedonic representations (Frömer et al. 2019). Taken together, these studies suggest vmPFC signals reflect the value of options contextualised by an individual's current task.

## 2.4 Context-dependent choice

If it is true that vmPFC plays a role in selecting context-dependent dimensions for choice, at what point does this become important? Or to put the question differently, when might we expect damage to vmPFC to be particularly impeding? There are three kinds of scenarios that I discuss below. First, when context-dependent choice cannot be known through observable information, vmPFC might be particularly important (Wilson et al. 2014; Schuck et al. 2016). Second, it may be critical in cases when decisions require complex integration of different dimensions (Pelletier and Fellows 2021). Finally, it may be important for guiding decisions in the face of crowded sources of value, competing for attention. In the following section, I review these three ideas.

### 2.4.1 Non-observable state information

Deciding whether to wear a new pair of shoes will depend on whether you are going on a hike (consider: are they waterproof?) or going to a fancy dinner (consider: are they formal?). In cases like this, the current context is not visually observable, but decisions must reflect knowledge of hidden state variables. Some authors have argued the ability to represent these hidden variables in vmPFC will become critical when decisions depend on information which is not observable (Wilson et al. 2014; Schuck et al. 2016).

An helpful way of viewing this is informed by the computational concept of a 'state' in reinforcement learning (Wilson et al. 2014). As briefly described in **chapter 1**, the notion of a state describes the collection of information which is relevant for a decision (Sutton and Barto 1998). The critical idea is that states contain all the information which determine (a) the probability of immediate reward, and (b) the probabilities of transitioning to other states in the task, and their values. This is what gives RL problems what is known as the 'Markov Property' – the feature that the future is only dependent on the current state, independent from the history of past choices and states. In order to have such a property, state information might include observable variables (such as the stimulus on the screen), but also unobservable variables (such as the current task rule).

According to this theory, activity in vmPFC more generally encodes variables defining the task-specific state, i.e. the set of relevant information to make value-guided decisions in a particular task (Wilson et al. 2014; Schuck et al. 2016). In a canonical study establishing this idea (Fig.2.3), participants were presented with superimposed pictures of faces and houses and asked to report the age (young/old) of one of the categories (Schuck et al. 2016). Critically, whenever two sequential images were of different ages, the participant had to switch to reporting the age associated with the other category. To make the correct response, the relevant

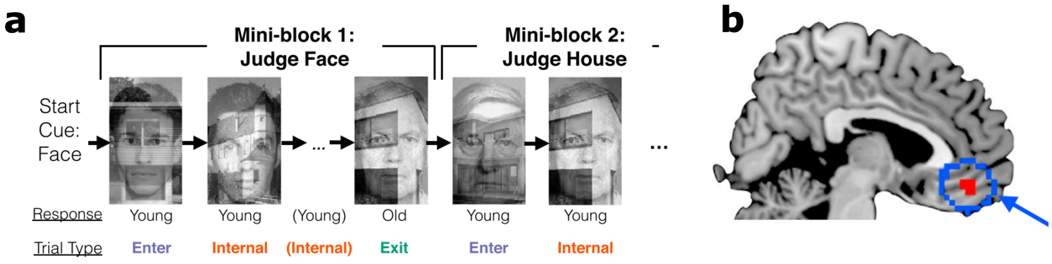


**Figure 2.3:** VmPFC encodes unobservable state variables relevant for making decisions. Reprinted with permission from Schuck et al. (2016). **a** Correct responses in this task depended on three hidden variables: the previous relevant category (house/face), the previous age, and the current relevant category. **b** The conjunction of these three hidden variables could be decoded in vmPFC.

information included the age and category of the previous image, as well as the age and category of the current image. All relevant but non-observable variables could be decoded from vmPFC, while no irrelevant variables could be reliably decoded. The authors argued this suggested vmPFC held the relevant 'state representation' – all the (unobservable) information needed to make the right response.

In support of this theory, Bradfield and colleagues compared versions of a devaluation study with and without observable outomces. Rats with lesions to mOFC continued to perseverate with performing the devalued action in the unobservable condition, but not in the observable condition (Bradfield et al. 2015).

### 2.4.2 Integrating dimensions

VmPFC is also critical in cases where choices depend on a complex integration of value dimensions (Pelletier et al. 2021). This can be seen as an example where being able to select the correct map will be important because the weighting of different dimensions is context-dependent. In one task, Pelletier and colleagues (2021) presented subjects with objects whose value was determined either by summing the individual dimensions, or through a configuration of dimensions. In the 'configuration' condition, an object's value was determined by the unique pairing of elements. Individuals with vmPFC lesions were impaired at making configural evaluations but not when evaluating the sum of the parts.

How could the vmPFC integrate dimensions to guide choice? Some have argued the value comparison process emerges as the product of a cognitive map – a space which allows dimensions to be integrated or compared (Behrens et al. 2018). There is currently much interest in the idea that the vmPFC – similarly to the hippocampus – contains 'grid-like' activity which is modulated in a hexagonal manner (Constantinescu et al. 2016; Behrens et al. 2018; Park et al. 2021; Veselic et al. 2023). This has been proposed as a potential mechanism for structuring the relational knowledge in order to form comparisons, in conceptual or value space (Behrens et al. 2018). For example, in one study (Park et al. 2021), participants learned about pairs of individuals in a 2D social space (individuals ranked by 'popularity' and 'competence'). Inferences about the value of individuals across previously unexperienced pairs were supported by grid-like activity in both entorhinal cortex and vmPFC. Recently it has been argued that value-based inference choices in rodents are also made in this way: grid-like codes in vmPFC flexibly re-map to reflect the inference in value space akin to the way that hippocampal grid cells remap to new locations in spatial maps (Veselic et al. 2023). However, the particular claim that vmPFC activity encodes all value using grid-like codes is contested (Lee et al. 2021; Vaidya and Badre 2022). One large study explicitly compared the subjective value hypothesis to a grid code hypothesis in vmPFC activity during an intertemporal choice task, and found no evidence for grid-like coding frameworks (Lee et al. 2021).

### 2.4.3 Ignoring irrelevant information

Another situation in which the ability to select relevant information will become critical is when irrelevant sources of value compete for attention. Distraction by irrelevant value has a larger effect on vmPFC patients. In one task people learned to associate three items with changing levels of reward. In both monkeys and humans, vmPFC damage caused interference from the value of the lowest (third) option, which should be irrelevant for the choice (Noonan et al. 2010; Noonan et al. 2017). A possible explanation is that damage impairs the capacity to orient

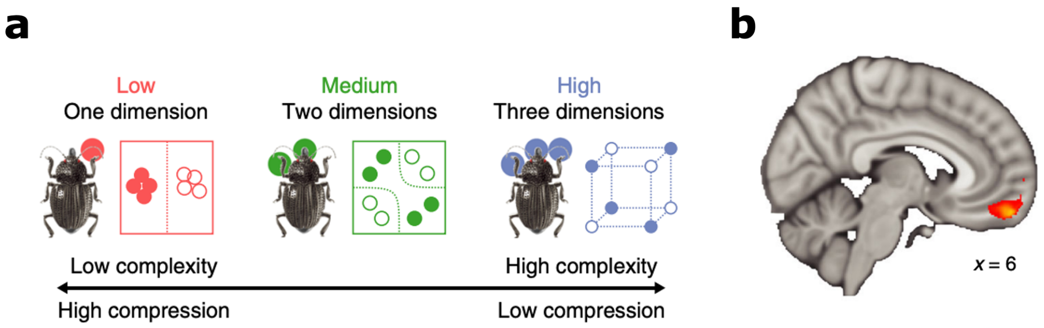


**Figure 2.4:** VmPFC and compression of irrelevant information. Reprinted with permission from Mack et al. (2020). **a** The task required categorising beetle stimuli on the basis of either one dimension (left), two dimensions (middle), or three dimensions (right), where dimensions corresponded to visual features such as antenna length. **b** VmPFC activity correlated with the degree of task compression, where neural compression was measured as the complexity of neural representation using principal component analysis.

attention to the relevant items for consideration, rather than be distracted by competing but irrelevant value (Gluth et al. 2018; Gluth et al. 2020b).

There is also evidence in healthy participants that vmPFC is involved in compressing across irrelevant information (Mack et al. 2020; Leong et al. 2017). In one study (Mack et al. 2020), participants learned to categorise beetles on the basis of visual features, such as the size of their antennae (Fig.2.4). Participants learned different versions of the task, where the beetle's category could depend on one feature, the conjunction of two features, or the conjunction of three features. Learning the task therefore required selecting the relevant feature(s), and compressing the irrelevant features, where the degree of compression scales inversely with the number of features required to make the categorisation. VmPFC was the one area where neural compression (the complexity of neural representation measured using principal component analysis) correlated with the inherent compression in the task (Fig.2.4). In addition, the extent to which participants selectively attended to the correct information (predicted by a computational model) correlated with subject-specific vmPFC compression. This suggests vmPFC may be involved in reducing the dimensionality of input to select only what is relevant for the current context.

To what extent do these areas actively suppress signals of irrelevant value? Mahmoodi and colleagues used a social decision task which required participants to judge pairs of faces on the basis of three different social dimensions (Mahmoodi et al. 2023). Continuous theta burst transcranial magnetic stimulation (cTBS) was used to causally disrupt activity in anterior mPFC and anterior insula (AI). AmPFC was found to be important for mediating interference from irrelevant dimensions, while AI enhanced the relevant dimension. On incongruent trials (where the ranking of the individuals being compared differed across the dimensions), disruption to amPFC increased interference from the irrelevant social information. Crucially, this was not due to a noisier comparison, but specifically an increase in the impact of irrelevant value on these harder incongruent trials. In comparison, disruption to AI also impaired performance, but this time through a decreased impact of the relevant dimension value difference, and did not affect use of irrelevant information. It is worth noting the region found here is considerably more dorsal than the vmPFC areas discussed above, but provides a lens for considering how mPFC generally might contribute to selecting relevant information.

These observations that vmPFC could help to guide behaviour in the face of competing 'maps' is consistent with findings from a different corner of the literature. Another perspective on vmPFC damage has focused on observations of 'confabulation' (Moscovitch 1989). In these cases, patients report and continue to believe wildly falsifiable claims even when the relevant information for discounting them is available. This has led to the development of theories that vmPFC is necessary for the use of 'schemas' or contextual frameworks to guide memory recall (Ghosh et al. 2014; Spalding et al. 2015; Farovik et al. 2015; Warren et al. 2014). The concept of a schema refers to structures of memories which link experiences in similar contexts (Bartlett 1932). For example, swimsuit and ice-cream belong to the schema of 'beach', while hammers do not. Various studies have suggested that vmPFC damage impairs use of the appropriate schema (Ghosh et al. 2014; Spalding et al. 2015). For example, in one study vmPFC lesion patients and controls were asked to report whether words were related to a schema ("a visit to the doctor"). Later, they performed the same task with a new schema ("going to bed"). Patients with vmPFC damage failed to reject lures belonging to the previous schema, suggesting a problem with retrieving information for the right context (Ghosh et al. 2014). Another way of interpreting this is that vmPFC lesion patients were unable to suppress information related to dimensions of the task which were now irrelevant to the current goal.

### 2.4.4 Summary

In the last section, I reviewed the idea that vmPFC plays a role in selecting the correct map to guide choices in the context of the current task or goal. I discussed three decision scenarios where this capacity may be particularly important. First, when the goal or other relevant information is not visually observable. Second, when choices depend on integrating value dimensions according to complex criteria. Third, when goals must be pursued in the face of irrelevant forms of information, competing for attention. This last case resembles the story earlier, where Person Two has the goal of buying groceries but ends up in the cinema because irrelevant information creeps into their state representation. This idea will be particularly important for understanding the results of **chapter 5** in this thesis – where vmPFC damage causes patients to abandon their goals more. In the final section of this chapter, I turn to a different question. If vmPFC is critical for selecting the right map to pursue a goal, how do we evaluate the goals themselves?

## 2.5 Escaping goals and dACC

If vmPFC selects goal-relevant material, do we escape bad goals? A body of literature has identified dACC (see Fig.2.1a for anatomy) with an activity profile which could support longer-term decisions to re-evaluate behavioural strategies (Heilbronner and Hayden 2016; Holroyd and Verguts 2021; Kolling and O'Reilly 2018). For example, dACC activity predicts decisions to abandon the current patch in foraging paradigms (Fig.2.5). In these 'patch-leaving' tasks, animals are

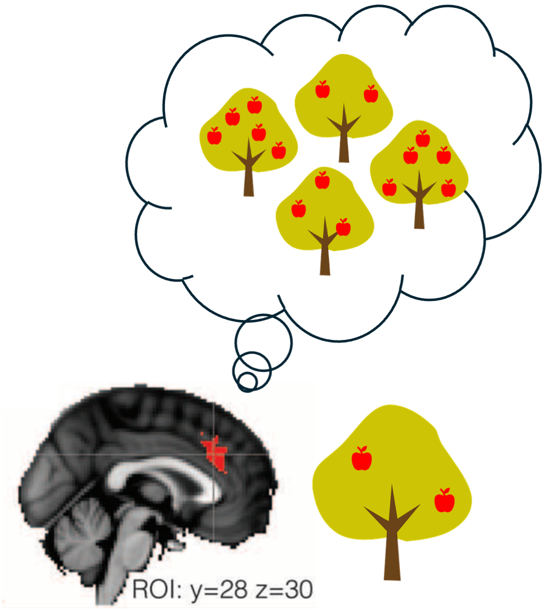


**Figure 2.5:** Evaluating abandonment in dACC. Brain image reprinted with permission from Kolling et al. (2012). In foraging paradigms, dACC correlates with the long-term value of leaving the current patch and pursuing alternatives, and is more active when participants choose to abandon.

presented with rewards which slowly deplete over time. At any point the animal can choose to stay and continue accepting the current reward, or leave, in which case they must wait for a long delay before the reward is reset to its original value. Over the course of the reward depleting, the firing rate of dACC neurons gradually increases until a threshold is reached and the animal chooses to leave (Hayden et al. 2011). This could be considered a case of dACC activity tracking the longer term value of the goal itself. In scenarios where animals are presented with choices to accept or reject prey, dACC signals also correlate with the value of rejected options, as if estimating the value of alternative paths of action (Blanchard and Hayden 2014). Finally, the dACC has been shown to track long-run trends in reward which could help predict when a goal is no longer worth pursuing (Kolling et al. 2016; Wittmann et al. 2016).

One recent study found separate neural circuits within rodent dACC supporting two kinds of foraging decisions: the decision to switch to a new strategy (goal disengagement), as well as the decision to commit to a new alternative (goal selection) (Tervo et al. 2021). These observations are consistent with the idea that the primary role of ACC is in goal selection, including disengagement if the goal goes awry (Holroyd and Verguts 2021).

The dACC is also active when people decide to switch between exploiting a current option and exploring the environment (Trudel et al. 2021). In fact, when people switch out of an exploitative state towards exploration, dACC activity predicts subsequent changes in task representation in the vmPFC (Muller et al. 2019), supporting the theory that it could be setting goals subsequently supported by vmPFC activity. Taken together, this set of studies suggests a potential role for dACC in the hierarchy, which involves the evaluation of goals themselves.

## 2.6 Summary

In this chapter, I reviewed the ways in which vmPFC and dACC could support pursuit of goals. In particular, vmPFC may play a critical role in selecting relevant information in the context of current goals. On the other hand, dACC could support longer-term decisions to select and re-evaluate goals.

In the following chapters, I investigate the contribution of both of these areas to goal pursuit. This involves an fMRI study in **chapter 4**, and studying the impact of vmPFC damage in lesion patients in **chapters 5-6**. The results build on the ideas of this chapter, suggesting vmPFC is necessary for constraining decisions to the context of current goals. In **chapters 3-5**, this is investigated using a new task where individuals must flexibly switch between goals. VmPFC shows sustained activity predicting commitment and attention to current goals, while dACC is implicated in decisions to abandon. When vmPFC is damaged, patients are more engaged by alternative goals, suggesting that in healthy people, vmPFC is playing a role in constraining the decision space to the currently pursued goal. This capacity will be critical in complex environments, where goals must be pursued in the face of vast quantities of irrelevant information. In **chapter 6**, I find that the deficits caused by vmPFC damage in planning are linked to the tendency to overlook relevant valuable features in complex settings. Bringing these chapters together, I will argue this area supports goal pursuit by selecting and integrating the relevant information to guide behaviour in light of the current goal.

# 3

# Balancing commitment and abandonment during goal pursuit

*This chapter sets up the behavioural paradigm and framework which is later used in the fMRI study (**chapter 4**) and lesion patient study (**chapter 5**). When striking the balance between commitment to a goal and flexibility in the face of better options, people often demonstrate strong goal perseveration. It is unclear how this goal commitment bias develops during pursuit of a goal, or how it responds to different forms of pressure. To study this, we develop an incremental goal pursuit task involving sequential decisions between persisting with a goal versus abandoning progress for better alternative goals. The task design allows us to disentangle the effects of different forms of pressure on peoples' tendency to persist, such as the rate of progress, or the impact from alternative options at different points in pursuit. By interleaving an additional spatial working memory task between decisions, we identify how attentional biases develop over the course of goal pursuit in a setting outside the decision context. We find that individuals with stronger goal commitment show higher goal-directed attention in the interleaved task. Increasing goal-directed attention also affects abandonment decisions: while pursuing a goal, people lose their sensitivity to valuable alternative goals, while remaining more sensitive to changes in the current goal, shown through both choices* *and reaction times. Rather than representing commitment biases as a (perhaps irrational) factor in the decision process itself, we argue it is better understood in terms of a more pervasive attentional effect: Mechanisms of selective attention prioritise processing of the current goal over alternative goals, resulting in reduced sensitivity to attractive alternatives.*

## 3.1 Introduction

In natural environments, many goals – whether it be pursuing prey, cooking dinner, or writing a thesis – are only obtained after persevering through a substantial period of unrewarded time and effort. In all these cases, optimal behaviour requires balancing commitment to the current goal against flexibility to abandon if the goal is no longer worth pursuing relative to alternatives. Psychiatry and neuroscience have tended to focus on failures of commitment during extended behaviours (Heron et al. 2019; Dalley and Robbins 2017; Kouneiher et al. 2009). However, behavioural economics provides us with ample examples of people showing too much commitment to a goal, particularly after investing time or money (Arkes and Blumer 1985; Mcafee et al. 2010; Ronayne et al. 2021). These 'sunk-cost' biases are not unique to humans, but have been found in rodents too (Sweis et al. 2018).

Why might animals show biases towards over-persisting with a goal? When behaviour is structured by sequential goals, constant re-evaluation can be both expensive and distracting. In consequence, it has been proposed that distinct phases of 'deliberation' (evaluation of available options) and 'implementation' (committing cognitive resources to achieving the chosen goal) might be present in both humans and non-human animals (Heckhausen and Gollwitzer 1987; OReilly 2020; Ludwig et al. 2020; Li et al. 2019; Molinaro and Collins 2023; Sweis et al. 2018). However, a picture involving entirely discrete decision phases fails to explain how animals remain flexible to goal abandonment when the situation requires it. As discussed in **chapter 1**, a plausible mechanism would allow for the agent both to preferentially allocate processing resources to goal completion, while retaining the necessary flexibility.

A candidate mechanism for such flexible focus on a goal is selective attention, specifically towards information about the chosen goal. Attentional selection need not be all-or-nothing, but can vary in strength as the need to exclude distractors varies (Lavie 2005), thus allowing for flexibility. In ecological scenarios, we are faced with different reasons for abandoning a goal: progress might be too gradual or might reverse; alternatively other options might become significantly more attractive. These different forms of pressure give rise to different emotional responses: frustration (with the current goal) in the former cases (OReilly 2020), and temptation (by alternative goals) in the latter. If selective attention to the chosen goal increases over the course of goal pursuit, this leads to a rather specific prediction about the interaction of 'temptation' and 'frustration' with increasing proximity to the goal: namely, sensitivity to the value of alternative goals ('temptation') should decrease more than sensitivity to the value of the chosen goal ('frustration').

The aim of this chapter was to investigate how commitment biases emerge over the course of goal pursuit. Specifically, we tested our hypothesis that attention and decision-making show these markers of increasing attentional orientation towards the current goal. To probe this, we developed a novel sequential choice task, where participants advanced incrementally towards completing a chosen goal in the face of alternative goal offers. We orthogonally varied the value of the current goal and the value of alternative goals at the decision, as well as continuously measured goal-oriented attention outside the decision period.

Our task design allowed us to disentangle the effects of different forms of pressure on peoples' tendency to persist, such as the rate of progress towards the current goal, or the impact from alternative options at different points in pursuit. Subjects were rewarded each time they amassed a target quantity of goods, which required collecting that good across several trials. Crucially, if subjects switched to a different type of good they sacrificed all the goods accumulated thus far, abandoning their existing goal. Since the aim was to maximise reward across the study, goal abandonment was often preferable if progress with the chosen good was unsatisfactory or an alternative good became more abundant. In addition, we looked at how goal pursuit affected people's performance in an intermediate spatial attention task performed at regular intervals (before each decision), which was irrelevant for goal progress.

We found that participants showed a universal 'goal commitment' bias towards persisting with their current goal, even in circumstances when they would greatly benefit from abandoning it. We were able to measure several markers of selective attention to the current goal. First, as predicted by the attentional account, decision-making and reaction times reflected goal-oriented attention: as participants approached goal completion, their responses remained relatively more sensitive to the value of the current goal, than to the value of alternatives. Second, using a separate spatial working-memory task, we found that even outside the decision period, stimuli related to the current goal were increasingly prioritized in attention. Finally, individual differences in commitment bias were predicted by goal-oriented attention measured outside the decision period.

## 3.2 Methods

### 3.2.1 Participants

Prior to collecting the cohort of participants for our MRI study, we piloted the task extensively on Prolific.co. Here we present the behavioural results from the MRI cohort, to lay the groundwork for upcoming chapters. A total of thirty-one participants (19 female; mean age 25 years, normal or corrected-to-normal vision) were recruited via email circulation on Oxford University mailing lists and social media. One participant was excluded from the recruited sample because they opted out of the study before the MRI scan, leaving a total of 30 participants whose data is analysed in this thesis. Ethical approval for the MRI study was obtained by the Oxford Central University Research Ethics Committee (Ref: R72921/RE001).

All participants gave written informed consent before the experiment. Participants were paid £15/hour plus a performance-dependent bonus between £8-12.

### 3.2.2 Experimental paradigm

#### Primary decision task

We developed a "fishing net" task where participants were incentivised to fill as many nets with seafood as possible over the course of the study Fig.3.1. Participants accumulated seafood "goods" over several trials, and gained a single point when a net was full. On each trial, participants chose between offers for three types of goods (octopus, crab, or fish), where the quantity available for each good was shown by a green bar. Once selected, the exact offered quantity would be added to the net. Importantly, only a single type of good could be collected in the net at once. This meant that if participants chose a different type of good to the type currently in their net, they would forfeit all their previously accumulated goods ('abandonment choice'). Alternatively, participants could choose to continue with the current goal by selecting the same type of good already in the net ('persistence choice'; Fig.3.1a for example). The horizontal order of the three creatures on the screen was randomised on every trial to avoid confounding persistence with motor perseverance.

At the start of each block, participants were shown the size of the net to be filled as an empty grey bar at the bottom of the screen. Blocks ended when a net was complete, and a point was won. Participants were therefore incentivised to fill as many nets as possible across the study, limited only by the number of choice trials in the study. The number of trials remaining in which the participants could continue to fill nets was shown in the top right corner of the screen throughout the study (Fig.3.2). Above this was shown the number of points earned (nets completed thus far), where each completed net was converted to a 25p bonus payment at the end of the study.

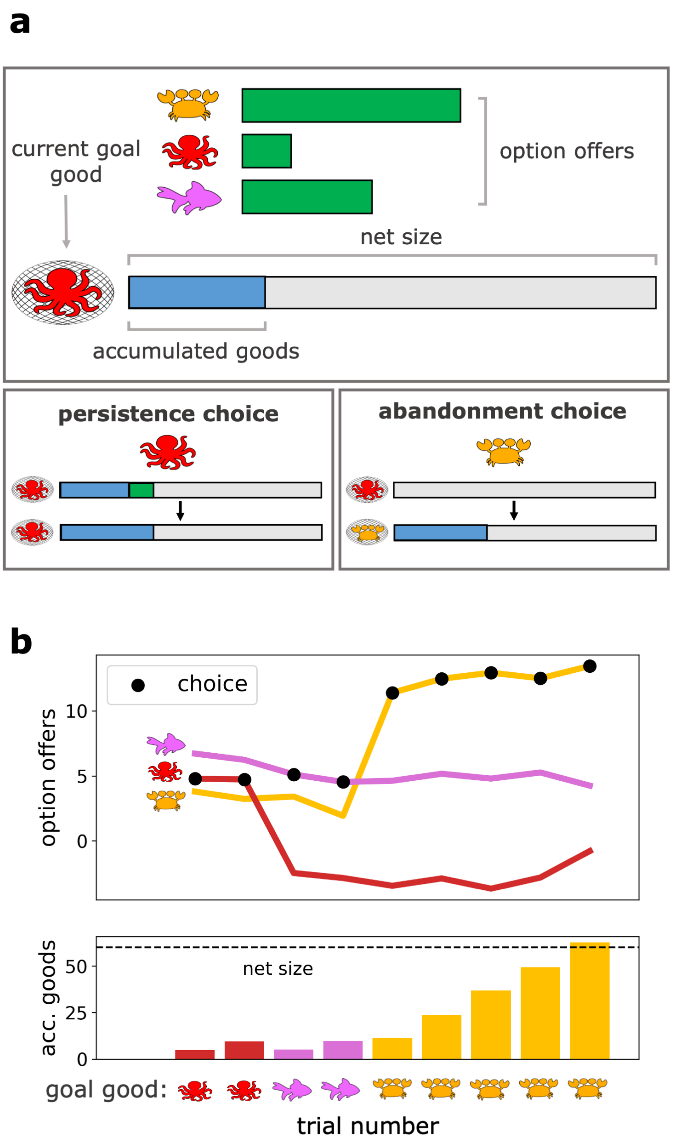


**Figure 3.1:** Decision task. **(a)** Participants performed a fishing net task which involved incrementally filling nets with seafood goods. Top panel: The current net was shown at the bottom of the screen (grey bar), and participants incrementally filled it with seafood across several trials. The net icon indicated the current type of good in the net (for example, octopus), while the blue bar indicated the aggregated quantity. On each trial, green bars indicated the current available quantities for each type of good (octopus, crab, or fish) which participants could add to their net. Critically, since the net could only contain one type of good, switching goods meant forfeiting the aggregated quantity. Bottom left: If participants continued with the same good, the offered quantity was added to the aggregated quantity. Bottom right: If participants chose a different good, the aggregated quantity was emptied before the new goods were added. Participants received a single reward when a net was completed, and the net size and option offers were re-set. **(b)** An example block where a participant switches goals twice. Top panel: Coloured lines depict the offers associated with each type of good across a block. Black dots depict the chosen good on each trial. During a block, the offers associated with each good varied gradually across trials with independent random walks, but could also jump to extreme high or low values (from where the random walk would continue). Bottom panel: Bars depict the goods accumulating in the net until completion.

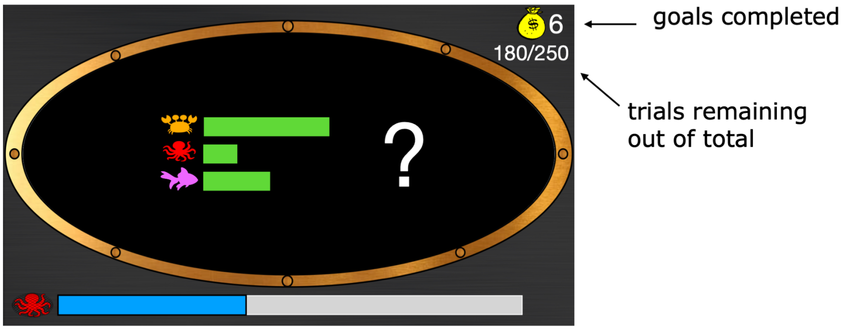


**Figure 3.2:** Choice visuals. At the time of choice, participants saw the three offers on the screen, the current contents and size of the fishing net. They were also shown the number of trials remaining, and the total points (corresponding to nets completed) in the top right. Participants were incentivised to use each choice as efficiently as possible to maximise points won within the limited number of trials.

**Offer trajectories** While the quantities offered for each type of good usually drifted gradually from trial-to-trial (random gaussian walk with low variance), sometimes the quantity would drastically change for a given a good (10% chance of a large shift up or down in quantity, independent for each type of good; see Fig.3.1b for example offer trajectories across a block). If the quantity associated with the current goal good collapsed or if an alternative good became much more bountiful, participants often benefited from abandoning their progress and switching to an alternative good. Participants were explicitly instructed that generally the quantities of creatures change slowly but sometimes a particular population of sea creatures would suddenly enter or leave the current fishing patch.

Specifically, for each block (commencing with a new net), the size of the net and the option offers differed. The net sizes were drawn from a uniform distribution (min=12, max=72). The initial values for the three options were drawn independently from a normal distribution at the start of each block ($\mu$=6, $\sigma^2$=1). From trial to trial, the offers for each option changed according to independent gaussian random walks ($\sigma^2$=0.8). In addition, on each trial there was an independent probability of any option changing more drastically in its associated offer ($p$=0.1 jump up, $p$=0.1 jump down), corresponding to an option becoming significantly more 'bountiful' or 'scarce' for fishing opportunities. The jump function consisted of drawing a random value between 3 to 9 points higher or lower than the option's starting offer, which corresponded to the new offer for that item. After a jump, the subsequent offers for that option would continue to change according to a random gaussian walk from the new starting location.

Offers were mostly positive (indicated by green bars), but could occasionally become negative (indicated by a red bar). If a negative offer was selected, the quantity of the bar would be subtracted from the net. Once a net was empty, nothing more could be lost so choosing a negative offer would lead to no change.

**Schedule variants** To minimise schedule-specific artefacts, we generated 5 different schedules which each consisted of 45 blocks of 100 trials. A block ended when the net was filled. In order to select pairs of net sizes and option offers for which completing the net was non-trivial yet feasible, we chose combinations where goals were completed in more than 3 trials and less than 15 trials when choice behaviour was simulated using the tree-search model. Participants on average viewed only 7 trials per block before completing the net. For each participant, separate schedules were randomly selected for the within-scanner and post-scanner sessions. In the lesion patient study (**chapter 5**), the same schedule was used across all individuals (including age-matched controls) due to the limited sample size for lesion patients. Each session ended after a pre-determined number of trials (300 in the fMRI session, 100 in the post-scan session, and 250 for all participants in the patient study), so no participant was able to complete all 45 blocks of a schedule within the available experimental trials.

### Attention task

In addition to the main decision task, participants performed an interleaved spatial attention task before every trial, providing a separate measure of attentional capture by the current goal (Fig.3.3). Participants viewed the stimuli associated with the three goods flash on the screen for 500ms in random locations. Subsequently, each creature icon appeared sequentially in the top right corner of the screen, prompting participants to respond by using their mouse to click at the location where they remembered it appearing. The three stimuli were probed in a random order. While the attention task involved the same stimuli as the decision task, participants were explicitly told that memory performance would not impact subsequent offers in the decision task.

### Experimental procedure

The data presented come from two sessions: one session inside the scanner (scanner session) and one session outside the scanner (post-scan session). It was not possible to interleave the attention task during the scanner session due to the practical difficulties of eliciting spatial responses with a button box. In the scanner session, the decision task was performed alone, while in the post-scan session, the attention and decision task were performed sequentially on every trial. The training, scan and post-scan task were all carried out in a single session lasting 2.5-3 hours total. Before the scan, participants were trained on the task for approximately twenty minutes. Participants practiced on three full example blocks (on average approx. 25 trials, dependent on performance) with the interleaved spatial attention task included, and one additional example block without the spatial attention task (scanner version). Comprehension questions were included at the end of training to ensure that participants had understood the task structure. Once this had been verified, participants entered the scanner and completed 300 trials of the decision task only lasting 50-60 minutes (scanner session). Participants then completed the spatial variant of the task for an additional 100 trials outside of the scanner, lasting approximately 20 minutes (post-scan session). Once the post-scan session was complete, participants filled out a short debrief questionnaire. The experimental task paradigm was created using PsychoPy (Version v2021.1.2).

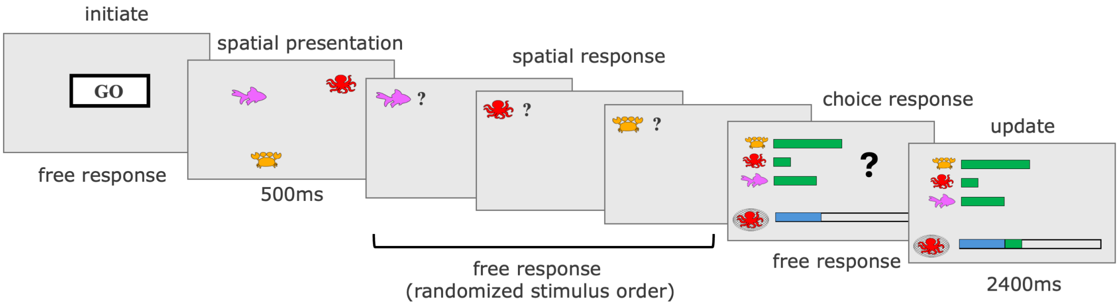


**Figure 3.3:** Task sequence. To measure the influence of goal pursuit on spatial attention outside the decision context, we interleaved an attention task before every decision. In the attention task, participants viewed the three goods flash on the screen in random locations, and were then probed on the location of each good. Participants knew their performance in the attention task had no impact on subsequent offers. Note the interleaved spatial task was only performed outside the scanner. Inside the fMRI scanner, participants performed the decision task by itself (due to the practical limitations of indicating spatial responses with a button box).

## 3.2.3 Models and Analysis

### Model-free analyses

We began by performing simple analyses to verify that people were sensitive to the primary elements of the task. Using a logistic regression, we predicted abandonment decisions on the basis of the current goal offer, the best alternative offer, the worst alternative offer, and the current contents of the net. We also looked at reaction times, to determine if the same predictors (with the addition of a binary regressor capturing whether the decision was to stay or abandon) affected how quickly people responded, using a linear regression model.

### Modelling

Because of the need to commit to a good for many trials in order to realise the reward (delivered on the completion of a full net), a good decision is based not only on the current offer, but also the quantity already in the net and projections of future offers. To understand how participants made such choices, we constructed a series of models reflecting increasingly complex possible strategies. Five models with increasing complexity were tested as candidates for describing peoples' subjective evaluation of the offers (Fig.3.4):

1. **Offer-max model.** The agent chooses the largest offer on screen, regardless of the accumulated contents in the net. The value of persisting is therefore equal to the offered quantity for the current goal good, while the value of abandoning to either alternative is also equal to its current offer.

2. **Myopic model.** The agent maximizes the contents of the net on the current trial. This means they will only switch if an alternative offer is greater than the combined accumulated goods and offer for the current goal good. For this model, the value of persisting is equal to the accumulated goods plus the goal item offer, while the value of abandoning to either alternative is simply equivalent to their current offers.

3. **Simple prospective model.** The agent calculates how much progress towards the goal each offer will entail, where progress is the proportion of the remaining unfilled net that will be completed after choice. Intuitively, this model values each option based on how long it will take to fill the net, if the offers were to stay the same over subsequent trials. The value of persisting is therefore the offer for the current goal item divided by the un-filled length of net. The value of abandoning to either alternative corresponds to their offers divided by the full length of net (since switching would require starting from scratch).

   $$V_{goal} = \frac{O_{goal}}{T - G} \tag{3.1}$$

   $$V_{alt} = \frac{O_{alt}}{T} \tag{3.2}$$

   Where $O$ refers to the current offer for that item, $T$ refers to the target quantity (size of the current net), and $G$ refers to the quantity of goods already accumulated in the net.

   A central difficulty for a model which estimates value in this way is dealing with negative offers. Negative offers would reverse the respective values, meaning that implausibly, negative offers associated with the goal good are valued lower than negative offers associated with alternative goods. To address this problem, we set the value of negative offers associated with alternatives to their raw (negative) offer, and the value of negative offers associated with the goal option to the proportion of progress they would be losing i.e. the offer divided by the accumulated value.

4. **Stochastic tree-search.** This agent uses information about offer trajectories to simulate possible futures for the different candidate options, choosing the option which is forecasted to complete the net fastest. Specifically, it samples possible future trajectories for the three options and calculates each option value as the (negative) average number of trials until net completion across the iterations, if it were chosen on this trial.

   The same statistics used for creating the experimental offers were used when the model simulates the future trajectories of the options (procedure described in Block Generation). In other words, this model possesses task knowledge of how offers are likely to change over time, and leverages that to compute a better estimate of how long each option will take to fill the net.

   Formally, this consists of calculating value in the following way according to a Monte Carlo procedure:

$$Q(s,a) = \frac{1}{N}\sum_{i=1}^{N} -T_i(s,a) \tag{3.3}$$

   Where $N$ is the number of iterations, which we set as 10000, $T_i$ is the number of trials to complete a net across a simulated trajectory, $s$ defines the current state (the size of the net, the accumulated contents inside the net, and the type of good in the net), and $a$ denotes a particular choice (octopus, crab, or fish). Note that if the agent chooses to persist, then the distance to goal completion will be shorter given that goods have already been accumulated. This means $-T_i$ will typically be higher (depending on the current offers) and thus persisting will typically be preferred.

5. **Stochastic tree-search with switching.** In model 4, the simulated trajectories assume that once an option is chosen, the agent will persist with that particular option until goal completion. In model 5, we extended the model to allow the possibility for an agent to switch within the simulated trajectory, at any time-step. In other words, the agent calculates each option value as the (negative) average number of trials until net completion if they started with choosing that option but are able to switch to another option subsequently. In other words, this model not only possesses knowledge of how the options could change, but also anticipates future switches.

   This model requires a separate value estimate to be used to determine choices *within* the simulated trajectories. Since it is too computationally demanding to use nested sampler models (i.e. the recursion of sampling trajectories within each sampled trajectory makes this model untenable), we instead chose to use the value from model 3, i.e. simple prospective value, for the value estimates determining switching within simulated trajectories.

Since the final tree-search model takes into account the total relevant statistics of the task and all possible actions to estimate value (using a monte carlo sampling procedure), it provides an approximation of the optimal choice in this setting (Metropolis and Ulam 1949; Browne et al. 2012).

### Model Fitting

The decision in this task can be framed either as a choice between three items, or as a choice to persist with the current goal versus abandon for an alternative goal. Initial behavioural analyses revealed that (i) Participants chose the worst alternative offer on fewer than 1% of trials and (ii) The worst alternative offer had no significant impact on either choice behaviour or reaction times. Since this suggests participants were treating the decision as a binary choice between the current goal and best alternative, we decided to model behaviour using a logistic regression capturing decisions to abandon or persist with the current goal.

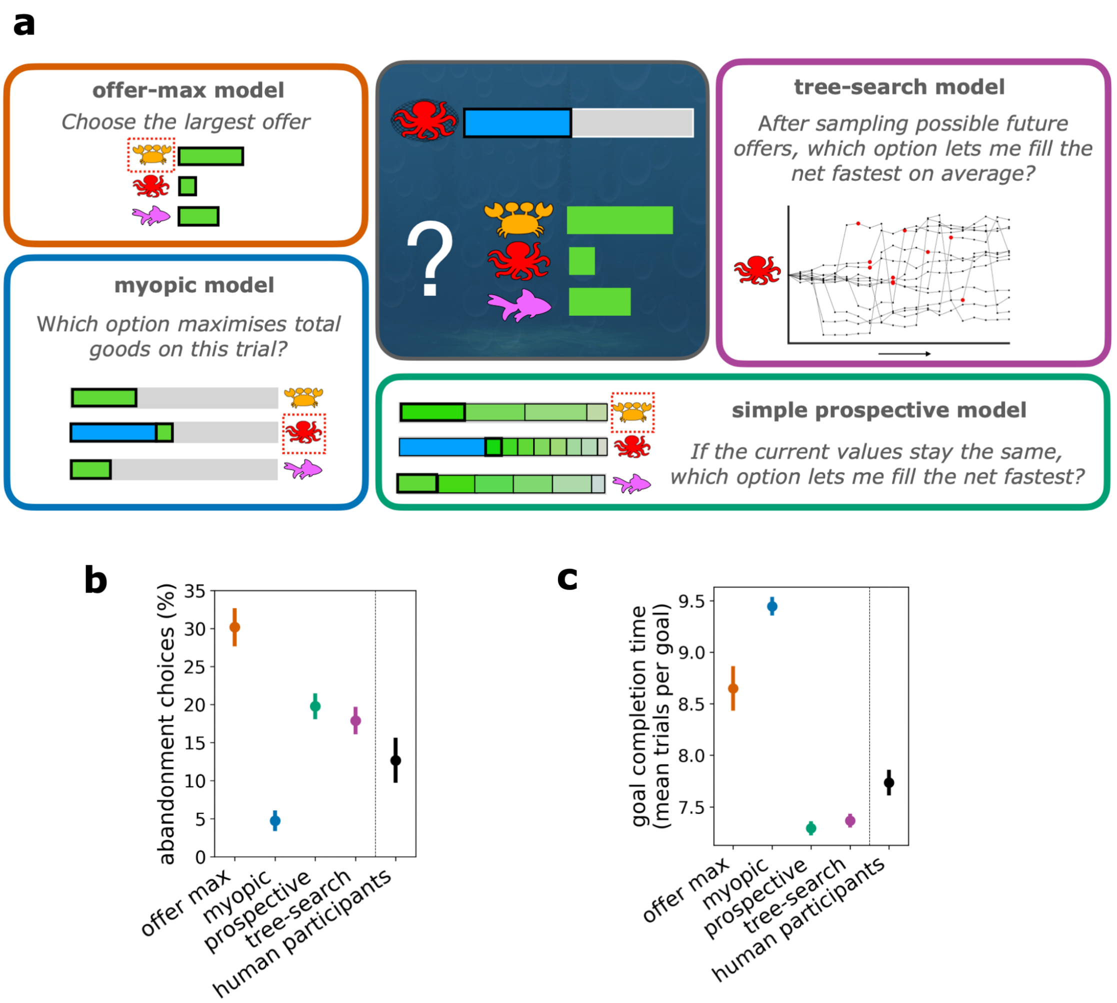


**Figure 3.4:** Behavioural Models. **(a)** Graphic of the four final models which were fit to participant data. **(b)** Simulated abandonment behaviour of the four models, compared to participants. Abandonment is quantified as the percent of decisions (after initial goal choice) to abandon the current goal. **(c)** Performance of the models, compared to participants. Performance is quantified as the mean number of trials to complete the goal, where lower scores indicate faster goal completion (better performance).

For each model, the value of switching was calculated as the model's value for the current goal subtracted from the model's value for the best alternative goal:

$$V_{abandon} = \beta_0 + \beta_1(V_{alt} - V_{goal}) \tag{3.4}$$

Where $V_{alt}$ is the value of the best alternative option, and $V_{goal}$ is the value of the current goal. To determine the best fitting normative model, we predicted choices using mixed effects logistic regression models, where intercept and slope were modelled as random effects across participants.

We then compared performance across models using a leave-one-out cross validation process to evaluate between models, since the models differed in their conceptual complexity but not in the number of fitted parameters. For each participant, we fit each of the mixed-effects model to the choices of all other participants ($n = 29$). For the held-out participant, we then computed the predicted abandonment value for each trial, and transformed this into the predicted probability of switching using the soft-max function:

$$P_{abandon} = \frac{1}{1 + e^{-SV_{abandon}}} \tag{3.5}$$

We took the absolute difference between the predicted probability of switching, and each held-out participant's true responses, and subtracted from 1 to compute the model accuracy for each participant separately. This allowed us to evaluate both the overall accuracy of each model in predicting choices, as well as the frequency of best-fitting models across participants. We also separately assessed the ability to predict abandonment trials and persistence trials, assessing the accuracy of each model for the two trial types separately.

**Model Validation Process** We performed a model recovery analysis, to confirm that we could successfully recover each generative model from its simulated data. This process revealed that no models were confused with each other, with the exception of models 4 and 5 (i.e. the tree-search models with and without the addition of switching during simulated trajectories). Since simulated behaviour from these models could not be successfully discriminated, we did not use model 5 for any empirical analyses, and do not report it further. Instead, we report the results from fitting the four models which could be successfully discriminated through model confusion: offer-max, myopic, simple prospective, and stochastic tree search. Methods and results of the model recovery procedure are included in the appendices (Fig.A.1).

**Persistence Bias** A major theme in this thesis is understanding over-commitment to a goal. To quantify this for each individual, we used the metric of persistence bias, defined as the tendency to persist with the current goal beyond the predictions of the tree-search model, which approximates optimal choice behaviour. Since the tree-search model also provided the best description of peoples' behaviour compared to the other models, it provided a useful metric for comparing differences across participants. For each participant separately, we fit a logistic regression model predicting their abandonment choices using the tree-search value of abandonment:

$$V_{abandon} = \beta_0 + \beta_1(V_{alt} - V_{goal}) \tag{3.6}$$

We quantified the point at which each participant was indifferent to abandoning the goal, compared to the tree-search model (i.e. the 'shift' on the sigmoid function in Fig.3.7a). In other words, this is the point at which a participant is equally likely to persist or abandon, quantified in terms of its deviation from approximately optimal choice. Mathematically, this is equal to:

$$bias = \frac{-\beta_0}{\beta_1} \tag{3.7}$$

Where $\beta_0$ and $\beta_1$ refer to the intercept and slope respectively from the logistic regression predicting participant abandonment choices (Eq. (3.6)). Note that persistence biases are highly correlated with model-free metrics of persistence, such as a participant's total number of abandonment choices (see Fig.A.5), but provide a more sensitive measure because they are not affected by differences in schedule and indicate deviation from approximately optimal choice. Since the persistence bias metric captures a deviation from optimal choice behaviour, higher persistence biases are correlated with worse performance (as shown in Fig.A.6).

Since persistence biases violated tests of normality, we use the one-sample Wilcoxon signed rank test to determine whether indifference points were different to zero, showing a bias compared to the tree-search model. See Fig.A.2 showing persistence biases and sub-component parameters can be accurately recovered after simulating behaviour within the empirical parameter range.

**Test-retest reliability of persistence** Persistence biases are an important individual difference metric in this thesis. We investigated the test-retest reliability of the persistence bias parameter across the two testing sessions (inside and outside the scanner), by using the intraclass correlation (ICC). The ICC captures the agreement across measurements while allowing for baseline differences across sessions (Shrout and Fleiss 1979) We used the ICC(2k) score (absolute-agreement, two-way random-effects model) as used in similar paradigms (Loosen et al. 2022) where the conventional approach considers scores below 0.5 as 'low', between 0.5 and 0.75 as 'moderate', and above 0.75 as 'good' (Koo and Li 2016). For all future analyses involving persistence biases, we used the values fit to the aggregated data across both sessions unless explicitly indicated otherwise. Persistence biases across the two sessions are shown in Fig.3.7b.

We also investigated the stability of the two sub-parameters from which persistence bias is derived (intercept and slope) across the two behavioural testing sessions (inside and outside the scanner). All three parameters show good test-retest reliability, although note that persistence biases have both higher recoverability from simulated data (see Fig.A.2) and higher test-retest reliability across empirical sessions (see Fig.A.4) than either subcomponent parameters on their own.

**Goal progress** How does goal pursuit affect decision processes? To investigate this, we performed a series of analyses to model how decision-making is affected as people progress towards the goal. We define goal progress as the proportion of the current goal completed (i.e. current net contents / net size). First, we asked whether people became more biased in favour of persisting with the goal (compared to what the tree-search model would do) as they progressed with the goal. To quantify the additional impact of goal progress on peoples' abandonment choices, we added goal progress as an additional regressor alongside tree-search value. We used chi-squared tests to determine whether this addition improved our basic mixed-effects model across participants.

$$V_{abandon} = \beta_0 + \beta_1(V_{alt} - V_{goal}) + \beta_2 GP \tag{3.8}$$

Where $GP$ refers to goal progress (i.e. proportion of goal completed), and $V_{goal}$ and $V_{alt}$ refer to the tree-search value for the current goal and best alternative goal respectively. Next, we asked whether people were not only more reluctant to switch over the course of goal pursuit (influence of goal progress), but also less sensitive to the value of switching (interaction between goal progress and value). We also asked whether there was a significant interaction between goal progress and tree-search value, indicating that people were less sensitive to the value of switching over the course of goal pursuit:

$$V_{abandon} = \beta_0 + \beta_1(V_{alt} - V_{goal}) + \beta_2 GP + \beta_3(V_{alt} - V_{goal})GP \tag{3.9}$$

On finding that people were less sensitive to value over time, we asked how this related to our attentional hypothesis: specifically, that people will ignore alternative goal value more than current goal value as attention is increasingly oriented towards the current goal. We split the aggregate value of abandonment into its two components associated with the current and best alternative goals. We fit this final model capturing asymmetry in value use at the level of individual participant. For each participant, we fit a logistic regression model which included the interaction between each source of value and goal progress:

$$V_{abandon} = \beta_0 + \beta_1 V_{alt} + \beta_2 V_{goal} + \beta_3 GP + \beta_4 V_{alt} GP + \beta_5 V_{goal} GP \tag{3.10}$$

$T$-tests were used to determine whether there was a significant difference between the disappearance of current-goal and alternative-goal value across goal pursuit, using coefficients fit for each individual separately. We first tested whether the interaction between value and goal progress differed from zero separately for each source of value, indicating a change in uses of both value components across goal pursuit. We then tested for the difference in slopes between these coefficients (difference in $\beta$ weight for $V_{alt}GP$ interaction and for the sign-flipped $V_{goal}GP$ interaction). Note we use the sign-flipped coefficients for $V_{goal}GP$ because the value of the current goal and the value of the best alternative have opposing impact on the likelihood of switching (see Fig.A.8 for an illustration of this analysis in simulated behaviour from the tree-search model, showing this effect is not an artefact of the task paradigm).

**Reaction time analyses** Earlier analyses had revealed that people slow down in the face of high offers for alternative goals, and for low offers for the current goal. We therefore asked whether peoples' sensitivity to these factors also depended on goal progress. Specifically, we predicted that reaction times would continue to be sensitive to the current goal value but stop being sensitive to alternative goal value over the course of goal progress.

To determine whether this was the case, we predicted trial-wise reaction times using the following regression analysis:

$$rt = \beta_0 + \beta_1 V_{goal} + \beta_2 V_{alt} + \beta_3 GP + \beta_4 V_{alt}GP + \beta_5 V_{goal}GP + \beta_6 switch \tag{3.11}$$

Where the final regressor *switch* captures whether the trial was a persist or abandonment choice. Earlier analyses revealed that people slow down more on abandonment trials, so this regressor was included as a control regressor.

Specifically, we were interested in whether the influence of alternative value would disappear more than the influence of the current goal value over the course of goal progress (difference in $\beta$ weight for the $V_{alt}GP$ interaction and for the sign-flipped $V_{goal}GP$ interaction). As in the previous analysis, we compared the sign-flipped coefficients for $V_{goal}GP$ because the value of the current goal and the value of the best alternative have opposing impacts on reaction times.

### Spatial attention task analyses

The spatial task results come from a separate behavioural testing session after the fMRI session, where participants performed the same decision task with the addition of an interleaved spatial attention task before making each decision (the 'spatial variant' described above). We used this task to measure the relative distribution of attention between stimuli associated with the current goal, and stimuli associated with alternative goals, across goal pursuit. We quantified spatial error as the Euclidian distance between the location of the participant's click and the true location at which the stimulus appeared, in normalised screen units. We quantified reaction times (RT) as the time in seconds ($s$) between when a stimulus was probed (appearing in the top left corner of the screen), and when the participant indicated their response.

We then categorised responses according to whether the probed stimulus was the current goal good or one of the alternatives. We excluded the first trial of every block from analyses, where no goods had yet been accumulated. Since the distribution of mean reaction times and mean error did not violate assumptions of normality, we used t-tests to determine whether mean error differed as a function of the status of the stimulus (i.e. whether the stimulus was the current goal item or an alternative goal item).

We then investigated whether the spatial error bias developed as a function of goal pursuit. We fit two linear models for each participant predicting (a) current-goal stimulus error and (b) alternative stimuli error using the number of trials participants had been pursuing the goal, in each case modelling error using the following linear regression:

$$error = \beta_0 + \beta_1 trials \tag{3.12}$$

Where *trials* corresponds to the trials invested in the current goal. Since the beta weights did not violate assumptions of normality, we used t-tests to determine whether the $\beta_1$ coefficients across participants differed for zero (showing error is dependent on the number of trials invested) for either the current goal stimulus or the alternative goal stimuli. We also tested for the difference between slopes, using a t-test to determine whether trials invested affected error differently for the current versus alternative goal stimuli.

Finally, we investigated whether goal biases in the spatial task were related to persistence biases in the decision task. To capture an individual's goal-oriented attention bias, we took the difference between an individual's mean error for the current goal stimulus, and their error averaged across the two alternative stimuli. We tested for a relationship between an individual's goal-oriented attention bias, and their persistence. Spearman's correlation was used because as previously noted, persistence biases violated assumptions of normality.

## 3.3 Results

### 3.3.1 Validation of task

Choices reflected sensitivity to the task structure (Fig.3.5a). People were more likely to switch when the offers associated with alternative goods were high (impact of best alternative offer on switching: $\beta$=0.53, $t(29)$=14.74, $p < 0.001$), and less likely to switch when the offers for their current good were high (impact of current goal offer on switching: $\beta$=-0.51, $t(29)$=-16.84, $p < 0.001$), or after having accumulated a higher quantity of goods in their net (impact of net contents on switching: $\beta$-0.76, $t(29)$=-27.67, $p < 0.001$). We found no effect of the worse alternative on abandonment decisions (impact of second-best alternative on switching: $\beta$=-0.009; $t(29)$=0.11, $p$=0.912).

Reaction times showed a similar profile (Fig.3.5b). We did not analyse reaction times inside the scanner, because people were forced to wait for a jittered period before responding. However, outside the scanner we found that people slowed down substantially on abandonment trials (mean $RT$=2.59 seconds, std=0.62) compared to trials where they persisted with the same goal (mean $RT$=1.52 seconds, std=0.28). On top of this, people responded more slowly when offered valuable

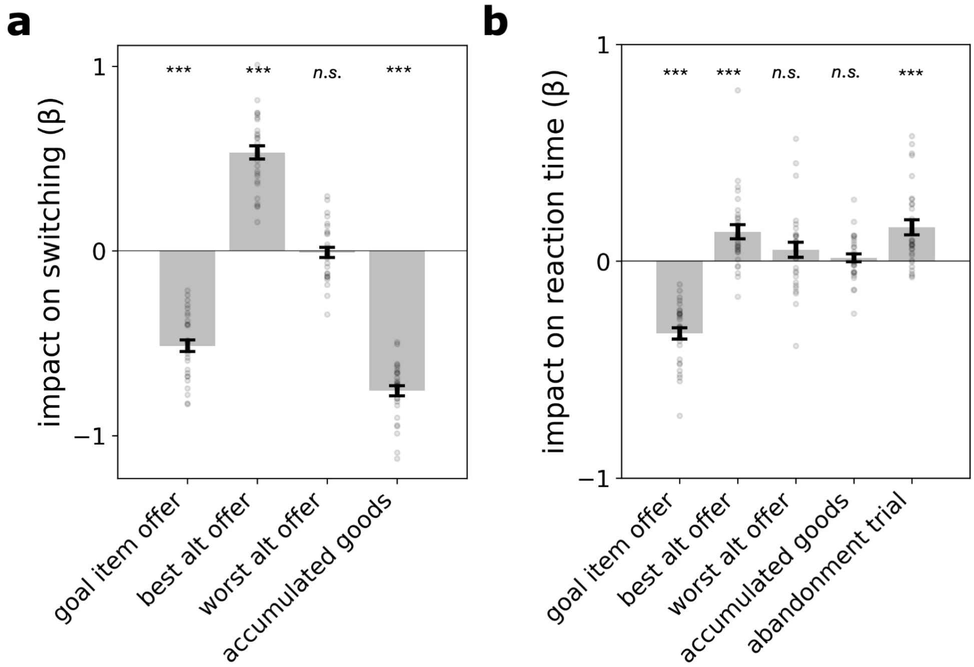


**Figure 3.5:** Simple choice analyses. **(a)** Results of logistic regression analysis performed for each participant, predicting goal switching as a function of simple task parameters. **(b)** Results of linear regression analysis performed for each participant, predicting reaction times as a function of simple task parameters. For both **b** and **c**, error bars show SEM across *beta* coefficients, dots show *beta* coefficients for individual participant, stars indicate significance for two-sided t-tests of *beta* coefficients against zero.

alternatives (impact of best alternative offer on RT: $\beta$=0.14, $t(29)$=4.21, $p < 0.001$) while responding faster when the goal offer was higher (impact of current goal offer on RT: $\beta$=-0.33, $t(29)$=-12.85, $p < 0.001$). Just as for choices, we found no effect of the worse alternative offer on reaction times (impact of second best alternative offer on RT: $\beta$=0.015, $t(29)$=0.80, $p$=0.431, *n.s.*).

On average, people were able to fill 52 nets (std=3.6) within the 400 trials total of the study (mean 7.67 trials to complete a net). Participants switched to collecting a different good on 13.4% of trials excluding the first trial (mean 47 switches per participant), with 53.7% of these switches occurring within the second and third trials of the block. There was wide variety in how much people switched, ranging from 24 to 63 switches (std=9.72) total during the study.

### 3.3.2 Behaviour is best described by a tree-search model

After confirming people made sensible responses, we turned to investigating the behavioural strategies people were using to make these choices. Participants' behaviour was best described by the most complex model we tested ("tree-search model"; see Fig.3.6 for comparison of model fits; see Fig.3.4 for graphic of models). This model samples possible future trajectories for the option offers using the true generative procedure, and selects the option which is predicted to fill the net fastest when averaging across the sampled trajectories, providing an approximation of the optimal decision (monte-carlo sampling). Note that if trials are divided into those where participants chose to abandon versus chose to persist, tree-search model captures choice behaviour best for both trial types (Fig.A.3). This suggests that beyond capturing the total number of abandonment choices better than other models (as shown in Fig.3.4), tree-search model also captures the *timing* of abandonment choices better. That is, just looking at trials where participants choose to persist, tree-search model also captures choice data best.

### 3.3.3 People show goal commitment biases

While general choice strategy was best described by the tree-search model, which approximated optimal choice behaviour, people tended to over-persist with their current goal beyond the predictions of the model (Fig.3.7a; persistence biases were significantly greater than zero: $t(29)=11.23$, $p<0.001$). Persistence biases were quantified as peoples' deviation from the tree-search model, in terms of their indifference point to abandoning the goal (see green dots on Fig.3.7a). While by definition the tree-search model is indifferent to abandonment at a value of zero, people tended to require a higher objective value of abandonment in order to actually abandon their current goal. Since the tree-search model provides an approximation of the optimal choice, people with higher persistence biases tended to perform worse (Fig.A.6; pearson's correlation between persistence and average goal completion time: $r = 0.61, p < 0.001$). This metric of persistence bias

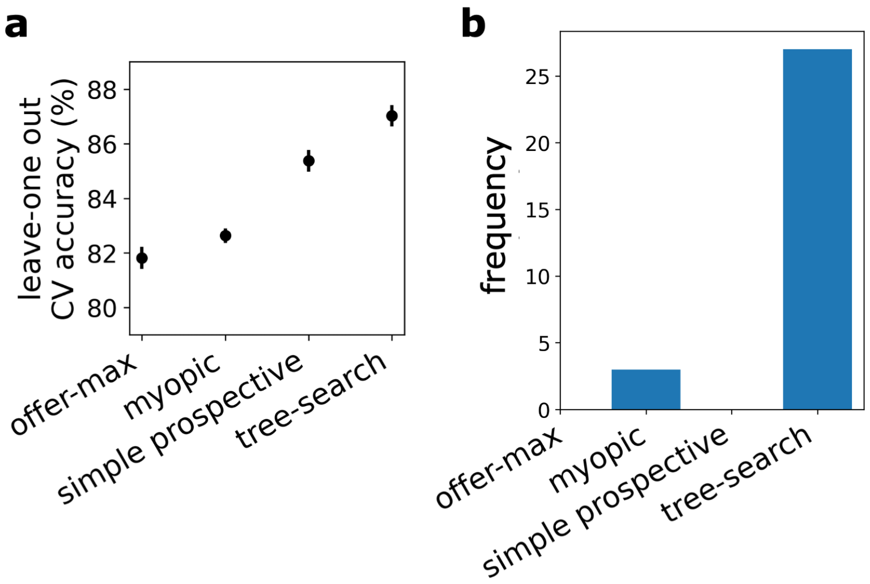


**Figure 3.6:** Model fits. **(a)** Cross-validation accuracy of each model. A leave-one-out procedure was used. For each participant, we fit each of the mixed-effects models to the choices of all other participants ($n-1$). Predictive accuracy for the left-out participant choices was computed using the fitted coefficients. Mean cross-validated performance across participants is plotted, with error bars depicting SEM. **(b)** Frequencies of the best fitting model across the population. For each participant, the best fitting model was assessed using the cross-validated accuracies. The tree-search model was the best fit to choices for 27 out of 30 participants.

had excellent test-retest reliability within participants across sessions (see Fig.3.7b, intraclass correlation coefficient $= 0.76$, $p = 0.002$, 95% confidence interval ($CI$) $= (0.25,\ 1.0)$).

### 3.3.4 People lose sensitivity to alternative goals over the course of goal pursuit

Our task design allows us to characterise commitment biases further, by asking what drives peoples' commitment to goals and what causes people to remain willing to abandon their goal. Specifically, we predicted that changes in alternative goals might influence behaviour less than changes in the chosen goal. How does progress towards a goal affect peoples' sensitivity to the value of alternative goals? We looked at how both decisions and reaction times might reflect diminished processing of alternative goals.

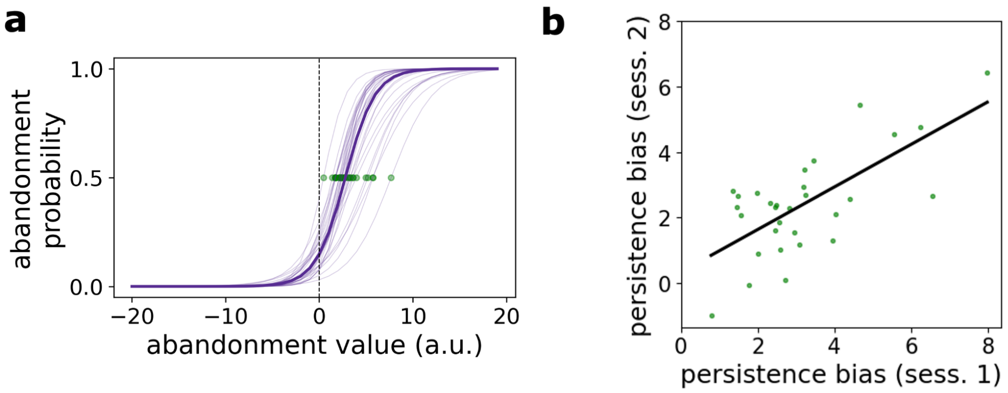


**Figure 3.7:** Persistence bias. **(a)** Probability of goal abandonment as a function of the tree-search value of abandonment. Although the tree-search model captured choices best, people showed an additional bias towards persisting. Bold line shows fits across all participants, transparent lines show individual participants (n=30). Green dots indicate indifference to abandonment, used as the index of individual persistence biases. **(b)** Test-retest reliability of persistence biases across sessions. Session 1 took place inside the fMRI scanner, while session 2 took place outside the fMRI scanner (alongside the interleaved spatial working memory task).

In general, as people progressed towards the goal, they became both more reluctant to switch, and less sensitive to the value of abandonment (Fig.3.8a; abandonment value is defined as the projected value difference between staying with the current goal and switching to the best alternative goal; main effect of goal progress on top of tree-search abandonment value: $X^2(1, N = 30) = 5.27$, $p = 0.022$; interaction between abandonment value and proportion of net completed, on top of both main effects: $X^2(1, N = 30) = 42.43$, $p < 0.001$). We then asked whether this loss of sensitivity equally affected value associated with the current goal versus value associated with alternative goals.

Pressure to abandon the current goal comes from two directions: an alternative good might become more attractive, pulling the agent towards the better option (‘temptation’) or the value of the goal good might collapse, pushing the agent away from the current goal (“frustration”; see Fig.3.1b for example). A rational agent should weigh these two forms of pressure equally when evaluating the options, since value is simply the estimated time in which the target can be completed with each option (i.e. the tree-search already factors in accumulated value; see

Fig.A.8c for analyses on model simulations). We predicted that value associated with alternative goals would impact behaviour less than value associated with the current goal over the course of goal progress.

We found that people indeed showed an asymmetry in their use of these value sources which developed during goal pursuit. As an individual neared goal completion, abandonment was driven less by offers of highly attractive alternatives than by the current goal collapsing, compared to the normative model (Fig.3.8c). To test this, we predicted abandonment choices in a regression model using the interaction between goal progress and each source of value (alongside the main effects). Both sources of value impact behaviour less over the course of goal progress (interaction between alternative value and goal progress: $t(29) = -7.97$, $p < 0.001$; interaction between current goal value and goal progress: $t(29) = 7.08$, $p < 0.001$). However, this loss of influence on behaviour affected alternative goal value more than current goal value (difference between slopes: $t(29) = -3.39$, $p = 0.002$; visualised in Fig.3.8c by binning the data). In other words, over the course of goal pursuit, the impact of temptation from alternatives fades more rapidly than the impact of frustration with the current goal.

We then looked at whether reaction times also show this marker of reduced sensitivity to alternative goals during goal pursuit. In general, we found that people slow down both when offered a valuable alternative, and when the current goal drops in value. We asked whether these trends were sustained across goal pursuit. We found that reaction times became less sensitive to good alternatives over time, but became more sensitive to the value of the current goal (Fig.3.8b). Taken together, both reaction time and choice data suggest people lose sensitivity to valuable alternative goals over the course of goal pursuit, whilst retaining sensitivity to their progress with the current goal.

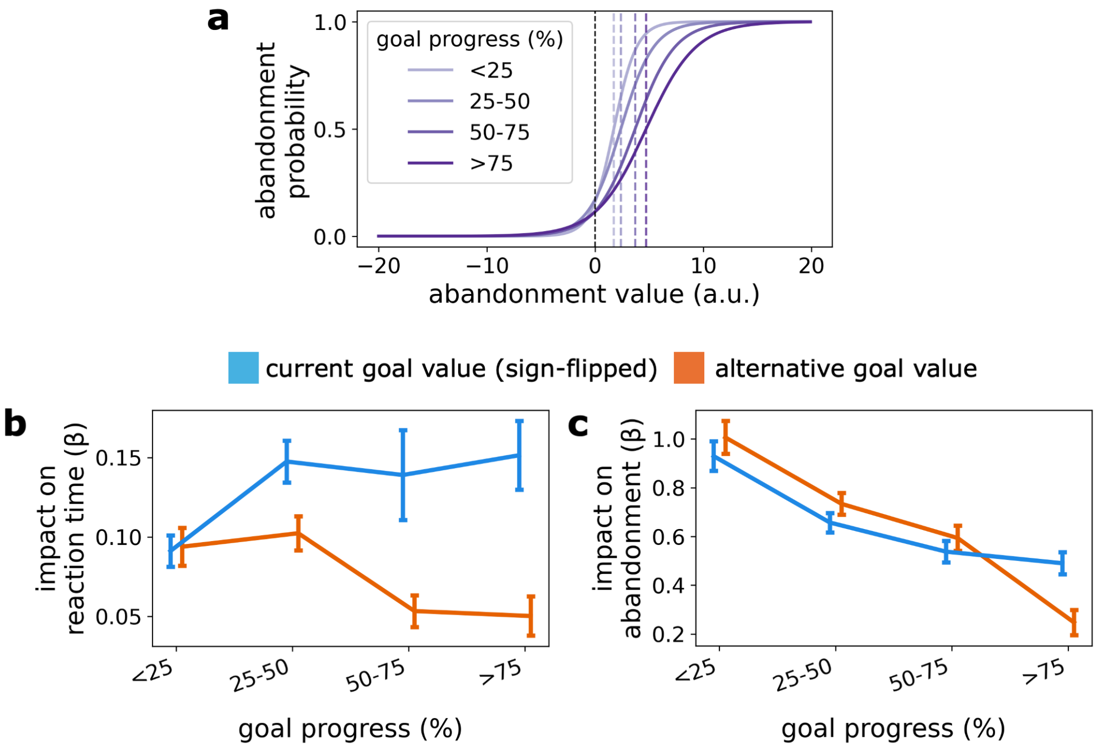


**Figure 3.8:** Development of persistence biases across goal pursuit. **(a)** Across individuals, persistence biases increased with goal progress (i.e. proportion of the net completed). Successive purple lines show probability of abandonment as a function of tree-search abandonment, binned by goal quartile (shown for illustration). **(b)** Impact of value on choice reaction times. Over the course of goal pursuit, reaction times maintained sensitivity to the value of the current goal, but lost sensitivity to the value of the best alternative goal. Orange and blue lines depict the influence of best alternative goal tree-search value and (sign-flipped) current goal tree-search value on reaction times across goal pursuit. Error bars depict SEM of beta weights. **(c)** Impact of value on choices to abandon the current goal. Over the course of goal pursuit, abandonment was driven less by temptation (high alternative goal offers) compared to frustration (collapse in the current goal value). Orange and blue lines depict the influence of best alternative goal tree-search value and (sign-flipped) current goal tree-search value on abandonment choices across goal pursuit. Error bars depict SEM of beta weights.

### 3.3.5 Goal commitment is linked to higher goal-directed attention

We predicted that attention and decision-making biases would be related during goal-pursuit. To measure attention biases, we investigated how attention was distributed between stimuli associated with the current and alternative goals in a decision-free spatial attention task interleaved between decisions. Since the spatial attention task was not possible to perform using a button box inside the scanner, we investigated these attention biases in a separate testing session conducted outside the scanner. In the post-scan session, trials of the spatial attention task were interleaved with new trials of the main decision task.

In the spatial attention task, participants were asked to report the location of briefly-flashed fish, octopus and crab symbols, using a mouse click. Indeed, participants were both more accurate and faster at reporting the location of the current goal stimulus compared to the alternative goal stimuli (Fig.3.9a,b; difference in accuracy for current goal vs alternative: $t(29) = 2.25, p = 0.032$; difference in RT for current goal vs. alternative: $t(29) = 3.30, p = 0.003$). This accuracy difference was primarily driven by progressive memory enhancement for the goal stimulus: spatial accuracy for the current goal stimulus increased with the number of trials participants had been pursuing the current goal (Fig.3.9c; effect of pursuit time on goal item accuracy: $t(29) = -2.65, p = 0.013$; there was no significant effect of pursuit time on accuracy for alternative stimuli: $t(29) = -0.033, p = 0.974, n.s$). In a direct comparison, there was a significant difference between slopes for the effect of goal pursuit on selected and alternative goal items ($t(29) = -2.37, p = 0.024$). This effect occurred despite the fact that the task occurred outside the decision period, and that participants knew their performance on this interleaved task would not affect subsequent offers, suggesting a true attentional bias towards the chosen goal, that increases with goal commitment.

This metric of attentional goal capture directly predicted individual differences in persistence biases: people who showed more attentional capture by the current goal demonstrated higher persistence biases (Fig.3.10; correlation between spatial bias and persistence bias. Note that this relationship holds even when attention-biases and decision-biases originate from separate behavioural testing sessions (using persistence biases fit to data from scanner-only session: Spearman's $r = 0.50, p = 0.005$; Using persistence biases from data aggregated across both scanner and post-scan sessions: Spearman's $r = 0.53, p = 0.003$). This demonstrates that an individual's tendency to over-persist with the current goal is related to their allocation of selective attention towards the current goal.

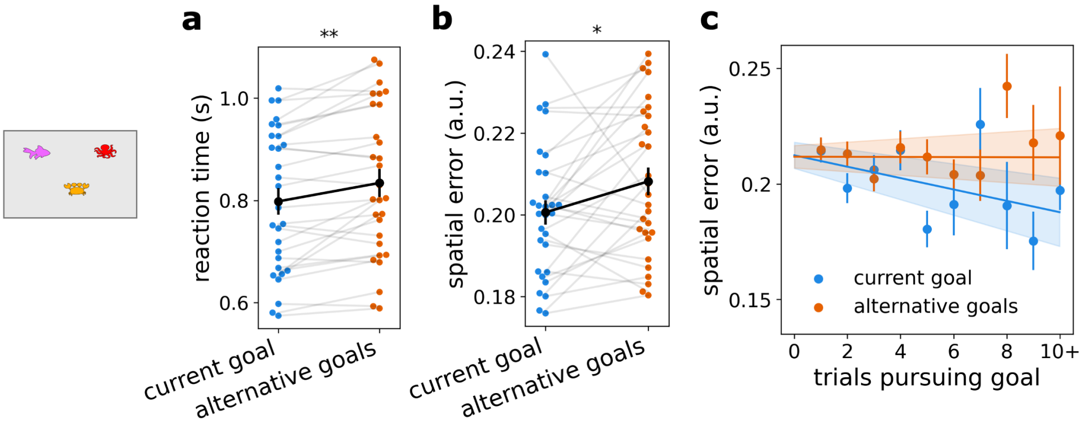


**Figure 3.9:** Performance in the attention task. **(a)** In the interleaved spatial task, reaction times were lower for the current goal stimulus (blue) compared to alternative goal stimuli (orange). Note the order in which stimuli are probed is randomised. **(b)** Error was lower for the current goal stimulus compared to the alternative goal stimuli. Error is measured as the Euclidian distance on the screen between the true location of the stimulus and the reported location. SEM of RT is depicted, stars indicate two-sided paired t-test. **(b)** As participants invested more trials in a particular goal, spatial error decreased for the goal stimulus (blue), but not for alternative goal stimuli (orange). Mean error is plotted against trials pursuing the goal; dots show binned means and SEM, with added regression lines (shaded region indicates SEM of regression lines across participants).

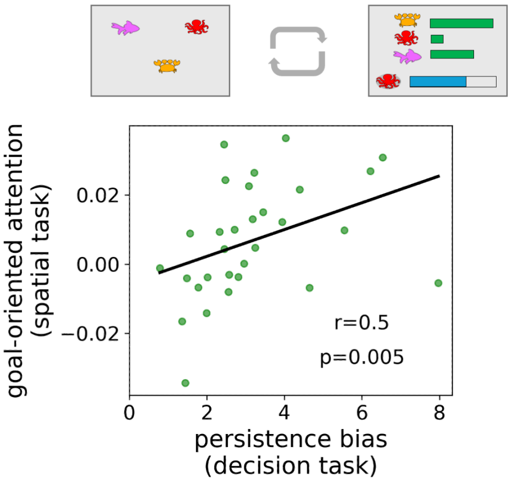


**Figure 3.10:** Individuals showing greater goal-oriented attention (in the interleaved attention task) had higher persistence biases (in the choice task). Attention bias corresponds to the difference in spatial error for the current goal compared to alternative goals. Persistence biases and attention biases come from separate testing session data (inside and outside the scanner respectively).

## 3.4 Discussion

Many rewards are only obtained after a period of persistent effort. Therefore a key challenge for agents is to maintain a balance between commitment with the current goal and flexibility if it ceases to be worthwhile. In this chapter, we developed a pair of complementary tasks to measure how attentional and decision-making biases develop together during incremental goal pursuit. We present evidence that the shift towards goal commitment relates to goal-oriented selective attention.

In the decision-making task, commitment to a goal is required in order to realise rewards, but to perform well at the task participants must also remain sensitive to changes in the value of the current and alternative goals. It is well known that people tend to over-persist with chosen goals (the 'sunk cost' fallacy) (Arkes and Blumer 1985). Consistent with this literature, participants in our task tended to persist with goals longer than was optimal. However rather than consider this behaviour as a biased weighting of decision variables, our evidence supports the theory that it emerges as a consequence of sustained alterations in attention favouring the current goal. Mechanisms of selective attention prioritise processing of the current goal over alternative goals, resulting in reduced sensitivity to attractive alternatives ('temptation'). This reduced sensitivity manifests in two behavioural metrics: (i) over time, reaction times stop responding to the value of alternative goals (while continuing to respond to the value of the current goal), and (ii) decisions to abandon the goal are driven less by temptation from alternative goals compared to frustration with the current goal.

We further probed this attentional account by interleaving the decision-task with an unrelated and decision-free spatial working memory task. We found that participants were better able to recall the location of stimuli associated with the current goal, and this tendency increased as they pursued the goal for longer. Furthermore, there were stable individual differences in persistence with a goal, which were predicted by individuals' sustained goal-directed attention outside the decision period. Individuals who were more biased to persist with a goal showed higher goal-oriented selective attention, even when these metrics were captured in separate testing sessions and despite the fact that performance in the spatial task did not influence the goal pursuit task.

People who over-persist show greater goal-directed attention. What is the direction of this relationship? On the one hand, attention could affect persistence: when attentional biases prioritize the current goal option, the likelihood of selecting the goal is increased. This is consistent with studies showing that exogenously guiding attention to stimuli increases the likelihood of their selection (Armel et al. 2008; Shimojo et al. 2003; Schonberg et al. 2014; Salomon et al. 2018). It is also consistent with our behavioural modelling showing that people lose sensitivity to the value of alternative goals, suggesting alternatives are not entering the decision.

Persistence could also affect attention: higher evaluation of the current goal relative to alternatives could lead to greater attentional capture by the goal option (even though value is irrelevant, such as during the interleaved spatial attention task). Previous studies have shown that historic value modulates visual attention, even in scenarios when value is not currently relevant or even detrimental to the task at hand (Anderson et al. 2011; Anderson and Yantis 2013; MacLean and Giesbrecht 2014; Pearson et al. 2016; Le Pelley et al. 2015; Gluth et al. 2018).

These explanations are not mutually exclusive and their interaction could escalate goal persistence. However importantly, the loss of sensitivity to changes in alternative goals (demonstrated both through reaction times and decisions) cannot be explained purely through higher evaluation of the current goal (i.e. an additive bias). Rather, it suggests that alternative value is contributing less to decision processes later in goal pursuit, consistent with our attentional account.

## 3.5 Contributions

This chapter includes work from an upcoming publication in Nature Human Behaviour: *Holton, E., Grohn, J., Ward, H., Manohar, S.G., O'Reilly, J, & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention.*

The conceptualisation, experimental design and analyses in this chapter were supervised by Nils Kolling, Jill O'Reilly and Jan Grohn.

# 4

# Neural mechanisms of goal pursuit

*This chapter investigates how goal pursuit is supported in neural activity, using fMRI. We find a network of medial prefrontal regions centred on vmPFC which continue to track progress with the current goal between decisions. We find that individual differences in behaviour discovered in* ***chapter 3*** *(both biases to persist with a goal and goal-oriented attention) are predicted by baseline activity in vmPFC. These findings offer a mechanism for how goal pursuit is supported in neural activity, through sustained representations of the current goal in vmPFC which bias attention and subsequent choice. We also examine how value-related activity at the time of decision is modulated by goal pursuit. Mirroring our behavioural finding that people lose sensitivity to alternative value as they near the goal state, ventral striatum also shows a reduction in alternative goal value signals (but not current goal value signals) over the course of goal pursuit.*

## 4.1 Introduction

### 4.1.1 Research aims

This chapter examines the neural mechanisms supporting goal pursuit. Our analyses are motivated by our main behavioural findings in **chapter 3**. Previously, we found an asymmetry in how value affected decisions that unfolded over the course of goal pursuit. Specifically, we found a relative loss in sensitivity to value relating to alternative goals (‘temptation’) compared to value relating to the current goal. Our first question was to ask how this behavioural asymmetry related to neural value signals.

Second, we found that goal pursuit had a sustained impact on behaviour, even outside the decision context. Specifically, we found goal-oriented biases in attention in an interleaved spatial attention task which was irrelevant for goal pursuit. We reasoned that neural regions involved in these pervasive changes in processing might similarly show goal-related activity persisting outside the decision period. To investigate this, we looked at how a critical contextual variable – namely progress with the current goal – was carried in neural activity between decisions.

Finally, behavioural analyses revealed an intriguing individual difference metric. Peoples’ biases towards persisting with the current goal showed stability across testing sessions, and also correlated with their goal-directed attention in the interleaved spatial attention task. How do these individual differences in capture by the current goal relate to neural signals?

### 4.1.2 Value signals during goal pursuit

Our study presents participants with the choice between persisting with a current goal versus abandoning progress for alternative goals. A helpful framework for understanding this form of decision comes from the foraging literature, which presents natural decisions as choices between a default option and a non-default option (Stephens and Krebs 1986; Kolling et al. 2012; Hayden et al. 2011). One previous study defined the notion of a default option as ‘the option that would be selected in the absence of further information processing about its value relative to alternatives’ (Lopez-Persem et al. 2016). In our study, continued pursuit of the goal would be the natural default under this definition.

Previous studies have implicated dACC in representing the value of switching to the non-default option (Kolling et al. 2012; Blanchard and Hayden 2014; Fouragnan et al. 2019; Kaiser et al. 2021; Hayden et al. 2011; Tervo et al. 2021; Trudel et al. 2021). For example, dACC tracks the value of leaving the current patch, and is involved in decisions to switch away in foraging tasks (Kolling et al. 2012; Hayden et al. 2011). We therefore predicted that dACC activity would correlate with the value of pursuing alternative goals in our task.

In contrast, activity in vmPFC often correlates with the value of the default option (Lopez-Persem et al. 2016). In general, many studies have found vmPFC flexibly represents choice values according to the current goal when the goal is determined by the experimenter (Grueschow et al. 2015; Rudorf and Hare 2014; Castegnetti et al. 2021; Trudel et al. 2021; Park et al. 2021). We therefore predicted that in our study, vmPFC would correlate with the value of persisting with the current goal. Finally the ventral striatum is also a centre of value-guided choice (Jocham et al. 2011), which is known to be sensitive to goal proximity (Howe et al. 2013). We predicted that activity in ventral striatum would relate to progress with the current goal, and show similar value-related activity to vmPFC (Piray et al. 2016).

We first confirmed that these brain networks for value-guided decisions correlated with the key elements of our task, as predicted by previous studies. We then turned to analyses inspired by our behavioural finding that people were less sensitive to alternative value as they progressed towards the goal. Specifically, we asked how neural value signals were modulated over the course of goal pursuit. We found that ventral striatum showed the same marker as behaviour: namely a reduction in sensitivity to alternative value over the course of goal pursuit, alongside sustained sensitivity to the value of the current goal.

### 4.1.3 Sustaining the goal

In the second set of analyses, we examined the sustained impact of goal pursuit outside of the decision. Medial PFC has been implicated in carrying information about the current context even when it is not visually observable (Wilson et al. 2014), and even between decisions (Bari et al. 2019). Therefore, we predicted that medial prefrontal areas would carry information relevant for goal pursuit in the inter-trial period. We focussed on progress with the current goal as a critical contextual variable carried over across trials. We found this contextual variable was held in a subset of medial prefrontal areas between decisions, with the peak of activity in vmPFC.

Finally, we investigated how individual differences in persistence and attention related to our neural findings. VmPFC activity has long been implicated in correlating with the idiosyncratic values that people place on different attributes of reward (Levy and Glimcher 2012), including individual differences in willingness to pay for items (Plassmann et al. 2007), aversion to loss (Tom et al. 2007), or delay discounting (Kable and Glimcher 2007). What are the mechanisms through which vmPFC activity could influence the decision process according to these preferences? A line of research has identified a key role for baseline (pre-stimulus) vmPFC activity in carrying contextual information which biases subsequent choices (Lopez-Persem et al. 2016; Vinckier et al. 2018; Abitbol et al. 2015). For example, one study linked peoples' prior expectations about their subjective preferences within categories of food or music to baseline activity in vmPFC (Lopez-Persem et al. 2016).

Given that we had already found sustained goal-related activity in vmPFC between decisions, pre-stimulus activity in this area was a prime candidate for investigating individual differences in goal pursuit. We found that differences in persistence with a goal, as well as goal-oriented attention, both correlated with goal-related baseline activity in vmPFC.

## 4.2 Methods

### 4.2.1 Study procedure

Thirty participants took part in the fMRI study. For a full description of the participants, task and training, see Methods from **chapter 3**. Note that participants did not perform the spatial attention task inside the scanner due to the impracticality of using a button box to report spatial locations. However, we kept the structure of

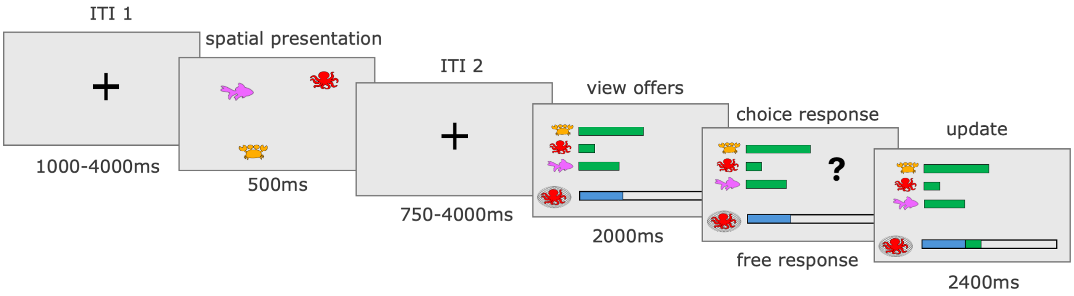


**Figure 4.1:** Task timeline in the scanner. To keep the task visually consistent with the spatial session outside the scanner, participants passively viewed the three sea creatures flash on screen during the inter-trial interval, but were not required to report the location of the creatures. To dissociate activity related to the decision from activity related to response indication, we included a two second buffer zone once the offers were presented, before participants could make their response. In the main fMRI analyses, activity was time-locked to the onset of the decision period, shown here as **(view offers)**. In the additional ITI analysis, activity was time-locked to **(ITI 1)**.

the task similar by including the flash of the three creatures on the screen, although participants were not probed on the locations (Fig.4.1). The original objective for this inter-trial spatial presentation was to measure attentional capture through eye tracking. However, regrettably due to malfunctioning eye tracking equipment, we were unable to attain this data. Inside the scanner, participants performed 300 trials of the decision task. On each trial, they selected the creature of choice using a button box where the first three buttons corresponded to the top, middle and bottom creatures on the screen.

### 4.2.2 Data acquisition

The fMRI data were collected at the Oxford Centre for Human Brain Activity using a 3T Siemens scanner with a multiband accelerated echoplanar imaging sequence with the following parameters: voxel resolution 2.4 x 2.4 x 2.4 $mm^3$, repetition time=1230$ms$, echo time=30$ms$, flip angle=60°, field of view=240$mm$, multiband acceleration factor=3, PAT factor=2, encoding direction=PA. A tilt angle of 30° was used to minimize signal drop out in the orbitofrontal cortex (Deichmann et al. 2003). Data was collected in two consecutive runs of approximately 25 minutes, where participants stayed in the scanner between runs. Fieldmaps were acquired using a dual echo 2D gradient echo sequence ($TE_1 = 4.92ms$, $TE_2 = 7.38ms$ and $TR = 590ms$, image resolution of $2.4mm$). T1-weighted structural images were also acquired at a voxel size of 1.0 x 1.0 x 1.0 $mm^3$, $TE = 3.96ms$, $TR = 1900ms$.

### 4.2.3 Pre-processing and analysis structure

Data were pre-processed using FMRIB's Software Library (FSL), using the FEAT software tool (Woolrich et al. 2001) Functional data were motion corrected using rigid body registration to the central volume (Jenkinson and Smith 2001; Jenkinson et al. 2002). Gaussian spatial smoothing was applied with a full-width half-maximum of 5mm, and high pass temporal filtering was applied with a cut-off of 60s. Cardiac and respiratory data were processed using FSL's Physiological Noise Modelling (PNM) tool to model the effects of physiological noise in the MRI data (Brooks et al. 2008). Since participants completed the MRI session in two runs, parameter estimates were first estimated at the level of run (first level), then combined within individuals as Fixed Effects (second level), and finally combined across subjects using FMRIB's Local Analysis of Mixed Effects (FLAME1+2; third level (Woolrich et al. 2004)). Multiple comparisons were corrected for using a $Z$ statistic threshold of 3.1, and a cluster probability threshold of $p = 0.05$. A double gamma HRF function was used.

### 4.2.4 Whole-brain analyses

Two whole-brain analyses of the data using general linear models (GLM) were conducted. An initial model free analysis was conducted prior to these analyses as a sanity check, and is included in the appendices (Fig.B.2). The results reported in this chapter are based on two analyses:

1. **Decision time analysis.** Our main analysis used value defined by the tree-search model, which was the best predictor of behaviour. Regressors included the tree-search value of the current goal, the best alternative, and worst alternative, goal progress, goal size, and whether the participant chose

to persist or abandon (1/0). Goal progress is correlated with tree-search value, and our behavioural analyses shows it is an additional predictor of abandonment beyond tree-search value (shown in Fig. 3.8a). For this reason, we chose to separate the goal progress component from value in the fMRI analysis. To do this, we residualised all forms of value using goal progress, and used goal progress as an independent regressor on top, enabling us to investigate signals of goal progress separately from information about option offers. In addition, since the tree-search value of an option is an approximation of its time to completion, it is highly dependent on the size of the net across different blocks. To account for this, we also residualised tree-search value to net size, and included net size as a separate regressor. In other words, for each value component (current goal, best alternative, worst alternative), we removed the components related to goal progress and goal size, and added these components as unique regressors. The final correlations between regressors are shown in Fig. 4.2a.

2. **Inter-trial analysis.** Behavioural analyses in **chapter 3** revealed pervasive effects of goal pursuit on attention, even outside of the decision period. We therefore hypothesised that neural areas important for these sustained goal-directed biases might carry information related to the goal between decisions. The inter-trial analysis included all the same regressors as analysis 1. However, in addition to including regressors time-locked to the decision, we also included regressors from the previous trial, time-locked to the previous inter-trial interval ('ITI 1' in Fig. 4.1).

All regressors were z-scored at the level of individual runs before fitting the GLM. In all analyses, reaction times were included as additional control regressors. In addition to the parametric regressors, five types of events were included in the final GLM as main effects: onset of the decision period, onset of the block, spatial presentation of the three stimuli (substituting the spatial task), the update of the net, and the end of the block. Finally, the following confound regressors were included in all analyses: Six motion regressors produced during realignment, the physiological EVs (processed by PNM) and motion outlier events. Motion outliers were detected using FEAT's fsl motion outliers tool. Metric values for detecting motion outliers were calculated for each time-point using the RMS intensity difference between each volume and the reference volume, and outliers were identified as volumes for which the metric value exceeded the 75th percentile + 1.5 times the inter-quartile range. We did not analyse the first trial of each block in parametric analyses of value, as the current goal was not yet defined.

### 4.2.5 Region of interest analyses

Our region of interest analyses aimed to address two questions based on the findings of our behavioural modelling and whole brain analyses. First, we asked how representations of value were modulated by goal pursuit. In other words, did neural areas represent value in a stable manner, or did these change dynamically as people progressed through the goal? Second, we asked how individual variability in persistence was related to our hypotheses about pre-stimulus vmPFC activity. The methodology for these two analyses is described below.

#### ROI selection and extraction procedure

We selected regions-of-interest in three key value-sensitive areas for further analysis. VmPFC, ventral striatum, and dACC all showed strong value-related activity at decision time in our whole-brain analysis. This is consistent with previous literature showing dACC is involved in value-guided abandonment (Fouragnan et al. 2019; Kolling et al. 2012; Tervo et al. 2021), and ventral striatum is a centre of value-guided choice (Jocham et al. 2011), known to be sensitive to goal proximity (Howe et al. 2013), and with meaningful projections to vmPFC (Piray et al. 2016). Given the relevance of these areas for decision-making during goal pursuit, we created regions of interest at the peaks of activity in these areas from our whole-brain analysis, to investigate how the observed activity changed over the course of goal pursuit.

We selected activity peaks for the value of persisting (current goal value–best alternative value; peaks in vmPFC and ventral striatum), and the value of abandoning (best alternative value–current goal value; dACC), following cluster correction. Illustration of the ROIs can be seen in Fig.4.3a,b,c, and all activity peaks are listed in the supplementary materials (Fig.B.4). Since our whole-brain analysis did not reveal any activation for the value of the third alternative in these areas, we did not include the third alternative in subsequent analyses. Regions of interest consisted of spheres with a 3 voxel radius (7.2mm3). In time-course analyses, activity in these spheres was up-sampled by a factor of 10, and cut into epochs which were aligned to the onset of the decision phase.

Activity in these value-related ROIs was then used to investigate the modulation of value signals over the course of goal progress. Any time courses displaying non-orthogonal contrasts are for illustration purposes only and no statistical tests were performed.

### Value modulation analyses

We found an asymmetry in the use of value in behaviour, where the influence of value related to alternative goals disappeared more than the influence of value related to the current goal, over the course of goal pursuit. Therefore, we asked whether neural representations of value in our ROIs also changed over the course of goal pursuit.

Following the analysis for behaviour, we predicted activity in each ROI with the interaction between goal progress and each source of value (tree-search value of best alternative and tree-search value of current goal). Following our analyses of reaction times, we also included a binary regressor controlling for whether the choice was to switch or abandon. This was important because abandonment trials have a large impact on neural activity beyond value, and there are substantially fewer abandonment trials later during goal pursuit. In addition, we used a confound regressor controlling for response times:

$$BOLD = \beta_0 + \beta_1 V_{goal} + \beta_2 V_{alt} + \beta_3 GP + \beta_4 V_{alt} GP + \beta_5 V_{goal} GP + \beta_6 switch + \beta_7 log(RT) \tag{4.1}$$

Where $BOLD$ refers to the neural BOLD activity within each ROI. All regressors were normalised before fitting the GLM. Activity in the ROIs was up-sampled by a factor of 10, and cut into epochs which were aligned to the onset of the decision phase. The model was subsequently fit to each time point within the epoched time-course data. We then multiplied the time-course of fitted beta coefficients by the double gamma HRF function, and summed the products to produce a single coefficient for each regressor (per participant). We selected the coefficients which correspond to the interaction between goal progress and each value source (current goal and alternative goal) at decision-time. We then tested whether the interaction coefficients were significant using one-sampled tests against 0 (one-sided to match our hypotheses for decreasing sensitivity to value over goal progress). Non-parametric methods (Wilcoxon signed-rank test) were used because the distribution of beta weights violated the assumption of normality. Note that although we report one-sided tests due to our behavioural prediction of finding a decreasing impact of value over goal pursuit, our findings remain the same with two-sided tests.

### Baseline activity analysis

Our previous whole-brain analysis found that activity correlating with goal progress was present in a subset of medial frontal areas during the inter-trial interval, with the peak of this activity located in vmPFC. Previous research has shown that baseline tracking of long-term task variables in vmPFC can influence choices (Abitbol et al. 2015; Lopez-Persem et al. 2016; Vinckier et al. 2018; Lopez-Persem et al. 2020). Our goal in this analysis was not to replicate previous findings showing that pre-stimulus vmPFC biases choices on a trial-wise level. Instead, we focussed on the specific question of whether this activity had relevance for the individual differences observed in behaviour.

Following previous paradigms, we define 'baseline activity' as the activity level measured in our vmPFC ROI at the time of stimulus onset, before the new offers or decision process itself influence the dynamics. Rather than using the hemodynamic response function which is inappropriate for analysing pre-stimulus activity (Abitbol et al. 2015), we built a GLM composed of a finite impulse response function (FIR) from four TR (4.92s) prior to stimulus onset, to six TR (7.38s) after stimulus onset. We then extracted regression estimates for the impact of goal progress at stimulus onset. This produced coefficients at the participant level (i.e. for each participant, the extent to which vmPFC activity at stimulus onset tracked goal progress).

Then we specifically tested for a relationship between these individual baseline coefficients and our behavioural measures (persistence bias and goal-directed attention). Note that the metric of attention is taken from a separate session, outside the scanner.

To test the specificity of our vmPFC baseline effect we did two additional analyses. First, we tested whether baseline activity tracking goal progress in the other two ROIs (ventral striatum and ACC) significantly predicted persistence. Second, we investigated whether goal-progress activity time-locked to the decision itself predicted individual behavioural measures. To quantify the decision-related activity, we used the HRF function. We multiplied the fitted beta coefficients for goal progress at each time-point by the double gamma HRF function, and summed the products to produce a coefficient for each participant (same procedure described in **Value modulation analyses** above).

### Mediation analysis

In the previous chapter, we found a significant relationship between inter-trial goal-oriented attention and persistence bias. In this chapter, we found a relationship between each of these behavioural metrics and baseline vmPFC activity in the fMRI scanner. To further investigate these relationships, we performed a mediation analysis. We asked the following question (based on temporal order): could attentional capture by the goal (measured before the decision) affect persistence (the choice) through baseline vmPFC activity at choice onset? In assessing mediation, we performed the following steps (Baron and Kenny 1986):

1. Impact of attention on persistence.

   $$persistence = \beta_0 + \beta_1 attention \tag{4.2}$$

   Where $\beta_1$ must be significant.

2. Impact of attention on baseline vmPFC

   $$vmPFC_{baseline} = \beta_2 + \beta_3 attention \tag{4.3}$$

   Where $\beta_3$ must be significant.

3. Mediation of relationship between attention and persistence by baseline vmPFC.

   $$persistence = \beta_4 + \beta_5 attention + \beta_6 vmPFC_{baseline} \tag{4.4}$$

   Where $\beta_6$ must be significant, and $\beta_5$ must be smaller in value than $\beta_1$ above (or insignificant).

Here, *attention* refers to an individual's goal-oriented attention bias defined as the mean difference in accuracy for reporting the location of the current goal stimulus compared to reporting the location of alternative goal stimuli (see attention task analyses from **chapter 3**). This metric comes from the spatial attention task performed between decisions in the post-scan session. *Persistence* refers to an individual's bias towards persisting with the goal (from the tree-search model; see persistence bias from **chapter 3**). Finally, the metric $vmPFC_{baseline}$ was the mean baseline effect of goal progress on vmPFC activity, as described in the section above.

## 4.3 Results

### 4.3.1 Neural activity related to goal pursuit at the decision

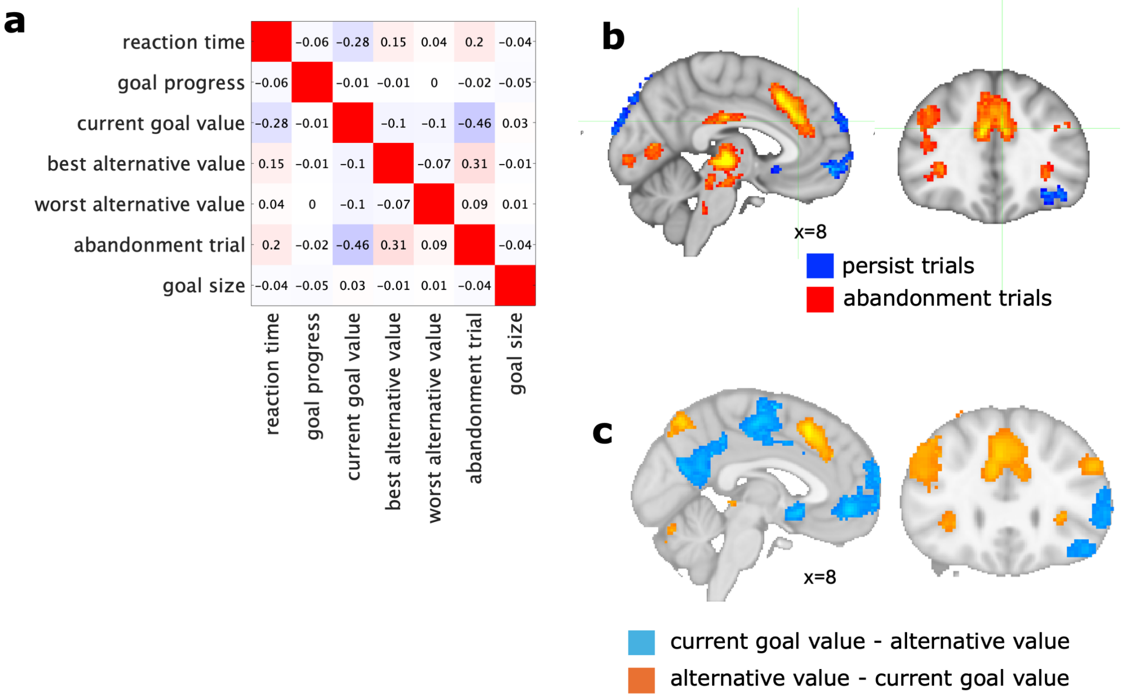
We investigated neural activity at the time of the decision in a whole-brain analysis (regressors time-locked to the onset of the offers). These analyses revealed a broad network of areas sensitive to goal pursuit. A version of this analysis with model-free regressors can be found in the appendices (Fig.B.2).

We found value-related activity consistent with previous studies engaging brain networks in choices between staying with a default versus switching to an alternative (Fig.4.2). Medial frontal pole, vmPFC and striatum increased their activity as the value of persisting with the goal increased (value of current goal–value of best alternative). These areas were also more active on persist trials (compared to abandonment trials). In contrast, dACC, presupplementary motor area (preSMA), bilateral dorsolateral prefrontal cortex (dlPFC), and bilateral insular, all showed the opposite profile: activity increased as the value of abandonment increased (value of best alternative–value of current goal), and activity was higher on trials where the participant chose to abandon the current goal. We included response times as an additional control regressor, previously used as a proxy for choice confidence (Shenhav et al. 2014). We found that dACC activity was also higher when participants were slower to respond, but we found no relationship between response times and vmPFC activity (See appendices; Fig.B.2).

Time-course illustrations of value-related activity in our three ROIs can be seen in (Fig.4.3). For interest, we have included in the appendices illustrations of the neural value profiles on trials where the participant chose to abandon the current goal versus trials where they chose to persist with the same goal (Fig.B.5).

### 4.3.2 Ventral striatum activity reflects reduced sensitivity to the value of alternative goals

As attention to the current and alternative goals varies with goal pursuit, we should expect to see changes in neural representations of these goals. In particular, in



| | reaction time | goal progress | current goal value | best alternative value | worst alternative value | abandonment trial | goal size |
|---|---|---|---|---|---|---|---|
| reaction time | | −0.06 | −0.28 | 0.15 | 0.04 | 0.2 | −0.04 |
| goal progress | −0.06 | | −0.01 | −0.01 | 0 | −0.02 | −0.05 |
| current goal value | −0.28 | −0.01 | | −0.1 | −0.1 | −0.46 | 0.03 |
| best alternative value | 0.15 | −0.01 | −0.1 | | −0.07 | 0.31 | −0.01 |
| worst alternative value | 0.04 | 0 | −0.1 | −0.07 | | 0.09 | 0.01 |
| abandonment trial | 0.2 | −0.02 | −0.46 | 0.31 | 0.09 | | −0.04 |
| goal size | −0.04 | −0.05 | 0.03 | −0.01 | 0.01 | −0.04 | |

**Figure 4.2:** Decision-related activity. **(a)** Matrix showing correlations between regressors in the main decision-time analysis. **(b)** Results from the whole-brain analysis showing cluster-corrected peaks for activity higher on trials where the participant chose to persist with the same goal (blue) versus trials where the participant chose to abandon for the best alternative option (red). **(c)** Activity correlating with parametric value regressors. Colours show cluster-corrected peaks for the contrasts capturing the value of persistence (blue: contrast of current goal value over best alternative value) and the value of abandoning (orange: contrast of best alternative value over current goal value).

behaviour we observed an intriguing asymmetry, namely that as goal commitment increased, sensitivity to alternative goal value ('temptation') was reduced more than sensitivity to the current goal value ('frustration'). We therefore asked how value signals relating to the current and alternative goals change as a function of goal pursuit.

Parallel with our behavioural results, we found an asymmetry between how goal pursuit affected signals relating to alternative and current goal value in the ventral striatum. Specifically, representations of alternative value disappeared in the ventral striatum over the course of goal pursuit, but activity continued to co-vary with the current goal value (Fig.4.4, left; interaction between best alternative value and goal progress: Wilcoxon signed rank, $Z = 2.37$, $p = 0.009$, $n = 30$, $r = 0.43$, one-sided; interaction between current goal value and goal progress:

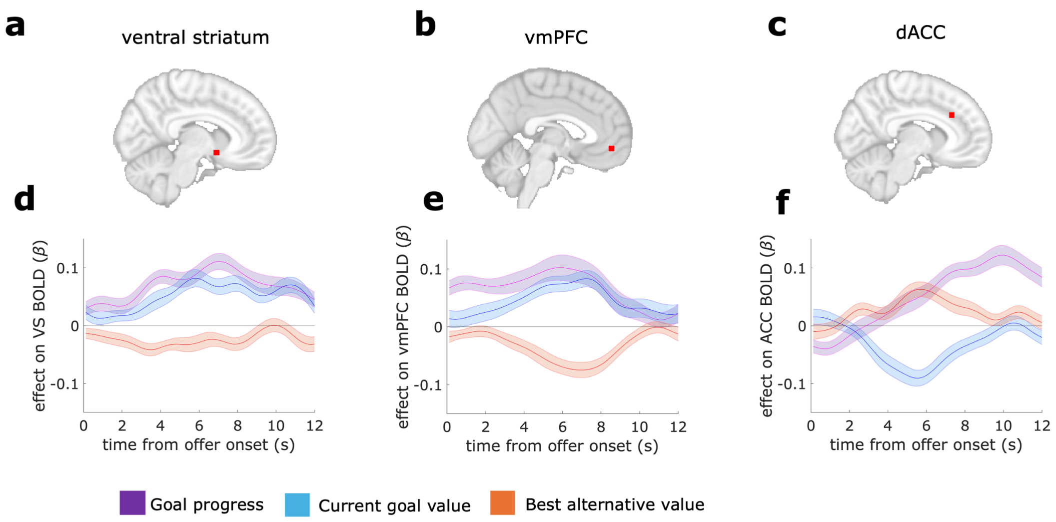


**Figure 4.3:** Time-courses of neural value activity. **(a,b,c)** We extracted regions of interest based on the peaks of value-related activity in our fMRI study. These consisted of the peaks of activity for the contrast of goal value over alternative value in the case of ventral striatum [8,8,-10], and vmPFC [-2,48,-8], and the largest sub-peak of activity in the dACC for the contrast of alternative value over goal value regressors [8,28,30]. **(d,e,f)** Time course analyses depicting the t-statistics for the regressors of goal progress (purple), current goal value (blue), and best alternative value (orange) in the three regions of interest (for illustration). Time 0 seconds corresponds to the onset of the offer stimuli. Mean beta weights are plotted, where shaded error show SEM across participants (n = 30).

Wilcoxon signed-rank, $Z = -1.03$, $p = 0.152$). This mirrored the behavioural finding that people became relatively less sensitive to temptation by alternative goods, whilst maintaining sensitivity to the value of the chosen goal, over the course of goal pursuit. In contrast, there was no significant change in the representation of alternative value over goal pursuit in either vmPFC ($Z = 1.19, p = 0.116$) or ACC ($Z = 0.41, p = 0.660$).

### 4.3.3 VmPFC activity tracking goal progress persists between decisions

Since behavioural analyses in **chapter 3** revealed pervasive effects of goal pursuit on behaviour (attention) outside the decision-making period, we reasoned that neural areas may also track variables related to goal pursuit between decisions. In particular, we looked at the regressor of ‘goal progress’ which captures an individ-

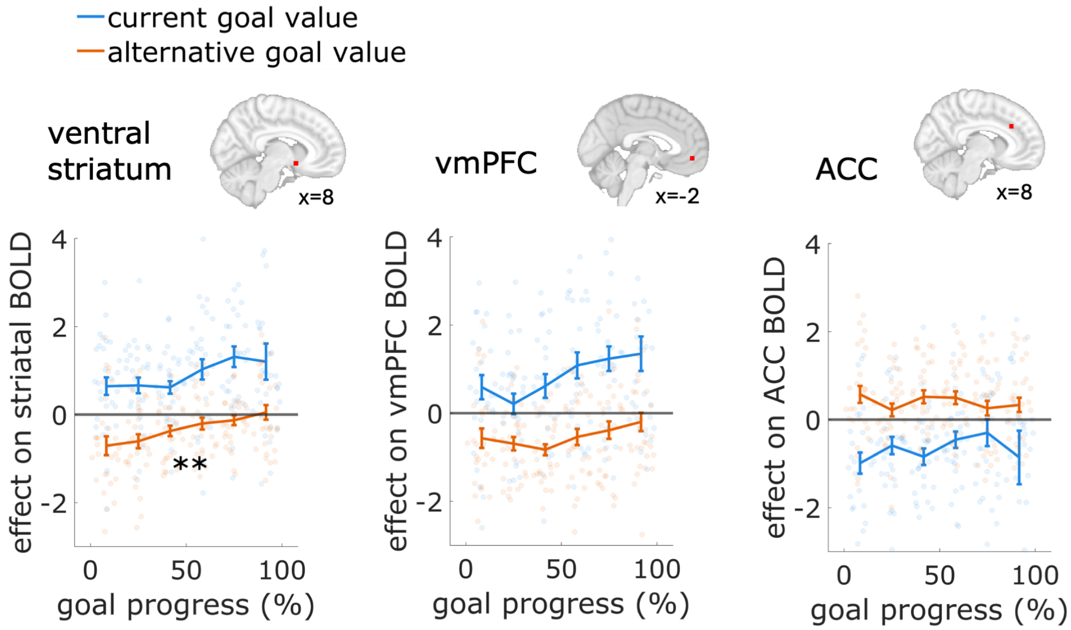


**Figure 4.4:** Modulation of value-related activity in ROIs over the course of goal pursuit, where the red dots in the brain images indicate the ROI location for activity shown in each plot. Here we show the effect of value on the BOLD signal (beta weight) as a function of the proportion of the goal completed, binned for illustration. Blue shows the impact of current goal value, while orange shows the impact of alternative goal value. Error bars depict s.e.m., while dots show beta weights for individual participants ($n = 30$ participants). In the striatum (left), there was a significant reduction in the representation of alternative goal value across goal progress (orange line; stars indicate significant interaction between alternative goal value and goal progress; Wilcoxon signed rank, $Z = 2.37$, $P = 0.009$, $n = 30$), parallel to the reduction in sensitivity to alternative goals seen in behaviour. In contrast, representations of the current goal value were maintained throughout goal pursuit in all ROIs.

ual's position in the goal (proportion of net completed), and is stable across trials.

Activity in a wide range of areas increased as an individual progressed towards completing the goal, including medial prefrontal cortex, striatum, and cingulate areas, as well as large regions of the occipital, and parietal cortices ('goal progress' regressor). However, the results of our inter-trial analysis reveal a much smaller subset of areas which included medial prefrontal cortex and right hippocampus continuing to track progress with the goal even between decisions (Fig.4.5). The peak of activity in medial areas was in the ventromedial prefrontal cortex, vmPFC.

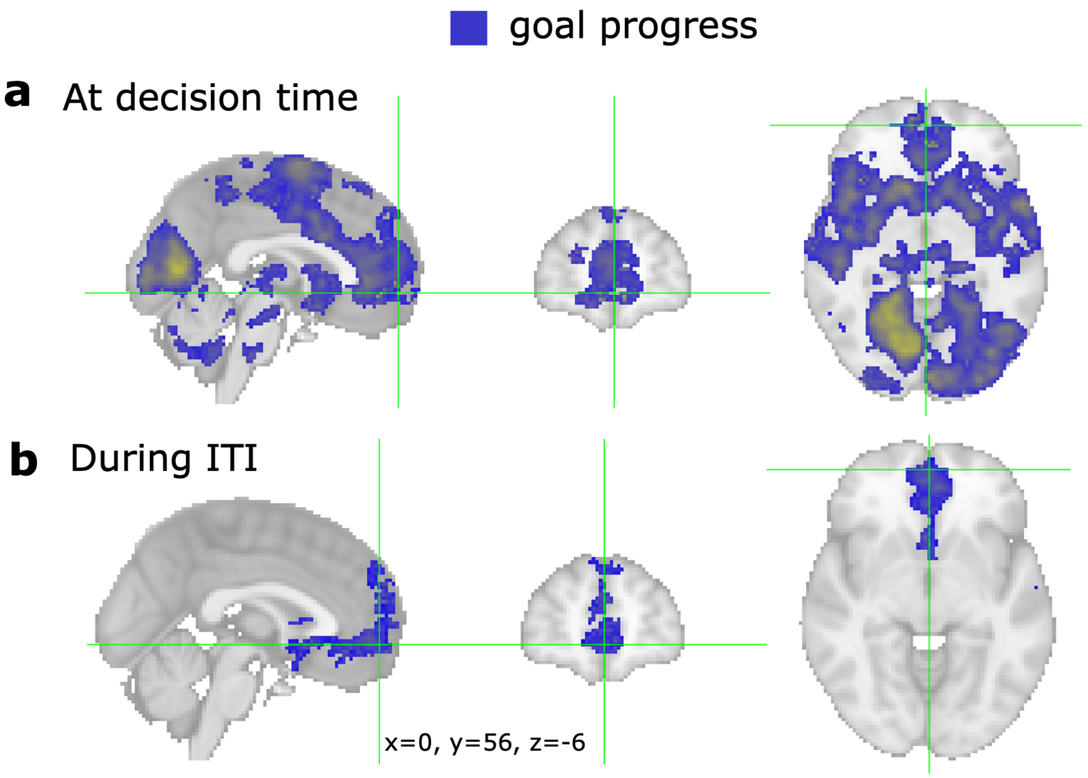


**Figure 4.5:** Activity tracking goal progress at decision versus between decisions. **a** Cluster-corrected activity representing goal progress time-locked to the onset of the decision period. **b** Cluster-corrected activity representing goal progress time-locked to the intertrial fixation cross. While there was widespread activity in the occipital and parietal areas at decision time (a), the majority of these areas did not track goal progress between decisions, where the highest peak was in the vmPFC.

### 4.3.4 Baseline vmPFC activity predicts the degree of goal-commitment across individuals

Previous studies have found that baseline vmPFC activity (activity before a decision) predicts biases or priors which affect subsequent decision-making (Abitbol et al. 2015; Vinckier et al. 2018). As vmPFC tracks goal progress between decisions, we hypothesised that the strength of this baseline vmPFC signal would predict the degree of commitment bias (unwillingness to switch goods) across individuals.

We extracted baseline activity on a trial-by-trial basis in our vmPFC region of interest, and quantified the extent to which pre-decision activity was tracking goal progress for each individual. We found this baseline goal-related activity correlated with an individual's bias towards persisting with the goal during the decision-making task (Relationship between baseline vmPFC and persistence bias: Spearman's $r = 0.49$, $p = 0.006$; Fig.4.6b).

We performed two control analyses to determine the specificity of this effect. First, we confirmed that baseline goal-related activity in dACC and striatum did not predict behavioural persistence (ACC: Spearman's $r = 0.19$, $p = 0.327$ *n.s.*; ventral striatum: Spearman's $r = 0.09$, $p = 0.644$ *n.s.*). Note these ROIs were previously selected because, like vmPFC, they also contained value-related activity. Second, we found that activity within vmPFC related to the decision itself (rather than pre-stimulus) did not show this relationship to behavioural metrics of persistence (Spearman's $r = 0.33$, $p = 0.079$, *n.s.*). For full details and plots of these control analyses, see appendices (Fig.B.6).

If baseline vmPFC activity also relates to the degree to which attention is oriented towards the current goal, we reasoned that it should also correlate with differences in goal-directed attention in the second, decision-free task. This was indeed the case – across participants the strength of the baseline goal-progress signal in vmPFC predicted greater accuracy for the current goal relative to alternative goals in the attention task (Spearman's $r = 0.50$, $p = 0.005$; Fig.4.6a). This was particularly striking as the spatial attention task was carried out in a separate session outside the scanner.

### 4.3.5 The relationship between goal-oriented attention and persistence is mediated by baseline vmPFC activity

We found relationships between goal-directed attention, pre-stimulus goal-related activity in vmPFC, and persistence with the goal. In this final analysis, we used a mediation analysis to probe these relationships. The measurement of baseline vmPFC activity is taken at the onset of the choice, which is between goal-oriented attention (before choice onset), and persistence bias (at choice time). We therefore conducted a mediation analysis to ask whether the relationship between goal-oriented attention and persistence could be mediated by baseline vmPFC. Indeed, our results support this idea, because the relationship between attention

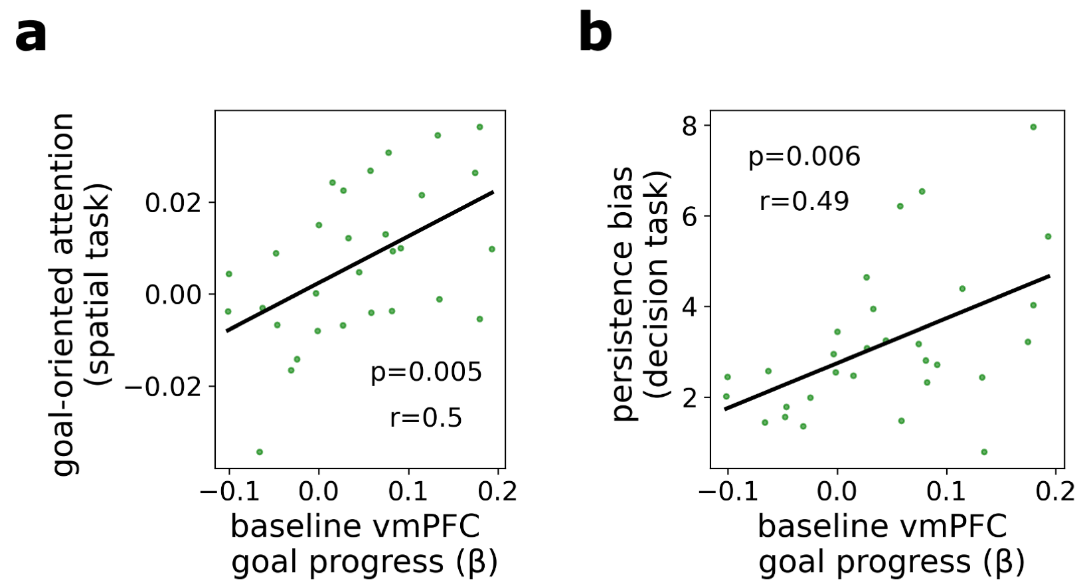
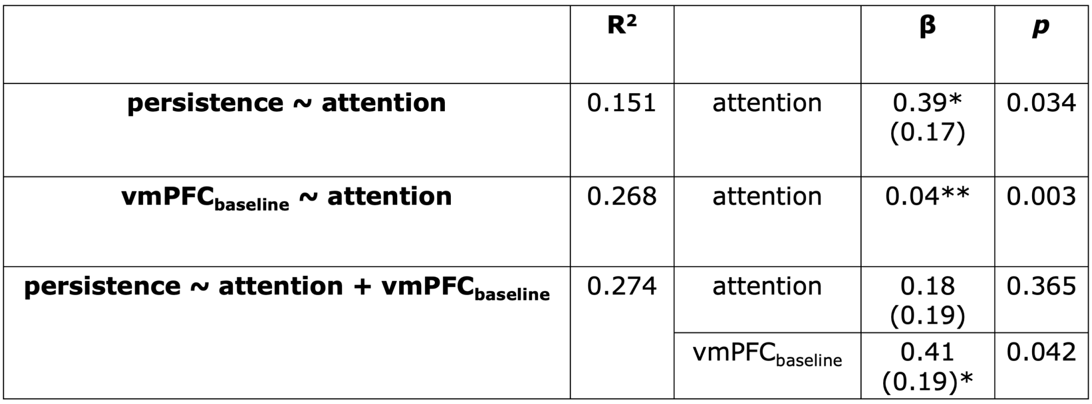


**Figure 4.6:** Correlation between baseline vmPFC activity and individual difference metrics of behaviour. **(a)** Baseline tracking of goal progress is plotted against goal-oriented attention bias, from the spatial attention task outside the scanner. Dots show data points for individual participants, line shows regression. **(b)** Baseline tracking of goal progress is plotted against persistence bias, which is the tendency to over-persist with the goal compared to the tree-search model. Dots show data points for individual participants, line shows regression.



| | **$R^2$** | | **β** | ***p*** |
|---|---|---|---|---|
| **persistence ~ attention** | 0.151 | attention | 0.39*<br>(0.17) | 0.034 |
| **vmPFC$_{\text{baseline}}$ ~ attention** | 0.268 | attention | 0.04** | 0.003 |
| **persistence ~ attention + vmPFC$_{\text{baseline}}$** | 0.274 | attention | 0.18<br>(0.19) | 0.365 |
| | | vmPFC$_{\text{baseline}}$ | 0.41<br>(0.19)* | 0.042 |

**Figure 4.7:** Mediation analysis table.

and persistence disappears when vmPFC activity is included in the same model, while vmPFC activity remains predictive of goal persistence (Table 4.7).

## 4.4 Discussion

Maintaining the balance between commitment and abandonment during goal pursuit may be supported by separate networks involving vmPFC and dACC respectively. We found that between decisions, a network of medial prefrontal areas centred on vmPFC continue to track progress with the current goal when this variable is no longer observable. We find that across individuals, baseline vmPFC activity (activity before the choice is presented) predicts both commitment to goals and goal-oriented attention in our task.

Our finding that vmPFC regions consistently track progress with the goal may relate to other studies finding sustained representations of contextual variables between decisions including in both rodent mPFC (Bari et al. 2019) and non-human primate vmPFC (Mehta et al. 2020). For example, one study found that neurons in rodent medial prefrontal cortex persistently represented long-term state variables in their firing rates including the values of individual actions (biasing choice) and the overall sum of action values (biasing response times) (Bari et al. 2019). These findings are consistent with theories that these areas carry relevant unobservable variables for guiding action, or in the language of reinforcement learning, the 'current state' (Wilson et al. 2014; Schuck et al. 2016). The manner in which contextual state variables sustained in persistent mPFC activity influences behaviour has been related to the dynamics through which recurrent activity in artificial neural networks biases subsequent processing (Blanco-Pozo et al. 2024).

Consistent with the idea that sustained vmPFC activity modulates behaviour, pre-stimulus activity in vmPFC predicted individual differences in attention and persistence in our task. This finding builds on growing evidence in both monkeys and humans demonstrating baseline vmPFC activity influences how options are processed and subsequently which choice is made (Lopez-Persem et al. 2016; Vinckier et al. 2018; Abitbol et al. 2015; Mehta et al. 2020). Baseline vmPFC activity has been argued to bias upcoming choices in line with prior contextual factors, including both stable preferences (such as tastes in music or food types; (Lopez-Persem et al.

2016)), and dynamic background states (such as satiety or mood; (Abitbol et al. 2015; Vinckier et al. 2018)). Our results provide evidence that another dynamic state, namely goal pursuit, modulates behaviour through this baseline activity. We argue that our results also offer a possible mechanism for these effects: sustained vmPFC activity represents the relevant current state, driving global changes in top-down attention and affecting how options are subsequently processed.

In various contexts, medial prefrontal cortex has been shown to support the selection of goal-relevant information at the time of decision, flexibly adapting to changes in the current goal (Grueschow et al. 2015; Rudorf and Hare 2014; Castegnetti et al. 2021; Trudel et al. 2021; Frömer et al. 2019), possibly through compression of goal-irrelevant information (Mack et al. 2020). Other studies have linked vmPFC activity to visual attention specifically, both responding to exogenous manipulations of attention (Lim et al. 2011; Hare et al. 2011), and in mediating the allocation of visual attention (Wolf et al. 2014). Preparatory activity in vmPFC has also been shown to predict memory-guided attention in cases where people anticipate upcoming events (Günseli and Aly 2020; Small et al. 2003). Here we present results bringing together these bodies of research, suggesting that the role vmPFC plays in selecting goal-relevant information is linked to visual attention.

Our results also reveal how neural value representations at the time of decision-making change dynamically across goal pursuit, consistent with attentional prioritization of the current goal. In **chapter 3**, we found that late in goal pursuit and compared to an optimal model, people demonstrated reduced sensitivity to the value of alternative goals compared to the value of the current goal. When the value of alternatives lost influence over behaviour, this was mirrored by a reduction in sensitivity to alternative value in the ventral striatum. While we are not aware of other studies showing this pattern, ventral striatum is known to respond to goal pursuit, for example through striatal dopamine ramps during goal approach (Howe et al. 2013; Hamid et al. 2016a).

Finally, we found both dACC and dlPFC positively co-varied with the value of abandonment, as well as being more active when participants choose to abandon their current goal. This is consistent with previous work showing that activity in these neural areas, and in dACC in particular, represents the value of alternative options (Fouragnan et al. 2019), and is more active when an individual disengages from the present behavioural policy (Kaiser et al. 2021; Stoll et al. 2016) or explores the environment (Trudel et al. 2021; Tervo et al. 2021). In fact, when people switch out of an exploitative state towards exploration, dACC activity predicts changes in task representation within vmPFC (Muller et al. 2019). Our findings are consistent with the idea that vmPFC maintains the current behavioural goal, while dACC underpins behavioural flexibility during goal pursuit. Consistent with this and in contrast to the striatal effects, we found relatively sustained representations of alternative option value throughout the goal in dACC.

## 4.5 Contributions

This chapter includes work from an upcoming publication in Nature Human Behaviour: *Holton, E., Grohn, J., Ward, H., Manohar, S.G., O'Reilly, J, & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention.*

The analyses in this chapter were supervised by Nils Kolling, Jill O'Reilly and Jan Grohn.

# 5

# Damage to vmPFC reduces goal persistence

*This chapter investigates whether vmPFC is causally involved in goal persistence in a population of patients with lesion damage. In* ***chapter 4*** *we found that vmPFC showed sustained goal-related activity which predicted individual differences in persistence with a goal among healthy individuals. Here we demonstrate this vmPFC region is responsible for increased commitment to a current goal: patients with damage to the same region of vmPFC have reduced goal persistence compared to lesion controls and age-matched controls. Importantly, we show this relationship between lower goal persistence and damage to this region of vmPFC is not explained by higher stochasticity. These patients choose to abandon their current goal at appropriate times, thereby performing better than patients with damage elsewhere who tend to over-persist with goals.*

## 5.1 Introduction

The aim of this chapter is to test whether vmPFC plays a necessary role in supporting goal commitment. In the previous chapter, we found that vmPFC tracked goal pursuit across trials, and predicted individual differences in how willing people were to abandon their goal as well as their degree of goal-oriented attention. However, vmPFC could be correlating with goal persistence without playing a direct causal role. In this chapter, we use the same paradigm in a group of twenty-three lesion patients and a population of age-matched healthy control participants to ask the critical question of whether vmPFC supports goal commitment. To answer this question, the chapter is divided into two halves, investigating commitment first and then addressing possible behavioural confounds.

In the first half of the chapter, we investigate whether neural damage to vmPFC is linked to reduced commitment to the current goal through two independent analyses. In the first analysis, we take a voxel-based approach to identify neural areas where damage predicts lower goal persistence, diagnosing a region in vmPFC. In the second analysis, we start with the pre-defined region of interest from our fMRI study and show that patients with damage here have lower persistence compared to lesion controls and age-matched controls. These two independent analyses converge on the same patients.

In the second half of the chapter, we perform control analyses to address the question of whether vmPFC damage could result in lower persistence because patients are responding more randomly. This is critical because previous studies have found patients with vmPFC damage show impairments at guiding decisions using value (Noonan et al. 2017; Fellows and Farah 2007), and in our task higher stochasticity will be associated with more abandonment of the goal. Crucially, we do not find that patients' use of value is noisier, as formalised through inverse temperature. Moreover, since people generally over-persist in this study (compared to an optimal model), we should find that if patients are truly less biased to persist rather than responding more randomly, they should show an advantage in performance. Indeed, we find that the vmPFC group perform better than other lesion patients, and no worse than healthy controls.

## 5.2 Methods

### 5.2.1 Participants

Twenty-six patients with brain lesions (mean age=58) and twenty-seven age-matched control participants (mean age=59) took part in the study. Of the lesion patients, one was excluded because they failed to pass the initial comprehension questions, and two were excluded because they were unable to complete the task. Of the remaining twenty-three individuals in the study, sixteen had damage within the frontal cortex and the remaining seven had damage to areas outside frontal cortex (see Fig.5.1a for overlap map of lesion damage). The patient population was recruited from a database of individuals who had previously visited the John Radcliffe Hospital and consented to be contacted for research studies. Ethical approval for the patient study was obtained by the London Fullham Research Ethics Committee (IRAS project number: 242551; REC Reference number: 18/LO/2152). All participants gave written informed consent before the experiment. Participants were paid £15/hour plus a performance-dependent bonus between £8-12.

Data collection took place virtually, over a single session where the participant completed an online version of the task (hosted on Pavlovia), while the researcher remained on the telephone throughout the session. The patient group received the same training and were required to pass the same comprehension questions before proceeding to the main task as described in **chapter 3**. The main task consisted of 250 trials total. Rather than use variable schedules, the same schedule was used across all participants, in order to maximise the control of conditions for comparison of lesion effects. The age-matched controls completed the same schedule and training procedure online, and were recruited through Prolific.co.

For a description of the task itself, see Experimental Paradigm in **chapter 3**.

### 5.2.2 Sanity checks

Before investigating the effects of neural damage in this task, we began by performing several sanity checks. First, as for the healthy fMRI cohort, we performed a simple regression analysis to determine whether participants were sensitive to the critical elements of the task: the offers for the current and best alternative, in addition to the contents of the net. To do this, we predicted abandonment decisions in a logistic regression using the contents of the net and the three offers (current goal, best alternative offer and worst alternative offer) as regressors.

Next, we confirmed that the tree-search model was also the best model of behaviours in both groups (lesion patients and age-matched controls). As described in **chapter 3** methods, we used a leave-one-out cross validation analysis to verify the best fitting behavioural model.

### 5.2.3 Analyses of persistence

#### Analysis 1: Voxel-wise lesion behaviour mapping

We began by investigating the relationship between brain damage and persistence biases independently from the priors that resulted from our fMRI study. To investigate areas causally relevant for persistence in the task, we performed a voxel-wise analysis predicting behaviour from maps of the patients' neural damage (Bates et al. 2003; Karnath et al. 2018). This process involves identifying voxels at which damage predicts a difference in the behaviour of interest compared to individuals with lesions sparing this voxel. Analagous to fMRI analyses, this creates a map of t-statistics (where damage predicts the behaviour of interest), which are thresholded and cluster corrected to control for multiple comparisons.

We focussed on goal commitment, quantified as an individual's persistence bias from the tree-search model, as our behaviour of interest. We used a threshold of $t > 2.3$ where damage predicted lower persistence biases (corresponding to $p < 0.01$, one-sided test because we were interested in where damage will reduce persistence biases).

We controlled for multiple comparisons by performing cluster correction using the False Discovery Rate method (FDR; Genovese et al. (2002); Karnath et al. (2018)). Using a permutation-based approach, we quantified the maximum cluster size we would expect from our lesion data-set due to chance, at the same significance threshold. On each permutation (total 1000 iterations), we shuffled individual persistence biases and performed the same voxel-wise regression analysis with the shuffled biases. We created a distribution of clusters found across all permutations, and defined the minimum cluster size for significance at the 95% cut-off of all clusters found by chance, resulting in a minimum cluster size cut-off of 255 voxels for our empirical analysis.

### Analysis 2: ROI-based analysis

Next, we performed a group-wise comparison where we split lesion patients based on whether they were damaged in the region pre-defined by our fMRI study. Our fMRI study had identified a subset of areas carrying signals relating to goal-pursuit even between decisions, focussed on vmPFC. We split all patients into two groups on the basis of whether they were damaged at an ROI centred on the peak of this inter-decision fMRI activity. Following the same procedure described in **chapter 4**, we extracted the region of interest with a 3 voxel radius ($7.2mm^3$) centred on the peak of activity tracking goal progress during the inter-trial interval in our fMRI study. We then tested for a difference in persistence biases between the two groups of patients, and against the age-matched controls. We used a one-sided permutation test to test for difference in means between groups, due to the small sample sizes and non-normally distributed biases (n.b. we used a one-sided test based on our hypothesis that damage to vmPFC would reduce persistence; although note the difference remains significant if we were to perform a two-sided test).

Our previous voxel-wise regression analysis identified a region of vmPFC which included damaged voxels from five unique patients. Our ROI-based lesion analysis independently identified four out of the five same patients when selecting on the basis of the pre-defined fMRI region. For the analyses which follow, we compared the four patients identified using the fMRI ROI (‘lesion in vmPFC ROI’) to other patients (‘lesion controls’) and age-matched healthy controls (‘age-matched controls’). Note that if we were also to include the fifth patient in the vmPFC sample (identified from the voxel-wise regression analysis), all the same conclusions hold.

We chose to perform further analysis on the four patients identified from the fMRI ROI to avoid circularity (because the voxelwise analysis explicitly searched for areas predicting reduced persistence). When relevant, the analysis with all five patients is included in the appendices.

### 5.2.4 Performance analysis

If damage to the identified region of vmPFC causes participants to be less biased towards over-persisting (rather than more stochastic and thus more likely to switch goals), we would counter-intuitively expect these patients to perform better than other groups. We quantified performance as the mean number of trials taken to complete a goal, where a lower value means goals were completed faster. Since all participants in the patient task completed the identical schedule, this measure is not vulnerable to schedule-specific artefacts. We then tested whether the vmPFC cluster patients performed better than patients with damage elsewhere, and better than lesion controls, using a one-sided parametric test (we used a one-sided test based on our hypothesis that reduced bias should improve performance but note the difference remains significant if we were to perform a two-sided test).

### 5.2.5 Control analyses

We performed several subsequent control analyses to verify that the four patients with damage in the vmPFC ROI were truly less biased to persist, rather than persisting less for other reasons (such as using a drastically different strategy, or responding more randomly).

First, we verified that the vmPFC lesion group were responding sensibly by quantifying the number of trials in which the worst alternative was selected (a clear irrational choice). We used two-tailed permutation tests to verify there was no difference between the vmPFC lesion group and other patients, and between the vmPFC lesion group and age-matched controls.

Second, we confirmed that behaviour among patients with damage to this region was still best explained by the same behavioural model as healthy individuals (the 'tree-search model'), and not by a more simple strategy, by fitting the four behavioural models described in **chapter 3** (see behavioural models).

Finally, we verified that the vmPFC patients were not more stochastic in their decision process. We quantified stochasticity as inverse temperature, which is the beta weight associated with the tree-search value in our logistic regression predicting abandonment. We used two-tailed permutation tests to verify there was no difference in stochasticity between the vmPFC lesion group and other patients, and between the vmPFC lesion group and age-matched controls.

### 5.2.6 Spatial task in lesion patients

Our patient group also performed the interleaved spatial task. We quantified spatial attention bias as the accuracy advantage for the current goal item over the alternative item, as described in **chapter 3** (spatial attention task analyses). We predicted the vmPFC group would show a lower accuracy advantage for the goal item over the alternative items in the interleaved task, since attention would be captured less by the goal.

## 5.3 Results

### 5.3.1 Lesion patients perform sensibly on the task

Both the lesion patients and the age-matched controls performed sensibly on the task (see Fig.5.1b). They were more likely to switch given high alternative offers (impact of best alternative offer: lesion patients, $\beta = 0.44$, $t(22) = 5.74$, $p < 0.001$; age-matched controls, $\beta = 0.34$, $t(26) = 8.77$, $p < 0.001$), less likely to switch when the current goal offer was high (impact of current goal offer: lesion patients, $\beta = -0.35$, $t(22) = -5.77$, $p < 0.001$; age-matched controls, $\beta = -0.35$, $t(26) = -8.65$, $p < 0.001$), and less likely to switch after accumulating more goods (impact of net contents: lesion patients, $\beta = -0.67$, $t(22) = -13.41$, $p < 0.001$; age-matched controls, $\beta = -0.64$, $t(26) = -29.19$, $p < 0.001$). For neither group did the worst alternative offer have a significant impact on goal abandonment choices (impact of

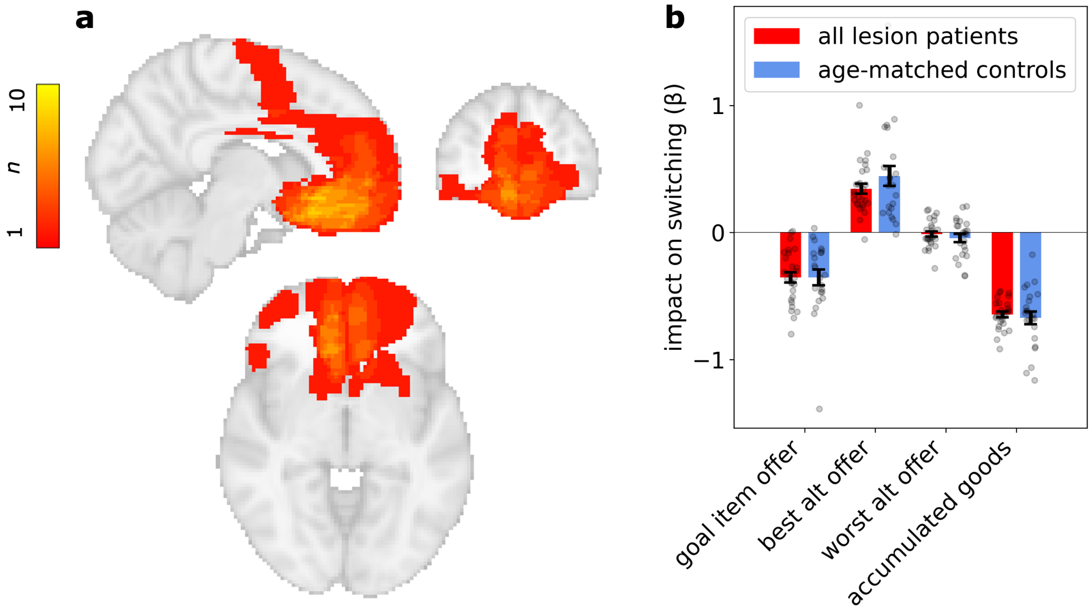


**Figure 5.1:** Lesion maps and simple behaviour. **(a)** Lesion overlap maps showing the location of lesions across the entire cohort (n=23 total) **(b)** Simple regression showing both lesion patients and age-matched controls were sensitive to the main elements of the task (offers for the current and alernative goal, as well as accumulated goods). Dots show beta weights for individual participants, error bars depict s.e.m.

second-best alternative offer: lesion patients, $\beta = -0.04$, $t(22) = -1.37$, $p = 0.183$; age-matched controls, $\beta = -0.01$, $t(26) = -0.61$, $p = 0.550$).

In addition, for both patients and controls, the tree-search model provided the best description of overall behaviour within each group (See appendices; Fig.C.1).

For the next set of analyses, we asked how the location of lesion damage affected persistence bias, defined as the tendency to persist with the chosen goal beyond the point at which the tree-search model would switch, as the key behavioural marker of goal commitment.

### 5.3.2 Damage to vmPFC reduces persistence biases

We began by investigating whether damage to particular areas reduced persistence in the lesion patient group, independent from any priors from our fMRI study. Using a voxel-wise regression analysis, we asked at what locations damage predicted a reduction in persistence bias. Independently corroborating the findings of our fMRI study, the only region where damage predicted a reduction in persistence bias was in vmPFC (Fig.5.2a; green cluster). The vmPFC cluster survived whole-brain cluster correction as an area where damage leads to reduced persistence biases (cluster threshold $t > 2.3$; $p < 0.01$, one-sided, cluster size = 269 voxels, threshold cluster correction size = 255 voxels, cluster peak = $[0, 42, -14]$, t-statistic at cluster peak = 2.74). The vmPFC region identified in the whole-brain voxel-wise analysis contained damaged voxels from 5 patients (out of a total of 23 lesion patients who took part in the study).

### 5.3.3 Convergence between lesion study and fMRI study

We then asked how much the region identified in our lesion patient study aligned with the findings of our fMRI study (Fig.5.2b). Our fMRI study had identified a subset of areas carrying signals relating to goal-pursuit even between decisions, focussed on vmPFC. We split all patients into two groups on the basis of whether they were damaged within a region of interest at the peak of this fMRI activity, found in vmPFC (ROI centred on the peak of the activity tracking goal progress during the inter-trial interval in our fMRI study; shown in Fig.5.3, left). There were four lesion patients with damage to this ROI.

Patients with damage to this region were less persistent than both patients with lesions elsewhere and age-matched healthy controls (persistence biases among patients damaged within fMRI ROI: $n = 4$, mean = 2.33, $std = 2.31$; persistence bias among other patients: $n = 19$, mean = 6.12, $std = 2.88$; persistence bias among age-matched controls: $n = 27$, mean = 5.29, $std = 2.74$; difference between vmPFC group and other patients: permutation test, difference in means = 3.79, $p = 0.012$, one-sided; difference between vmPFC patients and age-matched controls: permutation test, difference in means = 2.97, $p = 0.023$, one-sided).

We found these four patients who had damage within the region pre-defined by our fMRI study corresponded to four (out of the five total) patients identified from our independent voxel-wise patient analysis. Therefore our fMRI study and

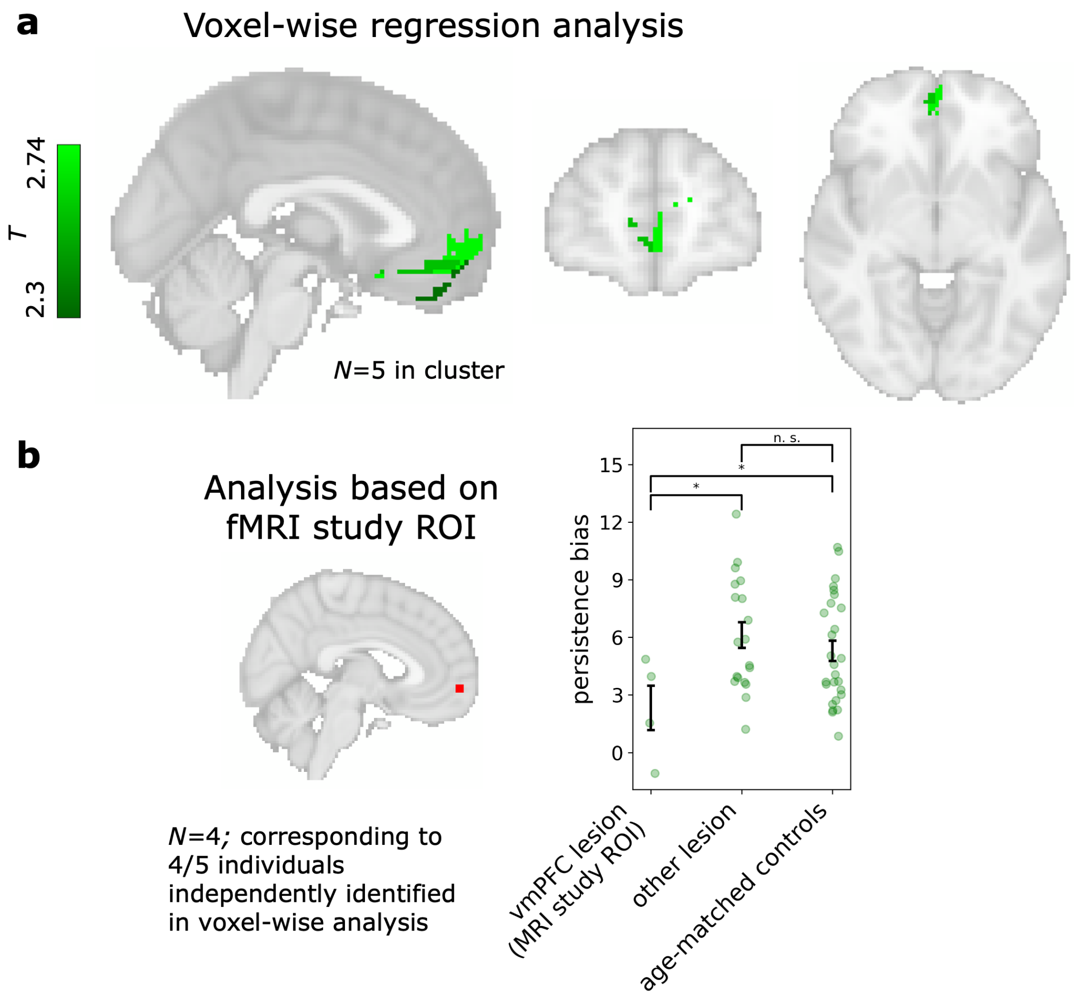


**Figure 5.2:** VmPFC damage reduces goal persistence. **(a)** Results from the whole-brain voxelwise analysis. Green shows areas where lesion damage predicts lower persistence biases. Above-threshold t-statistics (t > 2.3 before cluster correction) are displayed for illustrative purposes. We controlled for multiple comparisons by performing cluster correction using permuation based false discovery rate analyses. The vmPFC cluster survived whole-brain cluster correction. **(b)** Patients with damage to the vmPFC region identified in the fMRI study show reduced persistence bias. Patients were split into two groups depending on whether they were damaged within the ROI centred on the peak of BOLD activity tracking goal progress between decisions in healthy participants (shown with red dot, left image). This area was damaged in 4 patients, corresponding to 4 out of the 5 patients independently identified in the voxelwise analysis in (a). Patients with damage to this region showed lower goal commitment than patients with lesions elsewhere and age-matched controls (right plot). Error bars show s.e.m. in each group; green dots depict individual biases.

lesion patient study independently converge to identify the same vmPFC region as being relevant for goal commitment.

### 5.3.4 Reduced persistence following vmPFC lesions is not explained by higher stochasticity

Next, we ruled out the possibility that the vmPFC damaged group were simply performing worse in some general way, for example by making random choices or forgetting the goal. An important point to note is that, because participants in general over-persist, a reduction in persistence biases should lead to better task performance, if participants switch goals at points at which it is beneficial to do so (rather than making random switches due to, for example, task disengagement). This is exactly what we find: the four patients with damage to the vmPFC ROI in fact perform significantly better than patients with damage elsewhere, and no worse than age-matched healthy controls (Fig.5.3; performance is quantified as mean trials to fill a net, i.e. smaller values indicate goals are completed faster).

Finally, we used further post-hoc analyses to verify that a) vmPFC patients were not selecting the worst alternative (indicative of irrational switching), b) vmPFC patients were not using a different normative model to solve the task and c) vmPFC patients were not responding more stochastically as measured by inverse temperature.

We found the vmPFC group did not switch to the worst alternative option any more than the other groups, suggesting they were not switching goals randomly (Fig.5.4a; Number of switches to worst alternative; vmPFC group, mean $= 1.75$, $std = 1.1$; other patients, mean $= 1.3$, $std = 1.7$; age-matched controls, mean $=$ $1.1$, $std = 1.3$; vmPFC vs. other patients: two-sided permutation test, difference in means $= 0.49$, $p = 0.592$, $n.s.$; vmPFC vs. age-matched controls: two-sided permutation test, difference in means $= 0.64$, $p = 0.464$, $n.s.$). Consistent with the idea they were switching goals sensibly, the vmPFC group switched to choosing the best alternative option more than either of the other groups (Fig.5.4b; Number of switches to best alternative; vmPFC group, mean $= 37.3$, $std = 14.8$; other

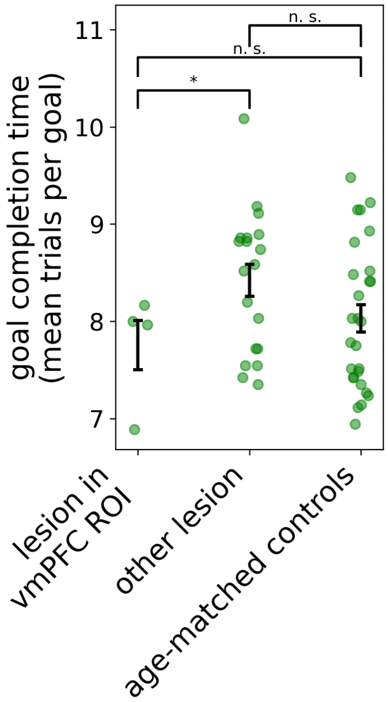


**Figure 5.3:** VmPFC lesion group perform better than other lesion patients, and no worse than age-matched healthy controls. Performance was measured as the average number of trials to complete a goal, where lower scores correspond to faster goal completion.

patients, mean = 19.3, *std* = 7.4; age-matched controls, mean = 19.9, *std* = 8.1; vmPFC vs. other patients: two-sided permutation test, difference in means = 17.99, $p = 0.010$; vmPFC vs. age-matched controls: two-sided permutation test, difference in means = 17.36, $p = 0.010$).

We did not find any evidence that vmPFC patients were using a simpler response strategy, as choices were again best described by the tree-search model (Fig.5.4c). We formally quantified stochasticity as inverse temperature when predicting choices using the tree-search model, and found the vmPFC group showed no difference compared to other patients or age-matched controls (Fig.5.4d; Inverse temperature; vmPFC group, mean = 0.57, *std* = 0.04; other patients, mean = 0.51, *std* = 0.22; age-matched controls, mean = 0.61, *std* = 0.19; difference between vmPFC group and other patients: two-sided permutation test, difference in means = 0.06, $p = 0.572$, *n.s.*; difference between vmPFC patients and age-matched controls: two-sided permutation test, difference in means = 0.04, $p = 0.633$, *n.s.*). Note the inverse temperature parameter has good parameter recoverability (see appendices; Fig.A.2).

Taken together, these results suggest that patients with damage to this region of vmPFC are not simply using a different task strategy or responding more randomly, but instead are less *biased* towards over-persisting with a goal.

### 5.3.5 Attention results in lesion patients are inconclusive

While as predicted, the vmPFC group did not show a significant accuracy advantage for stimuli related to the current goal (mean goal item accuracy - mean alternative items accuracy: $mean = 0.019$, $std = 0.030$, Wilcoxon: $n = 4$, $T = 2.0$, $p = 0.375$), we cannot interpret this result since we also did not detect goal-oriented spatial attention effects generally among patients with lesions elsewhere (mean goal item accuracy - mean alternative items accuracy: $mean = 0.032$, $std = 0.127$, Wilcoxon: $n = 18$, $T = 80.0$, $p = 0.568$). Accuracy in the spatial attention task for each group is shown in Fig.5.5, and reaction times for the spatial task are shown in the appendices (Fig.C.4). Predicted by the fact that neither patient group showed goal-directed attention, there was also no difference in attentional biases between groups (permutation test for difference in goal item accuracy advantage between groups: mean difference=0.013, $p = 0.524$, *n.s.*). Given we could not detect goal-oriented attentional effects in the lesion patient population generally, we could not determine whether lesion location affects goal-oriented attention in this study. As illustrated in Fig.5.5, accuracy was generally much lower in this population than for the fMRI cohort. Given the lack of goal-oriented attention effects in this older cohort with lesion damage, we did not perform any further analyses on the attention data.

## 5.4 Discussion

Across the last two chapters, we present multiple converging lines of evidence demonstrating that vmPFC plays a key role in supporting commitment to a current

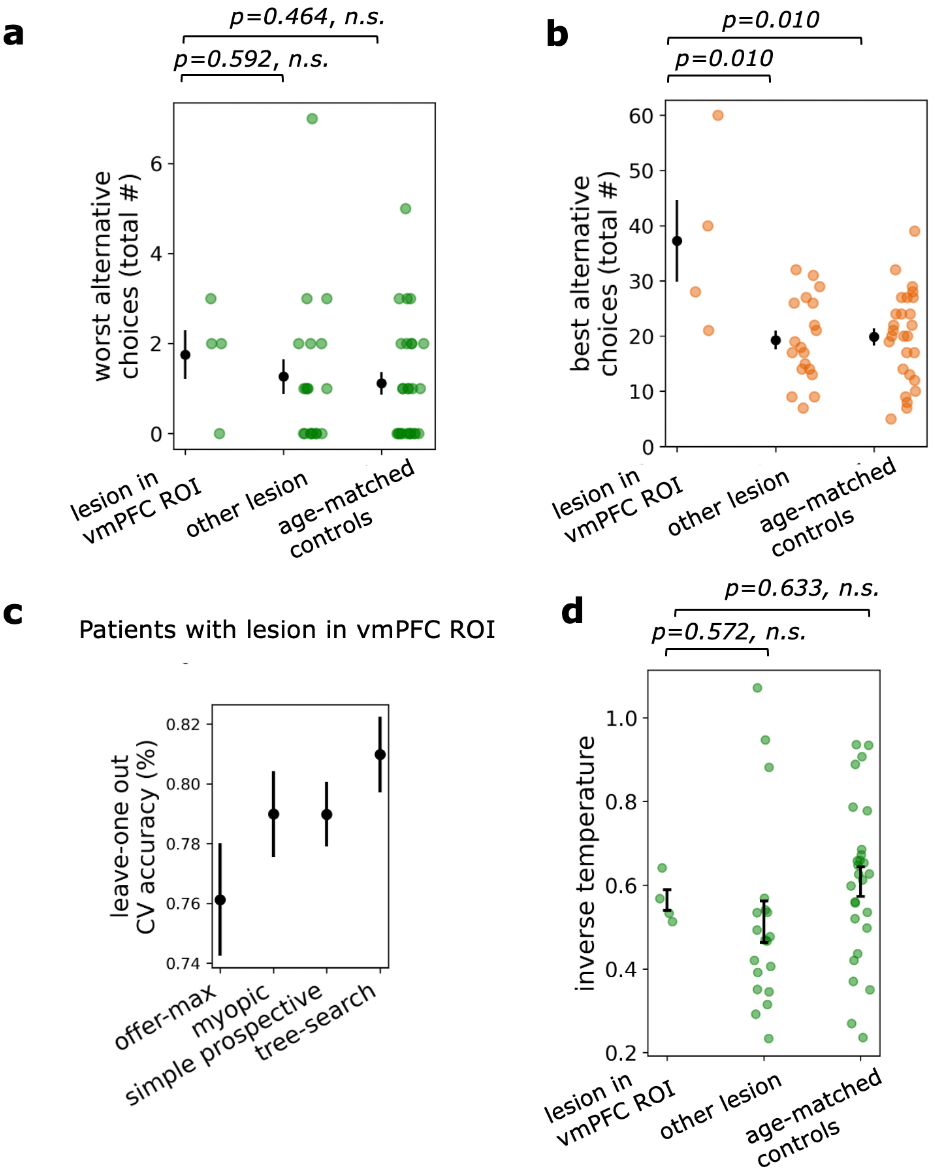


**Figure 5.4:** Post-hoc analyses on groups. The vmPFC group includes the four patients who had lesion damage in the pre-selected ROI from the fMRI study. **(a)** Number of times participants switched to the worst alternative offer (out of 250 trials total). The vmPFC group did not perform this kind of switch any more than lesion controls or age-matched healthy controls, suggesting they were not making random choices. Error bars show s.e.m. across groups, dots show number of switches per individual. **(b)** Number of times participants switched to the best alternative offer (out of 250 trials total). The vmPFC group were more likely to switch to the best alternative offer than either the lesion controls or age-matched healthy controls, supporting the idea they were making sensible switch choices. Error bars show s.e.m. across groups, dots show number of switches per individual. **(c)** Analyses of model strategy used by the vmPFC lesion group. As for the healthy fMRI participants and other groups, the tree-search model is the best description of behaviour. Error bars show s.e.m. of cross-validation accuracy. **(d)** Post-hoc analysis comparing inverse temperature across the three groups, where abandonment choices are predicted using tree-search value. This suggests patients with damage within the vmPFC ROI are not simply more stochastic since they show no difference in inverse temperature. Error bars show s.e.m., dots show individual data points.

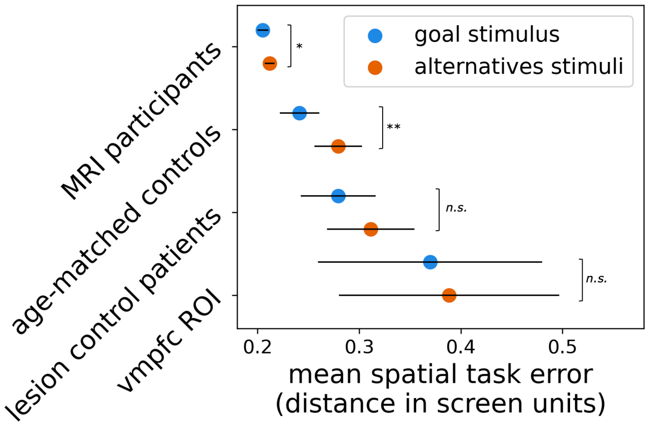


**Figure 5.5:** Spatial attention task in lesion patients. Neither the control lesion group, nor the vmPFC lesion group show a significant accuracy advantage for the current goal item. This contrasted with the MRI participants and healthy age-matched controls (stars depict significance for wilcoxon tests comparing goal stimuli accuracy against best alternative stimuli accuracy. MRI patients: p=0.035; EC patients: p=0.006; lesion patients: p=0.568; vmPFC patients: p=0.375), but limits our interpretation of the impact of lesion location on goal-oriented attention.

goal. In the last chapter, our fMRI study found that the vmPFC carried sustained goal-related information between decisions in our task, and baseline activity before the decision predicted the two independent behavioural metrics of goal capture: both an individual's bias to persist with the current goal and their bias to prioritize goal-related stimuli in attention. In this chapter, we demonstrated that vmPFC plays a causal role in the process of goal commitment: patients with damage to the same region are less biased to continue persisting with the current goal. Through a voxelwise analysis of damage in our patient sample, we identified a vmPFC cluster in which damage predicted reduced persistence biases. The area identified in patients closely corresponded to the area involved in persistence among healthy individuals, discovered in our fMRI study.

Our results expand on previous reports that lesions to this area in both humans and primates interfere with the ability to prioritize relevant decision variables, for example, in cases when a distracting alternative is introduced (Noonan et al. 2010; Noonan et al. 2017). However in our task, the capacity to maintain the relevance of alternative goals actually provides an advantage, since healthy people tend to ignore good alternatives in favour of persisting. While previous lesion studies have found this patient population to behave more stochastically (Noonan et al. 2017; Camille et al. 2011), notably lower persistence biases among vmPFC lesion patients in our task cannot be explained by an increase in stochasticity. In fact, we find that patients with vmPFC damage performed better than other lesion patients and no worse than age-matched controls. In a goal-pursuit context, healthy individuals may have a tendency to over-constrain the decision space by focusing only on the current goal and ignoring alternatives. In contrast, a lesion to this area of vmPFC may reduce capture by the current goal, allowing good alternative options to maintain their relevance throughout goal pursuit. We note that, while this is beneficial in our task, it is likely to be advantageous to constrain the task space in ecological goal-pursuit settings. As discussed in **chapter 1**, biases towards goal commitment could result from efficient neural resource allocation (attending to pursuing the chosen goal) and help to effectively structure behaviour over time.

These findings are relevant for understanding why vmPFC damage can lead to difficulties with real world goal-directed behaviours, for example in classic tests like the multiple errands task (Tranel et al. 2007). As discussed in **chapter 2**, a key part of planning and pursuing goals in the real world consists of selecting the relevant information in complex environments with many distractions. The ability to sustain the intended goal by constraining options to the relevant variables will be critical for pursuing goals in the real world.

### 5.4.1 Summary

The findings from the last three chapters suggest that goal pursuit leads to global changes in how the environment is processed, prioritizing the current goal in attention and decisions. In **chapter 3**, we found that goal commitment was related to pervasive changes in goal-oriented attention related to commitment. In **chapter 4** we found goal-directed biases in both decision-making and attention were related to pre-decision vmPFC activity. In **chapter 5** we demonstrated that vmPFC was causally responsible in supporting commitment to a goal, as patients with damage to this region are less persistent. We argue that goal-directed selective attention, supported by vmPFC, could provide a mechanism for sustaining goal pursuit. While goal commitment may manifest in seemingly irrational tendencies to persist with a selected goal, the ability to filter information to prioritize the chosen task would be critical in ecological settings.

## 5.5 Contributions

This chapter includes work from an upcoming publication in Nature Human Behaviour: *Holton, E., Grohn, J., Ward, H., Manohar, S.G., O'Reilly, J, & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention.*

Nils Kolling, Jill O'Reilly, Jan Grohn and Sanjay Manohar all contributed to the supervision of analyses in this chapter. Harry Ward contributed to data collection.

# 6

# Effects of mPFC lesions on the components of planning

*In this chapter, we turn to a different element of pursuing goals: planning future action. We investigate how medial PFC damage affects the components of planning in two pre-existing paradigms, namely the two-step task and four-in-a-row. While the two-step task involves planning over a small horizon of binary choices, the four-in-a-row task presents subjects with long horizons of upcoming choices in a complex state space. This presents a more naturalistic planning setting requiring multiple cognitive processes for effective decision making. Damage to mPFC did not affect behaviour in the two-step task, although we found that model-based planning was generally attenuated across all participants in our older cohort. However, medial PFC damage did selectively impair performance in the four-in-a-row paradigm. On closer inspection, we found this was explained by an increased tendency to overlook valuable features. This was contrasted with relatively preserved capacity to simulate future states, or recognise good heuristics for choice. We link this to a general theory of mPFC involvement in selecting task-relevant information and guiding goal-oriented attention.*

## 6.1 Introduction

A central idea in previous chapters has been that vmPFC supports the selection and representation of relevant information for guiding goal-directed decisions. In **chapter 2**, we argued this capacity will be particularly critical in complex environments where the sheer number of options and potential for interference is high. This includes scenarios where the space of current options is large, or where evaluating options requires planning into the future over an exponentially growing decision tree.

In this chapter, we refer to our population of patients identified with an anatomical vmPFC mask as 'mPFC' rather than 'vmPFC' for two reasons: (1) although a vmPFC mask was used to classify lesion patients, there is widespread damage extending into dorsal areas as shown in Fig.6.1, and (2) it is important to clarify that the lesion group is not selected using the same procedure as the previous chapter, because we did not have a prior hypothesis about the location of relevant vmPFC activity (compared to the lesion study in **chapter 4**, where we could directly test our hypothesis from the fMRI study in **chapter 3**).

In this chapter, we test the impact of mPFC lesions on the components of planning, comparing behaviour to lesion patients with damage elsewhere in the brain, and age-matched healthy controls. To investigate planning, we use two pre-existing paradigms from the field of computational cognitive science: the two-step task (Daw et al. 2011) and four-in-a-row (van Opheusden et al. 2023). Four-in-a-row requires subjects to plan in a substantially larger state space than the two-step task. We find that mPFC damage selectively affects planning in four-in-a-row, but does not affect behaviour in the two-step task. Notably, however, our older cohort generally show low markers of model-based planning in the two-step task across all groups.

The planning deficits that mPFC patients show in four-in-a-row are related to tendencies to overlook relevant valuable features in complex state spaces. This contrasts to relatively preserved capacity for other components of planning, including simulating the future or or recognise good heuristics for choice. An impaired ability to incorporate all the goal-relevant information could contribute to many of the real-world examples of planning deficits in mPFC patients.

### 6.1.1 Medial prefrontal lesions and planning

Across a range of settings, mPFC damage has been associated with alterations in planning or sequential decision-making. A traditional example of this is the Multiple Errands task, where mPFC patients struggle to structure their own behaviour across a sequence of real-world errands in a shopping mall (Shallice and Burgess 1991; Tranel et al. 2007). Patients with damage to mPFC also perform worse in classic neuropsychological "planning" tasks such as the Tower of London (Owen et al. 1990). However, real-world planning is a composite ability, requiring multiple cognitive processes. While many studies point to general disturbances in planning, the behaviours are consistent with a number of root cognitive deficits including model-based simulation of the future, applying heuristics to guide choice, or selecting and integrating all the right elements of the problem space. It remains an unsolved research question which of these components leads to planning deficits following mPFC damage.

A body of work has focussed on specific mPFC impairments in thinking about the future, for example in imagining events at distant future time-scales (Fellows and Farah 2005) or producing detail about the future (Bertossi et al. 2016a; Bertossi et al. 2016b; Bertossi et al. 2017). Some studies have also suggested that temporal discounting is disturbed in mPFC patients, finding patients prefer smaller more immediate rewards over larger future ones (Peters and D'Esposito 2016; Sellitto et al. 2010; although see Fellows and Farah 2005). These findings might suggest that mPFC lesion patients have difficulties simulating possible futures.

However an intriguing alternative account is that planning deficits in mPFC patients are not caused by impairments in simulating the future per say, but by impairments in selecting what to simulate or what information is relevant.

There are examples from other domains where down-stream deficits prove to be a consequence of up-stream disturbances in how information is selected, such as in goal-directed attention. For example, social theories of mPFC damage were originally supported by evidence showing patients struggle with facial emotional recognition (Hornak et al. 1996; Heberlein et al. 2008). However, this was later shown to be mediated by alterations in orienting visual attention towards the eye region of faces (Wolf et al. 2014), and the deficit in emotional recognition can be ameliorated when attention is explicitly oriented to these features (Wolf et al. 2016; although see Vaidya and Fellows 2016).

Disturbances in selecting goal-relevant information may also explain some of the inconsistent findings regarding the effects of mPFC lesions on value-based decision-making. Specifically, mPFC patients show interference from irrelevant value in multi-option scenarios (Noonan et al. 2017), contrasting to relatively intact decisions between binary choices, when decisions do not depend on the selection of relevant task variables (Vaidya and Fellows 2015a). In addition, when decisions depend on integrating multiple dimensions rather than using a single dimension, mPFC patients show more pronounced impairments (Vaidya et al. 2018; Pelletier and Fellows 2019). Could deficits in complex planning tasks be explained by alterations in guiding the decision-process to the relevant information?

### 6.1.2 Cognitive planning tasks

Planning depends on mental simulation of the future using an internal model of the environment (Hunt et al. 2021; van Opheusden et al. 2023). A paradigmatic laboratory planning task is the ‘two-step task’, designed to disentangle habitual ‘model-free’ learning from ‘model-based’ learning which utilises knowledge about the structure of the environment (Daw et al. 2011). In this paradigm, receiving reward depends on making a series of two sequential choices. The ability to plan across the two choices using structural knowledge of the task can be dissociated from simple repetition of reinforced (rewarded) action (although see Akam et al.

(2015) and Russek et al. (2017) for examples of model-free algorithms resembling model-based behaviour).

The two-step task allows for simulation over only two steps ahead at the most, and individuals are presented with only binary choices at each step. Full exploration of the decision tree is therefore a realistic possibility. In naturalistic environments, planning is far more complex. Evaluating different actions can require simulation over an indeterminate future horizon, and with limitless option possibilities. Given the size of this state-space, real-life planning is likely to depend on multiple different abilities, including selecting appropriate information as well as using approximate methods of calculating value and limiting tree-search (Pearl 1988; Sutton and Barto 1998; Koller and Friedman 2009; Russell and Norvig 2016; Gershman et al. 2015). Studying these complex behaviours in the laboratory while retaining computational tractability has been a challenge.

A recently developed planning paradigm known as 'Four-in-a-row' involves a much larger state space than previous laboratory paradigms (1.2 x $10^{16}$ non-terminal states), making planning over all possible scenarios unfeasible for players (van Opheusden et al. 2023; Ma et al. 2022). This task is a two-player game (played against a computer opponent), which is an extension of the traditional game 'tic-tac-toe'. The aim is to place four consecutive pieces of your own color on a grid before the opponent does. To play effectively, subjects must rely on heuristics to evaluate states quickly, simulate future moves, as well as guide attention to the relevant information.

### 6.1.3 Aims

A resounding message from many decades of frontal lesion research is the need for planning tasks which deliver the necessary complexity to reveal subtle behavioural alterations, but also the interpretability to diagnose specific cognitive functions. More complex paradigms supported by rich computational models of behaviour provide new opportunities to strike this balance between complexity and interpretability (van Opheusden et al. 2023). Our first aim was to determine whether planning deficits in mPFC patients differed based on the complexity of the planning task by investigating performance in the two-step task and four-in-a-row. Both tasks were run in mPFC lesion patients, control lesion patients, and an age-matched healthy control sample.

Given that mPFC damage impaired performance in four-in-a-row (although not in the two-step task), our second aim was to determine which components of planning contributed to the deficit. The computational model developed for the four-in-a-row task allows planning to be broken down into three distinct components. The first component is the capacity to identify good heuristics for evaluating states. For example, knowing that a bus stop is a valuable feature if your aim is to travel across town is based on a heuristic (you can recognise its worth without knowing its exact value in your journey). The second component is the number of steps planned into the future (the depth of forward simulation). This captures model-based simulation of future possibilities. The final component is the tendency to overlook valuable features on any trial ('attentional lapse'). Rather than reflecting systematic deviations in knowing what is 'valuable', this third component reflects tendencies to miss important information on any trial. We find that mPFC impairments in planning are specifically linked to this final component: the tendency to overlook valuable features.

## 6.2 Experiment 1: Two-Step Task

### 6.2.1 Participants

For the two-step study, I analysed an unprocessed data-set previously collected by members of the Manohar Lab from 2015-2018, at the John Radcliffe Hospital in Oxford. The data-set consisted of a total of fifty lesion patients (mean age=58) and twenty age-matched healthy control participants (mean age=68). Data from one patient was not included because they failed to complete the task, leaving a total of forty-nine lesion patients. Brain lesions were registered manually by a neurologist prior to study recruitment. Patients were a priori assigned to the medial prefrontal

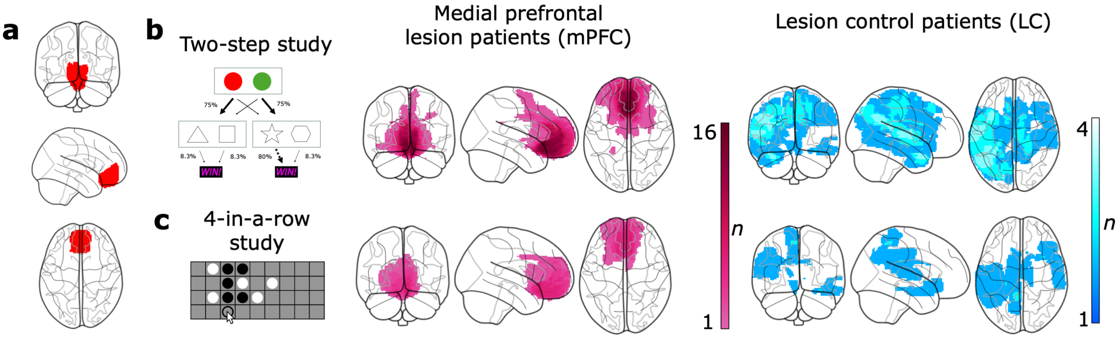


**Figure 6.1:** Lesion maps and simple behaviour. **(a)** Anatomical mask of medial prefrontal cortex from the Harvard-Oxford cortical atlas deployed in FSL (Jenkinson et al. 2012). Patients were categorised a priori into mPFC and LC groups on the basis of whether they had neural damage inside the mask. **(b)** Overlap of brain lesion masks for the mPFC patients (left, pink) and lesion control patients (right, blue) who participated in the two-step study. Colour bars shows number of patients with damage in each voxel. **(c)** Same as (b) for the Four-in-a-row study analysed later in this chapter.

cortex (mPFC) group and lesion control (LC) group. The Harvard-Oxford cortical structural atlas from FSL was used for mPFC lesion classification, by identifying patients who had neural damage within the binarized mask of ventromedial frontal cortex (map shown in Fig.6.1a; threshold $> 0$; Jenkinson et al. (2012)). All individuals with damaged voxels within the mask were assigned to the mPFC group (mPFC lesion group: $N = 30$; Fig.6.1b, left), while those where mPFC was spared were assigned to the comparison lesion group (Lesion control group: $N = 19$, Fig.6.1b, right).

### 6.2.2 Experimental Methods

All participants completed a variant of the two-step task (Daw et al. 2011), designed to measure habitual versus goal-directed decision-making. The task involved making repeated two-stage decisions in order to earn rewards (Fig.6.2a). On each trial, participants first chose between two colours and then between two shapes. Of crucial significance, each colour in step-one led to a specific set of shapes in step-two with a 75% probability (“common transition”), but led to the opposite pair of shapes in 25% of trials (“rare transition”). Of the four possible shapes which could be offered in step-two, only one shape had a high probability of reward at any point in time. This required participants to think strategically about which choice in step-one was most likely to lead them to the set of offers in the second step inclusive of the high reward option.

The task is designed to measure the extent to which decision-makers are using “model-free” or “model-based” estimates of value. A learner who uses a “model-free” strategy will be more likely to repeat their first step choice after being rewarded at the end of the trial, regardless of whether the transition between steps was common or rare. However, a decision-maker who uses a model of the task structure will be more likely to repeat their first step choice after being rewarded on a ‘common’ transition, but switch first-step choices when rewarded on a ‘rare’ transition.

To facilitate learning in the patient population, the reward probabilities in this version of the task were stationary for long periods with abrupt shifts in reward (Castro-Rodrigues et al. 2022; Akam et al. 2015; Doody et al. 2022), rather than drifting continuously. Specifically, at any point in time one arm would be associated with a high reward probability (80% chance of payout) while each of the other three arms would be associated with low reward probabilities (8.3% chance of payout). The high reward option was associated with the same arm for a period of 32 trials, before switching to a different arm (unannounced to the participant). The entire study consisted of 288 trials (9 blocks of 32 trials). Participants received standardised instructions (see appendices; Fig.D.1). Model parameters had good recovery in this version of the two-step task (see appendices; Fig.D.2).

### 6.2.3 Two-step task statistical analyses

#### Performance and simple behavioural analyses

As a metric for performance, we looked at accuracy for the ‘planning’ choice at step-one. Accuracy was operationalised as the proportion of correct choices for step-one i.e. choices which, if the common transition occurred, would lead to the rewarding shape offered at step-two. Given the probabilistic reward structure of the task, this metric of performance is less noisy than overall reward. Step-one choices rather than step-two choices were examined because only step-one choices can capture planning across two steps, through a model of the transition structure between steps.

Next, we quantified the extent to which participants were sensitive to the transition structure of the environment through reaction times. Participants highly sensitive to the environmental model may be expected to slow down more after surprising 'rare' transitions compared to predicted 'common' transitions (Nussenbaum et al. 2020). For each participant, we computed the average step-two reaction times following a rare transition versus following a common transition. Paired t-tests were used to determine whether reaction times differed as a function of transition type.

Second, analysis of stay probability was used to assess model-free versus model-based behavioural strategies (Daw et al. 2011; Otto et al. 2013; Worbe et al. 2016; Friedel et al. 2014; Castro-Rodrigues et al. 2022). This is defined as the probability of repeating the first-step choice on a trial as a function of the outcome (reward versus no reward) and transition (common versus rare) on the previous trial. We used a logistic regression model, where the outcome, transition and transition–outcome interaction were all used as predictors to model stay choices. We then looked at differences in the transition-outcome interaction across groups.

### Reinforcement learning model

Choices were predicted using a reinforcement learning model with separate components capturing model-based and model-free learning (Daw et al. 2011).

The task involves three states, with only one first-step state, and two second-step states. In the following notation, $s_{1,t}$ corresponds to the first-step state taken at trial $t$ (which is always the same), while $s_{2,t}$ corresponds to the second-step state (dependent on the first choice and transition). In each state, there are two available actions ($a_A$ or $a_B$). Below, $a_{1,t}$ refers to the first-step action taken at trial $t$, and $a_{2,t}$ refers to the second-step action taken at trial $t$. Finally, the rewards for the two stages are denoted by $r_{1,t}$ (always zero after the first-step), and $r_{2,t}$ (one or zero).

**Model-free algorithm** The model-free algorithm updates the value of state-action pairs according to a SARSA ($\lambda$) temporal-difference reinforcement learner (Rummery and Niranjan 1994). At each stage $i$ of the two-staged trial $t$, the value for the chosen action ($a_{i,t}$) is updated:

$$Q_{mf}(s_{i,t}, a_{i,t}) = Q_{mf}(s_{i,t}, a_{i,t}) + \alpha\delta_{i,t} \tag{6.1}$$

Where $\alpha$ is a learning rate parameter, and the reward prediction error, (RPE; $\delta_{i,t}$), corresponds to the following:

$$\delta_{i,t} = r_{i,t} + Q_{mf}(s_{i+1,t}, a_{i+1,t}) - Q_{mf}(s_{i,t}, a_{i,t}) \tag{6.2}$$

Note that the prediction error is driven by different sources of information after the first versus second stage choices. At the first stage choice, reward is never received, so the update is driven by the second-stage value, $Q_{mf}(s_{2,t}, a_{2,t})$. At the second stage choice, the update is driven entirely by the reward received, $r_{2,t}$ (while the value of the next state is set to zero because the trial ends after two stages). Finally, at the end of the trial, the value of the first-stage choice is also updated with an eligibility trace. In other words, the RPE from the final choice is used to update the first choice, multiplied by an eligibility parameter ($\lambda$) (Sutton and Barto 1998):

**Model-based algorithm** The model-based algorithm updates its values for step-one using a model of the task structure – that is, the probabilities associated with transitioning between steps. For example, if the state in step-two was unlikely to occur after the first choice (rare transition of 25%), the algorithm correspondingly updates the value of the action in state one that most *commonly* reaches the rewarded state. Below, $s_A$ and $s_B$ denote the two possible second states. The value of the step-one actions ($a_j$) are computed in the following way according to the Bellman equation:

$$Q_{mb}(s_1, a_j) = P(s_A|s_1, a_j) \max_{a \in a_A, a_B} Q_{mf}(s_A, a) + P(s_B|s_1, a_j) \max_{a \in a_A, a_B} Q_{mf}(s_B, a) \tag{6.3}$$

This is re-computed at every trial from the current estimates of value. At step-two, model-based learning is equivalent to model-free learning, since the second step value purely reflects an estimate of the immediate reward (Daw et al. 2011).

**Choice algorithm** The influence of model-based versus model-free strategies can be quantified in the first step. The probability of choosing each first-step action is determined by a combination of model-based value, model-free value, and a 'repetition bias'. We follow Decker et al. (2016); Potter et al. (2017); and Nussenbaum et al. (2020) in modelling separate weights associated with the influence of model based value ($\beta_{mb}$) and model-free value ($\beta_{mf}$), alongside a parameter capturing biases to repeat the previous step one choice ($p$). We then model the probability of choosing each possible action at the first step using a softmax on the weighted contribution of these components:

$$P(a_i|s_{1,t}) = \frac{\exp\left(\beta_{mf} * Q_{mf}(s_{1,t}, a_i) + \beta_{mb} * Q_{mb}(s_{1,t}, a_i) + p * rep(a_i)\right)}{\sum_{a'} \exp\left(\beta_{mf} * Q_{mf}(s_{1,t}, a') + \beta_{mb} * Q_{mb}(s_{1,t}, a') + p * rep(a')\right)} \tag{6.4}$$

At the second step, the model-free value is used to predict choice, and a separate softmax temperature is fitted:

$$P(a_i|s_{2,t}) = \frac{\exp\left(\beta_{step-two} * Q_{mf}(s_{2,t}, a_i)\right)}{\sum_{a'} \exp\left(\beta_{step-two} * Q_{mf}(s_{2,t}, a')\right)} \tag{6.5}$$

The final model had six free parameters including a first-step weight for model-free value ($\beta_{mf}$), a first-step weight for model-based value ($\beta_{mb}$), a first-step weight for persistence (repeating the previous choice; $p$), a second-step softmax temperature for model-free value ($\beta_{step-two}$), learning rate ($\alpha$) and eligibility parameter ($\lambda$).

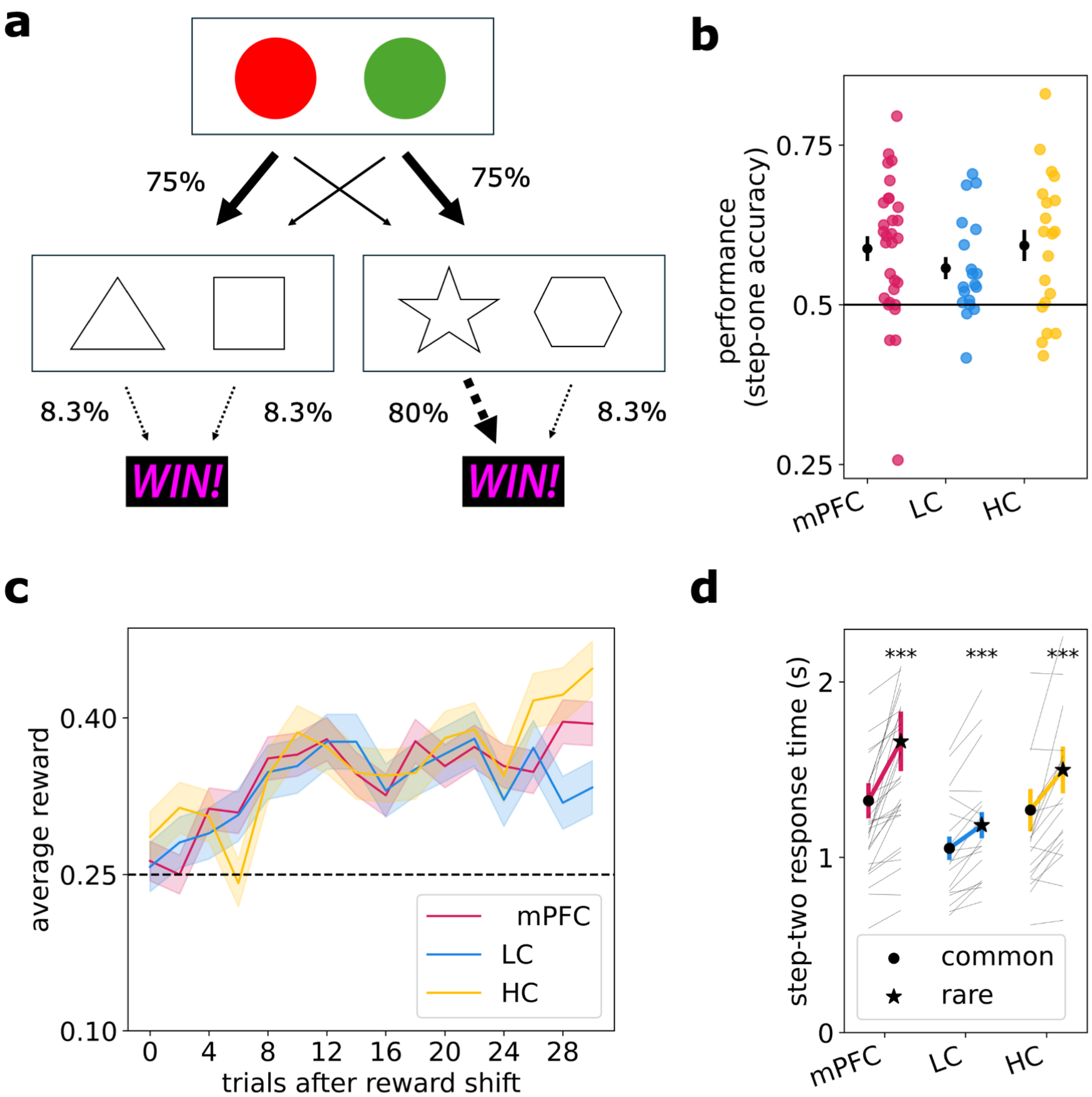


**Figure 6.2:** Two-step task. **(a)** Depiction of the task, where participants made two sequential decisions between colours (step-one) followed by shapes (step-two), with the aim of maximising wins. Arrows between step-one and step-two illustrate the common transitions (bold arrow, .75 probability) and the rare transitions (narrow arrow, .25 probability). Arrows between step-two and the reward outcome illustrate the probability of winning after selection of each shape. At any point in time, a particular shape (in this example, the star) was associated with high probability of winning the reward. The high-reward shape shifted every 32 trials. **(b)** Performance plotted by group. Performance is quantified as the proportion of correct choices at step-one i.e. choosing the colour which leads to the high reward set of shapes after a common transition. Error bars depict s.e.m., dots show individual data points, line depicts chance performance. **(c)** Average proportion of rewarded trials plotted by the trial number following a shift in reward contingency. The rewarded shape changed every 32 trials. Colour depicts group, confidence intervals depict s.e.m. **(d)** Step-two choice reaction times plotted for the three groups, split by whether the transition experienced was common or rare (circle shows common, star shows rare). Error bars depict s.e.m, and transparent lines show reaction times for individual participants. Stars depict significance, corresponding to the fact that in each group participants slowed down after experiencing a rare transition.

### Model Validation and Fitting

While the behavioural model described above has been used extensively in the literature, this dataset was collected using an adapted version of the two-step task which would be easier for patients with brain damage to learn, as used in similar studies with patient populations (Castro-Rodrigues et al. 2022). In this version, reward probabilities were stationary with abrupt shifts, rather than continuously drifting. Since previous studies have shown that model-free behaviour can be mistaken for model-based behaviour in environments with stationary probabilities (Akam et al. 2015), it was important to validate the model within the experimental paradigm to show that model-based behaviour could still be recovered by fitting the reinforcement models. Full details of model validation are included in the appendices, showing that all parameters including model-based weights are recoverable in this version of the task (see model validation).

A Bayesian hierarchical modelling framework was used to fit the reinforcement learning models to behaviour, allowing data pooling across participants to improve individual parameter estimates. Full details of model priors and fitting procedure are included in the supplementary materials. Models were coded in the Stan modelling language (Carpenter et al. 2017), and fitted to each dataset using the Cmdstanpy interface. Following previous studies (Decker et al. 2016; Potter et al. 2017), we did not include the first 9 choice trials in the analysis. Further details about the model-fitting procedure is included in the appendices (see model fitting).

### Lesion Group Comparisons

We began by investigating differences in performance and in sensitivity to the transition structure as a function of lesion. To test for differences in performance, we compared accuracy at step-one. ANOVA was used since this metric of performance was normally distributed. For transition structure sensitivity, we used the difference in reaction time following a common versus rare transition. Since the differences were not normally distributed, Kruskal-Wallis tests were used.

Two metrics were used to test the hypothesis that model-based planning would be lower in mPFC patients. Non-parametric one-sided Mann-Whitney U tests were used because both metrics violated assumptions of normality. First, we tested whether mPFC patients were less likely to modulate their stay-probabilities by the transition structure (lower transition-outcome interaction from the analysis of stay probabilities). Second, we formally quantified model-based planning from the reinforcement learning model, and tested if mPFC patients had lower model-based weights. In both cases we assessed the hypotheses that (a) mPFC patients had reduced model-based metrics compared to lesion controls and (b) mPFC patients had reduced model-based metrics compared to age-matched controls.

We had no other a priori hypotheses about parameters from the reinforcement learning model. Exploratory analyses were performed to determine if there was an effect of group on any of the reinforcement learning parameters, using Kruskal-Wallis non-parametric tests since assumptions of normality were violated in each case.

### Test for model-based behaviour

Since model-based weights recovered in the reinforcement learning model low, we performed a final analysis to determine if our data supported the use of model-based strategies at all. We performed a model comparison of two hierarchical RL models: one model containing all six parameters described in reinforcement learning models, and one five parameter nested model which excluded just the model-based beta weight (where the model-based weight is set to zero). The two models were compared using leave-one-out cross-validation (LOO-CV), where out-of-sample predictive accuracy is estimated using expected log predictive accuracy across data points (expected log pointwise predictive density; elpd). To quantify this, we used the ArviZ package (Kumar et al. 2019), which uses Pareto smoothed importance sampling (PSIS) to implement LOO-CV (Vehtari et al. 2017).

### 6.2.4 Results

#### Medial PFC damage does not affect performance in the two-step task

We found mPFC lesions did not impair performance at the two-step task. All groups in our study were sensitive to the general reward probabilities and transition structure of the task. In each group participants picked the correct colour at step-one more frequently than chance (Fig.6.2b; one-sample two-sided t-test comparing the mean proportion of correct step-one choices against 0.5; frontal patients: $t(29) = 4.44$, $p = 0.0001$; lesion controls: $t(18) = 3.23$, $p = 0.005$; age-matched controls: $t(19) = 3.70$, $p = 0.002$), showing a basic understanding of the reward structure. However, lesion damage had no effect on performance (ANOVA for effect of lesion group on step-one choice accuracy; $F(2, 66) = 0.71$, $p = 0.495$).

We also found all groups to be sensitive to the transition structure linking step-one and step-two as indicated by their reaction times (Nussenbaum et al. 2020). In all groups, reaction times in step-two were slower following the more surprising 'rare' transition compared to a 'common' transition (Wilcoxon signed-rank of mean reaction times after rare versus common transitions; frontal patients: $Z = 4.68$, $n = 30$, $p < 0.001$; lesion controls: $Z = 3.22$, $n = 19$, $p = 0.001$; age-matched controls: $Z = 3.47$, $n = 20$, $p = 0.0005$). There was no difference between lesion groups in sensitivities to the transition structure as shown in reaction times (Fig.6.2d; Kruskal-Wallis comparing the effect of group on the difference in step-two reaction time following common versus rare transitions: $H(2) = 2.35$, $p = 0.309$).

#### No difference in model-based strategies following mPFC damage

Next, we quantified whether mPFC lesions affected sensitivity to the task transition structure shown in choices. A learner who uses a 'model-free' strategy will be more likely to repeat their first step choice after being rewarded, regardless of whether the transition between steps was common or rare (although see Akam et al. 2015). However, a decision-maker who plans uses a model of the transition structure will be more likely to repeat their first step choice after being rewarded following a 'common' transition, but switch to the opposite choice when rewarded following a ‘rare’ transition. We found that mPFC lesions did not affect peoples’ model-based behavioural adjustments as measured through analyses of repetition probabilities (Fig.6.3; outcome–transition interaction on stay probability: mPFC $<$ lesion control; $n_1 = 30$, $n_2 = 19$, $U = 318.0$, $p = 0.754$; mPFC $<$ all controls; $n_1 = 30$, $n_2 = 20$, $U = 368.0$, $p = 0.912$).

Finally, we fit reinforcement learning models to the data to formally quantify the contribution of model-based and model-free reasoning, while also controlling for other behavioural factors such as persistence (the tendency to repeat the previous action), and learning rate (how quickly people update their behaviour). Medial prefrontal lesion patients did not show reduced model-based planning when formally operationalized in a reinforcement learning model (Fig.6.4, top left; model-based beta weight from RL model: mPFC $<$ lesion control; n1=30, n2=19, U=394.0, p=0.988; mPFC $<$ healthy controls; $n_1 = 30$, $n_2 = 20$, $U = 356.0$, $p = 0.868$). Our findings show that across a range of behavioural metrics, mPFC patients neither had worse performance nor showed a reduction in model-based planning in the two-step task.

### No effect of lesion on learning in the two-step task

We reported tests for the a priori hypothesis that model-based planning would be lower in mPFC patients, which we found was not the case. We had no other a priori hypotheses about parameters from the reinforcement learning model. We therefore performed exploratory analyses to determine if there was an effect of any group using Kruskal-Wallis non-parametric tests since assumptions of normality were violated in each case. None of the parameters from the RL model showed any significant differences as a result of group (Fig.6.4; Kruskal-Wallis tests; model-based weight: $H(2) = 5.19$, $p = 0.075$; model-free weight: $H(2) = 0.60$, $p = 0.741$; step-two beta: $H(2) = 0.640$, $p = 0.726$; persistence bias: $H(2) = 5.49, p = 0.065$; alpha: $H(2) = 1.03$, $p = 0.597$; lambda: $H(2) = 0.08$, $p = 0.959$).

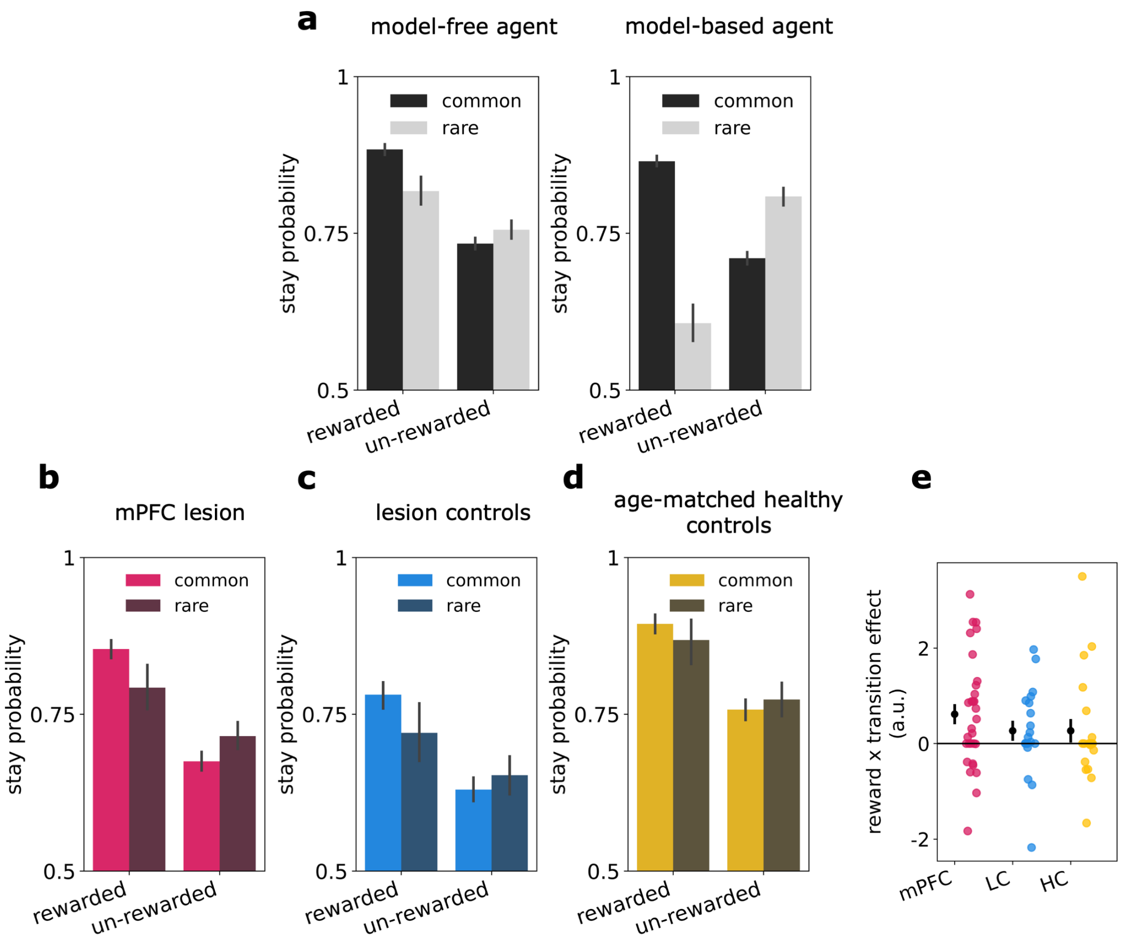


**Figure 6.3:** Analyses of stay probabilities. **(a)** Simulated choices from a model-free agent (who ignores the transition structure between step-one and step-two), and a model-based agent. The probability of repeating the same step-one choice (y-axis) is plotted as a function of whether the previous trial was rewarded or unrewarded (x-axis). Shade depicts whether the previous outcome had followed the common transition (dark) or rare transition (light). A model-based agent will be more likely to repeat their step-one choice if a rare transition led to reward, and less likely to repeat their step one-choice if a common transition led to no reward. **(b,c,d)** Same analysis as (a) performed on empirical data in the three patient groups. **(e)** Results from the stay-probability analysis where the probability of staying is analysed depending on the type of transition and whether the previous trial was rewarded. The outcome-transition interaction beta weights are plotted as a function of group, where higher beta values indicate more model-based modulation of behaviour. Error bars depict s.e.m, dots show individual data points, line depicts no effect.

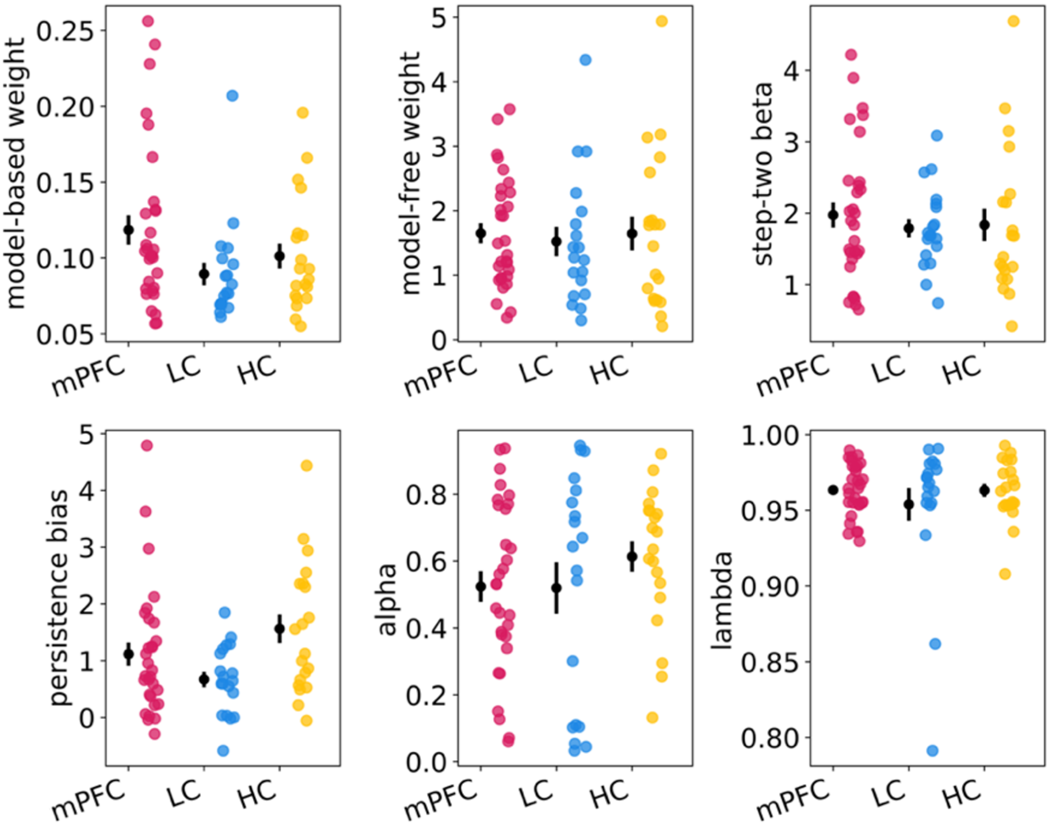


**Figure 6.4:** Parameters from the reinforcement learning model, plotted by group (mPFC lesion patients, lesion controls, healthy controls). No parameters show significant effect of group.

### Negligible model-based behaviour across older participants

All participants in our task were older than standard cohorts, which has been shown to be related to reduced model-based reasoning (Eppinger et al. 2013). In addition, previous studies have found that the use of model-based strategy is highly dependent on a range of experimental factors such as task framing and instructions (Feher da Silva and Hare 2020; Castro-Rodrigues et al. 2022). When plotting stay probabilities in our experimental cohort of older participants, all groups more closely resembled model-free planners than model-based planners (Fig.6.3a-d). To determine whether participants were using model-based strategies at all, we directly compared the hierarchical reinforcement learning models with and without this parameter. The full (6-parameter) model containing the model-based beta weight was no better than the 5-parameter model without the model-based component (elpd_loo with model-based component: -20852.76; elpd_loo without model-based component: -20852.56). In other words, we did not find evidence that participants were generally using model-based value to guide behaviour in this study, despite showing sensitivity to the transition structure through reaction times.

### 6.2.5 Interim Discussion

We did not find any effect of mPFC lesions on any aspect of behaviour in the two-step task. While this is consistent with the theory that mPFC lesions affect planning more in complex state spaces, this study comes with a crucial caveat. Importantly, use of model-based strategies were negligible across all groups in our cohort. Our findings are consistent with previous studies showing reduced use of model-based strategies in older adults (Eppinger et al. 2013). Previous papers have also reported drastically improved model-based reasoning after receiving explicit instructions explaining the structure of the task, compared to no instructions at all (Feher da Silva and Hare 2020; Castro-Rodrigues et al. 2022). Participants in our task received a level of instruction somewhere in the middle of these two extremes (Fig.D.1), so it is therefore possible participants would have used more model-based strategies if more explicit instruction were given. In any case, the general lack of model-based planning in this study limits the conclusions we can draw about the relationship between model-based planning and mPFC lesions within the two-step setting.

Despite these limitations to interpreting the results, it is important to note that mPFC lesion patients performed as well as lesion controls and age-matched controls in this probabilistic decision-making task. Overall, all groups performed well above chance, and showed sensitivity to the task structure through reaction times. They were also able to update their choices after a shift in the rewarding outcome, which has a structure analogous to reversal learning. The fact that mPFC patients were no worse at any of these elements supports previous findings that mPFC damage does not affect learning of contingent relations between actions and outcomes or "credit assignment" (Noonan et al. 2017). In addition, it supports the idea that decisions between binary choices based on a single attribute are not affected by mPFC lesions (Pelletier and Fellows 2019).

## 6.3 Experiment 2: Four-in-a-row

Compared to the two-step task, the four-in-a-row paradigm provides a substantially more complex setting to investigate planning. Since full simulation of the state space is computationally impossible, planning requires selecting possible futures to explore, and using heuristics to evaluate positions. This naturalistic setting therefore provides more opportunity to capture the planning deficits lesion patients show in real life, as well as provide insights into the underlying cognitive components causing these deficits.

### 6.3.1 Participants

For the four-in-a-row study, we collected data from a total of eighteen lesion patients (mean age=59, 9 females) and thirty age-matched healthy control participants (mean age=58, 16 females). The patient population was recruited from a database of individuals who had previously visited the John Radcliffe Hospital and consented to be contacted for research studies. As previously described for the two-step study, the Harvard-Oxford cortical structural atlas mask of ventromedial prefrontal cortex was used to a priori assign patients to the mPFC group or lesion control group (mPFC lesion group: N=10; Fig.6.1c, left; two-step lesion control group: N=8; Fig.6.1c, right).

### 6.3.2 Experimental methods

All participants played 'Four-in-a-row', which is a computer-based game where the player's aim is to place four pieces of their own colour in a line (vertically, horizontally or diagonally) before the computer opponent (Fig.6.5a). Participants alternated turns with the computer opponent to place pieces of their colour (black or white) on an empty space in a four-by-nine grid. A board of this size has approximately 1.2 x $10^{16}$ non-terminal states. Each game could terminate in a win (participant gets four pieces in a row), loss (computer opponent gets four pieces in a row), or draw (no one gets four in a row before the grid is full). Participants alternated between playing black pieces, and playing white pieces, where the first move of each game was always black.

The task was programmed in JavaScript and participants completed the game in a web browser (hosted on Amazon Web Services). For the patients, the researcher remained on the telephone throughout the session to help with any technical difficulties with the task, but all instructions and training were standardised and completed online. All participants received the same training which involved computer-based instructions, two practice games, and five comprehension questions. After training, participants completed forty games total, starting at an easy level and advancing to more challenging opponents based on a staircase algorithm. The healthy control participants matched for age were recruited from Prolific.co, and received the same training and study procedure with the only difference that the researcher was not present on the telephone.

The set of AI opponents comprised of 200 difficulty levels, previously developed by Opheusden and colleagues (van Opheusden et al. 2023). Participants were matched to opponents of an appropriate playing strength using a staircase procedure. The 200 difficulty levels were divided into five categories of playing strength (with 40 agents per category). For the two practice games, the level was initially set to 1 (easiest possible). At the beginning of the study, the first opponent was randomly drawn from category 2 (i.e. levels 40-79). After each game, the opponent was updated based on the outcome: after a loss, a new opponent was drawn from the category below; after 1 win or a draw, a new opponent was drawn from the same category, and after 2 wins a new opponent was drawn from the category above.

### 6.3.3 Four-in-a-row statistical analyses

#### Task performance

We assessed task performance through estimations of playing strength using the Elo System (Elo 1978; van Opheusden et al. 2023). Each player is given a relative score based on their history of wins, losses and draws against the same pool of opponents (computer levels). To quantify elo ratings, we used the Bayeselo algorithm (`https://www.remi-coulom.fr/Bayesian-Elo/`), originally developed for rating chess players but which has previously been used to rate performance in four-in-a-row (van Opheusden et al. 2023). Each category (1-5) was treated as a single 'player'. This measure of playing strength does not factor in differences in cognitive strategies, but rather derives a ranking across the total set of players.

#### Planning Model

To formally quantify differences between groups in the cognitive components of planning, we used the model developed to accompany Four-in-a-row by in Van Opheusden and colleagues. The model comprises a heuristic evaluation function and a best-first search algorithm. Given the size of the state space, it is impossible to plan across all possible futures in this task. For this reason, the model assumes agents search the space efficiently. This rests on two assumptions, namely that simple features are used to estimate the value of moves, and that the most promising moves are explored first during planning. The components of the cognitive model and reliability of parameters have been extensively validated in a previous paper (van Opheusden et al. 2023). Below, we outline the main parts of the planning model.

**Heuristic value function** The heuristic value function determines the value of each board state $V(s)$ according to a combination of heuristic 'features'. The algorithm posits 5 evaluative features: connected 2-in-a-row (i.e. two consecutive pieces), un-connected 2-in-a-row (i.e. two non-consecutive pieces which could form the basis of a 4-in-a-row win), 3-in-a-row, 4-in-a-row, and proximity to the board center. It is assumed that people approximate the value of the board through a weighted sum of the value of these features. For each participant, individual weights are associated with each feature. In addition, features are scaled differently (with a scaling constant $C$) depending on whether their colour belongs to the current 'active' player, or the other 'passive' player during the simulated move, capturing the fact that features are more valuable if they belong to the player whose current move it is. The final value function is as follows:

$$V(s) = w_{centre}V_{centre} + c_{black}\sum_{i\in F} w_i f_i(s, black) - c_{white}\sum_{i\in F} w_i f_i(s, white) + \mathcal{N}(0,\,1) \tag{6.6}$$

Where $c_{black} = C$ and $c_{white} = 1$ whenever black is to move in state $s$, and $c_{black} = 1$ and $c_{white} = C$ whenever white is to move in state $s$. The last term $N(0, 1)$ adds Gaussian noise with mean zero and unit variance.

**Tree-search algorithm** Guided by the value function, the tree-search algorithm constructs a partial decision-tree using best-first search (Dechter and Pearl 1985). On each iteration, the value function determines which position to explore, resulting from the sequence if both players choose their highest-value moves in the current tree. All legal moves from the selected position are evaluated, and values are back-propagated to predecessor nodes up to the root of the tree using minimax rule. Moves which are lower than the best move minus a 'threshold' ($\theta$) are pruned. This reflects the fact that people cannot do an exhaustive search over the state space, and aligns with empirical evidence that people 'prune' branches with initial low values (Huys et al. 2012). Finally, at the end of each iteration, there is a probability of the search being terminated with a stopping probability parameter.

**Sources of noise** In addition to the parameters related to the value function and tree search, there are two additional parameters related to sources of noise. 'Feature drop' models selective attention by capturing the probability of missing a feature on a particular trial (a particular feature is 'dropped' from $V(s)$ at all points in the tree). Finally, lapse rate parameter captures the probability of choosing a random move on any particular trial.

**Summary parameters** Following the original paper developing this model (van Opheusden et al. 2023), we focus our analyses on three final summary parameters: depth, heuristic quality, and feature drop rate. These summary parameters have better reliability and test-retest stability than the lower-level sub-parameters (van Opheusden et al. 2023). To calculate these summary parameters, each individual's set of sub-parameters are used to simulate moves across 5482 unique states. This simulation is repeated 10 times, to minimise variability in noise. The final parameters are then calculated in the following way:

1. **Planning Depth.** This parameter captures the average length of the decision tree across simulated searches (i.e. how far a participant tends to plan into the future). To derive depth, sub-parameters for each participant are used to simulate moves by building a decision tree. Planning depth is formally quantified as the length of the forwardly simulated sequence (depth of decision tree), averaged across simulations.

2. **Heuristic Quality.** This parameter captures the difference between subjective weights for the five heuristic features, and the optimal weights. The subjective value is calculated across the 5482 states using each participants weighted combination of features. The optimal state values from Van Opheusden et al. 2023 were calculated by running the model with no noise and no pruning until convergence on the state value. Heuristic quality is the correlation between the subjective state value (the participants weighted combination of features), and the objective optimal value.

3. **Feature Drop Rate.** This parameter directly corresponds to an estimated sub-parameter in the model, capturing the probability that a participant overlooks a feature instance on the board. When a feature is dropped, its weight is temporarily set to zero in a particular move, mimicking an

attentional lapse. See appendices for an illustration of how the feature drop rate parameter captures behaviour, specifically when an opportunity for winning the game is over-looked, or a three-in-a-row threat from an opponent is missed (Fig. D.4).

**Model fitting** The pipeline for fitting the four-in-a-row model is computationally expensive (fitting one participant's data for a single model requires approximately $10^{14}$ floating-point operations). Model fitting was carried out by our collaborators at NYU (Bas Van Opheusden), on the NYU high-performance cluster (Intel Xeon E5-2690v2 CPUs 3.0 GHz) with a parallel implementation of inverse binomial sampling, which uses 20 cores.

### Comparisons of behaviour across groups

Across all comparisons of group behaviour, non-parametric tests were used because variables violated assumptions of normality. First, we established whether the three groups differed in performance. For this initial test, we used Kruskal-Wallis to determine if Elo ratings differed as a function of group (mPFC group, lesion controls, healthy controls). We followed this result with two-sided Mann Whitney U tests. The critical test was whether individuals with mPFC lesions performed worse than other individuals with lesion damage (mPFC patients versus LC). Following this, we verified that mPFC patients differed to healthy controls (mPFC patients versus HC).

Since larger lesions are often associated with lower performance(Karnath et al. 2018), as an additional control, we verified that performance was truly related to the location of damage rather than the size of the lesion. To do this, we predicted Elo ratings using the number of voxels damaged within the mPFC, as well as the number of voxels damaged in total:

$$Elo = \beta_0 + \beta_1 damage_{mPFC} \beta_2 damage_{total} \tag{6.7}$$

Where $damage_{mPFC}$ refers to the number of voxels damaged within the ventromedial prefrontal mask, and $damage_{total}$ refers to the total number of voxels damaged across the brain.

Once it was established that mPFC lesion patients had a performance deficit, we used one-sided Mann-Whitney tests to investigate the three different hypotheses, namely that mPFC lesion patients planned at lower depth, were more likely to drop valuable features, or had lower heuristic quality. Again, we used the critical test to ask whether there was an effect of mPFC lesion within the lesion population (mPFC patients versus LC), following up to verify that the difference existed compared to healthy controls (mPFC group versus HC).

### 6.3.4 Results

#### Medial PFC damage impairs planning in Four-in-a-row

Medial PFC lesion patients had lower playing scores compared to both control lesion patients, and age-matched controls (Fig.6.5b). To quantify playing strength, we used the Bayeselo algorithm, originally developed for rating chess players but which has previously been used to rate performance in four-in-a-row (van Opheusden et al. 2023). A Kruskal-Wallis test indicated there was a difference in Elo ratings between the three groups ($H(2) = 7.20, p = 0.027$). Lesion patients with mPFC damage had lower Elo ratings compared to patients with damage elsewhere (mPFC Elo Median=–66.0; patient control Elo Median=33.5; difference in ratings: Mann Whitney $U = 14$, $p = 0.021$), and compared to a healthy age-matched population (median age-matched control Elo Median=27.5; difference in ratings: Mann Whitney $U = 71.0$, $p = 0.014$), while the difference between the control lesion patients and the healthy age-matched population was not significant (Mann Whitney $U = 127.0$, $p = 0.816$).

To eliminate the possibility that our results were driven by differences in lesion size, we controlled for total lesion volume in a regression analysis. Within the patient population we found lower Elo scores were predicted by larger vmPFC lesions ($\beta = -0.08$, $p = 0.033$) but not by larger lesions in general ($\beta = 0.00$,

$p = 0.664$). These findings suggest that damage to the vmPFC is predictive of lower performance in this task.

### Planning deficits following mPFC damage are linked to attentional lapses

What is causing the performance deficit shown by mPFC lesion patients in Four-in-a-row? The behavioural model of the task formally characterized three separable cognitive components involved in the game (van Opheusden et al. 2023). Better performance in the task is associated with planning further into the future (higher 'depth'), better evaluation of the board heuristics (higher 'heuristic quality') and a lower tendency to miss valuable feature on any trial (lower 'feature drop rate'). We therefore tested whether any of these trends were true for the mPFC group. The full description of these parameters can be found in the section summary parameters.

We found that mPFC patients were more likely to miss valuable features than lesion controls (Fig.6.5e; higher feature drop rate: one-sided Mann-Whitney; $n_1 = 10$, $n_2 = 8$, $U = 72.0$, $p = 0.002$), and compared to healthy controls (one-sided Mann-Whitney; $n_1 = 10$, $n_2 = 30$, $U = 246.0$, $p = 0.001$). We also found that mPFC lesion patients planned to a lower depth than lesion controls (Fig.6.5f; lower depth: one-sided Mann-Whitney; $n_1 = 10$, $n_2 = 8$, $U = 15.0$, $p = 0.013$), but this did not survive comparison with age-matched controls (one-sided Mann-Whitney; $n_1 = 10$, $n_2 = 30$, $U = 103.0$, $p = 0.073$). Medial PFC patients did not demonstrate a lower ability to use good heuristics compared to either lesion controls or healthy controls (Fig.6.5g; patient controls: $U = 23.0$, $p = 0.072$; age-matched controls: $U = 104.0$, $p = 0.078$).

While we did not analyse the model sub-parameters in general (because these are less reliable than the summary parameters (van Opheusden et al. 2023)), we examined one relevant sub-parameter which controls for an alternative form of 'noise'. The 'lapse rate' parameter captures the probability of making a random move on any trial. While lesion patients generally had higher lapse rate than healthy age-matched controls, there was no difference as a function of lesion location in mPFC (analysis included in appendices; see Fig.D.3). This suggests that mPFC patients

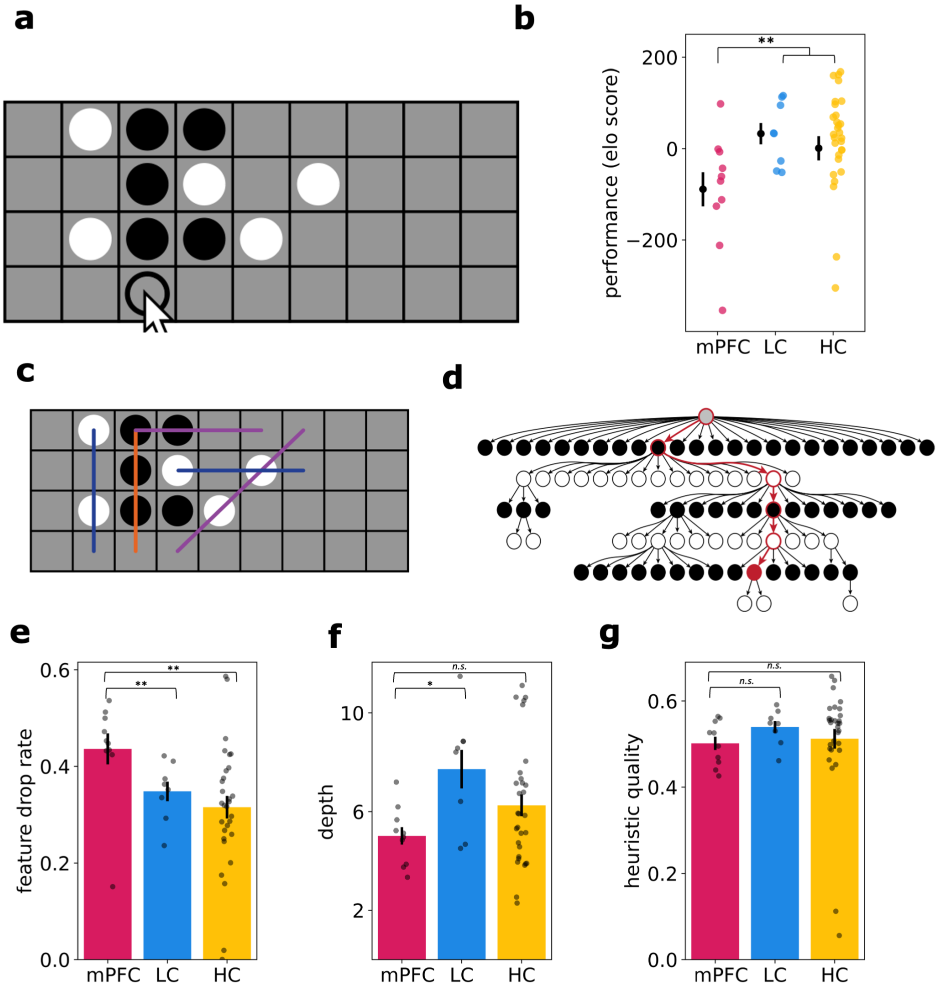


**Figure 6.5:** Main results from the Four-in-a-row study. **(a)** Depiction of the task, where participants aimed to place four pieces of their colour in a row. Arrow depicts example winning move (reprinted with permission from van Opheusden et al. (2023)). **(b)** Elo ratings (a metric of performance, or playing strength) as a function of group. Error bars show s.e.m., dots show individual participant ratings. Stars show statistical significance, where mPFC lesion patients perform worse than both lesion controls and age-matched controls. **(c)** The computational model consists of a heuristic value function (for evaluating states) and a tree-search algorithm (for simulating future moves). The value function corresponds to a linear combination of heuristic features critical for playing the game. Coloured lines depict example features, where purple shows connected two-in-a-row, blue shows un-connected two-in-a-row, and orange shows three-in-a-row. Within the model, heuristic quality refers to how closely an individuals weights for each feature match the optimal weights. Feature drop rate refers to the probability of over-looking a feature on the map, on any trial. **(d)** The tree-search algorithm constructs a partial decision-tree using best-first search (see section Planning Model for full description). Depth refers to the average length of forward search, illustrated in the example with the red trajectory. **(e,f,g)** Summary parameters from the planning model, plotted by group. Error bars show s.e.m., dots depict individual participant parameter estimates. Stars depict result of testing the three hypotheses, namely that mPFC lesions are associated with higher feature drop rate, lower depth, and lower heuristic quality than the two control groups.

are selectively more likely to have a lapse in attention which means they miss a particular valuable feature, rather than a tendency to respond more randomly.

### Differences in feature drop rate cannot be explained by failures to block opponents

Medial PFC patients in our study are more likely to miss valuable features while planning. This is consistent with the theory that mPFC is critical for orienting attention to the relevant features of the problem. However, there is an alternative explanation dependent on the type of features dropped. The feature drop parameter does not differentiate between whether a dropped feature corresponds to a high value feature for the opponent (thus lower value for the player) or a high value feature for the player (see Fig.D.4 for illustration of these two cases). In other words, patients could be specifically missing features related to the opponent winning, compared to features signifying their own win. Since mPFC lesions have also been linked to deficits in theory of mind, this is a critical alternative explanation to rule out. To investigate this, we performed a simple analysis independent of the model.

Non model-based analyses of wins and losses are difficult in four-in-a-row because opponent difficulty changes as a function of the players' performances, according to a staircase algorithm. Moreover, opponent difficulty is a complex construct that does not map neatly onto any interpretable strategies, but rather was derived through iteratively ranking AI opponents using Elo ratings (van Opheusden et al. 2023). However, we reasoned that if participants were specifically impaired at representing opponent theory of mind, this should be apparent in how frequently they make moves to block an opponent. Failures to block an opponent directly lead to an opponent win, and are therefore roughly balanced across the experiment as a result of the staircase algorithm. However successful blocking moves provide a more indirect metric of participants' ability to spot opponent opportunities, and therefore provide a cleaner metric of opponent tracking in this game.

We counted the number of times each participant successfully blocked an opponent's chance of winning (blocking an opponent 3-in-a-row) across the forty games.

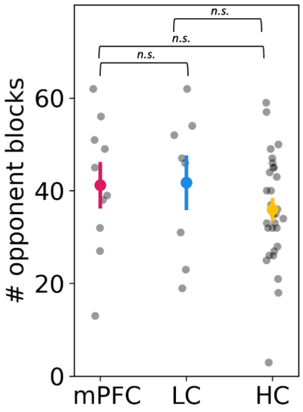


**Figure 6.6:** Four-in-a-row opponent blocking moves. Total number of opponent blocks across the study, plotted by patient group (mPFC lesion, lesion control, healthy control). Opponent blocks are events where the player puts a tile in the fourth remaining position next to an opponent three-in-a-row, terminating the opponents chance of winning the game with four pieces in a row. The fact that mPFC patients were not worse at blocking opponent wins suggests their attentional lapses are more pervasive rather than being explained by a specific tendency to neglect opponent moves, or a deficit in theory of mind.

We found no difference between the three groups, supporting the idea that patients are not specifically failing to block the opponent (Fig.6.6, one-way Anova on total blocks by group: $F(2, 45) = 1.00, p = 0.373$). This suggests that the tendency to miss relevant information is more pervasive, and cannot be explained by a specific deficit in failures to model the opponent's strategy.

### 6.3.5 Interim Discussion

We found mPFC patients were impaired at planning in four-in-a-row. We investigated three possible hypotheses for why planning was impaired: deficits simulating the future, poor evaluation of features generally, or the tendency to overlook valuable features when planning. We found patients were far more likely to drop valuable features than both lesion controls and age-matched healthy controls. We also found that mPFC patients planned to lower depth than lesion controls, but this did not survive comparison with age-matched healthy controls. We also investigated the possibility that mPFC patients were specifically missing features which signal an opponent's win, consistent with deficits in theory of mind. We found mPFC patients blocked opponent wins as many times as the control groups, suggesting the attentional deficit is more general than a failure to miss opponent moves. In other words, these analyses show mPFC patients are more likely to miss valuable features when deciding where to place their tile, despite showing no systematic disturbances in board evaluation (heuristic quality), or consistently planning less far into the future (depth). This is consistent with the idea that goal-directed attention is disturbed following lesions to vmPFC, which is the case for all our mPFC patients.

## 6.4 General Discussion

Classical studies of frontal lesion patients have emphasised the need for tasks that are sufficiently naturalistic to elicit complex behaviours, yet can provide specificity about cognitive functions. One such complex behaviour is planning, where mPFC patients show disturbances across a range of settings. We investigated how mPFC damage affects planning capacities using two paradigms from computational cognitive science: the relatively simple two-step task and substantially more complex four-in-a-row task. Four-in-a-row provides a computational framework designed to disentangle the contributions of different cognitive components to complex planning (van Opheusden et al. 2023).

Patients with damage to mPFC demonstrated impaired performance in four-in-a-row, but not the two-step task, compared to lesion control patients and age-matched healthy controls. We investigated how the four-in-a-row impairment related to the cognitive components identified by the model, through three possibilities: that mPFC damage leads to shorter depth of planning, worse heuristic evaluation overall, or a higher probability of missing relevant features in decisions. We found strong evidence for the third hypothesis, namely that mPFC patients were more likely to overlook critical relevant information on the board.

Four-in-a-row characterizes planning in a vast state space across multiple steps, providing a more ecological assessment of complex planning than previous laboratory planning tasks (van Opheusden et al. 2023; Ma et al. 2022). In contrast, the ‘two-step’ task consists of a maximum horizon of two steps, with only two options presented at each step. Across a large sample, we found that mPFC patients showed no behavioural differences in the two-step task compared to controls. These results are consistent with the idea that mPFC damage affects planning more in complex state spaces. However, our interpretation is limited by the caveat that no groups demonstrated strong model-based planning in our two-step study. This is likely to be related to the age of the population, where older cohorts generally show lower model-based planning (Eppinger et al. 2013).

In the four-in-a-row study, mPFC patients were not systematically worse at identifying what constituted a good move than controls (‘heuristic quality’). This metric captures the difference between each participant’s value function (how they weigh up the heuristic features on the board), and the optimal value function. This finding suggests mPFC patients showed no systematic differences in identifying the kinds of moves which would help them to win (for example, completing a three-in-a-row). Instead, they had a higher probability on each trial of missing relevant information (‘feature drop’).

Why were mPFC patients more likely to miss valuable features when planning in four-in-a-row? One possibility is that these missed features reflect lapses in orienting attention to goal relevant information. In complex environments with myriad options available, selective attention plays a critical role in narrowing down the dimensionality of the task (Niv 2019; Leong et al. 2017). Damage to mPFC has been shown to selectively impair decision-making in multi-option choice settings where irrelevant information is distracting, but not in simple binary choice settings (Noonan et al. 2017; Vaidya et al. 2018). This idea is also consistent with previous studies showing mPFC damage impairs the allocation of attention to reward-predictive features of the environment (Vaidya and Fellows 2015b), while activity in mPFC predicts orientation of attention to current goals in healthy individuals (Günseli and Aly 2020).

While our results are consistent with the possibility that task-relevant features are only dropped in visual attention (i.e. missed opportunities on the board), it is likely this deficit in selecting relevant information generalises outside of the visual context. For example, selection of task-relevant variables may also be impaired within memory or within simulation of future states. Our current paradigm cannot distinguish between failures to select in visual attention and failures to select within the simulation. However, mPFC is likely to be involved in sustaining key contextual variables which also guide internal computations (Wilson et al. 2014). A critical idea is that mPFC could represent a 'cognitive map' of current task-relevant information for guiding behaviour; whether that behaviour depends on selection in visual feature space, selection among model-based simulations, or selection within memory (Behrens et al. 2018; Wilson et al. 2014).

As well as showing a tendency to miss valuable features, we found that mPFC patients planned to lower depth than lesion controls, but not healthy controls. It is possible that with more power from a larger sample we would see lower depth in planning in mPFC patients compared to healthy controls too. Problems simulating the future would lead to suboptimal planning and has been one of the proposed explanations for planning deficits in mPFC patients (Bertossi et al. 2016a; Bertossi et al. 2016b; Bertossi et al. 2017; Fellows and Farah 2005).

One possibility is that mPFC is not directly involved in simulating future trajectories, but rather supports the goal-directed recruitment of these simulations taking place in other areas such as hippocampus. Consistent with this idea, damage to hippocampus has been shown to impair model-based planning in rodents (Miller et al. 2017) and most recently in humans (Vikbladh et al. 2019). For example, Vikbladh and colleagues demonstrated that patients with damage to the hippocampus following epilepsy treatment showed lower model-based planning in the two-step task compared to controls. This deficit was proportional to the extent of their hippocampal damage. Greater functional coupling between hippocampus and mPFC has also been shown to predict better inferences in healthy individuals (Zeithamova et al. 2012). It is notable that many tasks involving model-based inferences find activity in both mPFC and hippocampus (Barron et al. 2013; Barron et al. 2020; Park et al. 2021). While mPFC may be required to modulate processing through representation of the current state, other areas may perform more specific computations. For example, mPFC could coordinate context-dependent computations in other areas such as simulation or memory recall in the hippocampus (Zeithamova et al. 2012), or value learning in the basal ganglia (Blanco-Pozo et al. 2024).

## 6.5 Contributions

The two-step task dataset was collected by Matthias Raemaekers, Rowan Board and Patricia Lockwood. Harry Ward and myself collected the four-in-a-row dataset. Bas Van Opheusen and Xinlei Lin fit the computational planning model to the four-in-a-row data set. Jan Grohn, Sanjay Manohar, and Weiji Ma contributed to supervising the analyses included in this chapter. Jill O'Reilly gave feedback on the writing and interpretation.

# 7

# General Discussion

## 7.1 Goals and state representations

In **chapter 1**, I described a tension in the way that different disciplines have characterised goals. Psychologists have traditionally defined goal-directed behaviour in terms of its flexibility. To be pursuing a goal, rather than responding habitually, is to remain sensitive to the value of the goal and the contingencies in the environment. Yet philosophers and behavioural economists characterise goals as being inflexible: people fail to abandon goals in light of changing information. One response to this conundrum is to consider what kinds of information animals remain flexible to during goal pursuit.

In this thesis I have proposed that goal pursuit involves constraining the state representation to prioritize goal-relevant information in attention. This specifically predicts that animals will show reduced flexibility in response to sources of value which are irrelevant for the currently pursued goal. In **chapter 3**, I presented evidence that over the course of goal pursuit, people lost sensitivity to the value of alternative goals both in reaction times and choices. Consistent with this, selective attention was increasingly captured by the current goal, even outside the decision setting. These findings are consistent with the idea that goal pursuit has a global effect on how information is processed which enables prioritization of current goals at the expensive of goal irrelevant information.

In classic devaluation paradigms or for Tolman's rats running through the maze, competing goals are not present. These scenarios therefore do not provide opportunities for observing how goal pursuit could affect flexibility in response to the value of alternative goals. However, an interesting recent study presented rodents with choices between two competing homeostatic goals (Richman et al. 2023).

Rodents deprived of both food and water were presented with free choices between competing needs on every trial (Richman et al. 2023). A good choice strategy would be to select the option which relieved the most urgent need at every trial, as predicted by accounts of homeostatic reinforcement learning (Keramati and Gutkin 2014). However, instead, Richman and colleagues found that behaviour was best characterised by persistent bouts of trials where a particular type of good was chosen repeatedly, before the animal switched over to addressing the other need. In other words, while behaviour was sensitive to the animal's underlying needs for the two homeostatic goals, it was not defined by them. Rather, behaviour was structured by periods of sustained pursuit, where the currently pursued goal could be decoded in neural activity between trials. These underlying goals, sustained across time, seem to interfere with the 'optimality' of the animal's behaviour by delaying the choice of more urgent needs.

## 7.2 Commitment biases and optimality

Why would people – and potentially other animals – show these suboptimal behaviours to over-persist with current goals? In **chapter 3** I found that most healthy people over-persist with goals, while in **chapter 5**, a group of individuals who show lower goal persistence demonstrated a behavioural advantage. This chimes with classic behavioural economic accounts which consider 'sunk cost biases' to be an irrational weighting of decision variables (Arkes and Blumer 1985). However, there are various reasons why strong commitment to current goals might be beneficial in ecological settings. One reason could be that continual deliberation between possible goals is expensive, and resources are better allocated to achieving chosen goals, as an example of 'resource rational' decision making (Gershman et al. 2015; Lee and Daunizeau 2021). Another reason could be that in natural environments where reward is uncertain or estimates are noisy, having a fixed bias to favour continuation with the current goal allows for better outcomes overall (Johnson et al. 2013).

If commitment to goals can be more 'rational' in some settings than others, an interesting question is whether people are capable of adjusting their levels of commitment to reflect this. For example, 'resource rational' accounts would suggest that commitment biases should be reduced if deliberation is inexpensive or where fewer options are present. We might also predict that when people have more confidence in their value estimates or more certainty about the future, commitment to a selected goal is weaker.

This approach has proved to be informative for understanding other sources of individual variability in cognitive biases including preferences for risk or delay discounting. For example, both people and animals foraging in ecological environments flexibly adapt their preferences for risk in light of relevant variables such as their own energy or resource level (Caraco 1981; McNamara and Houston 1992; Kolling et al. 2012). Another example is the finding that children adjust their willingness to wait for better future rewards (delay discounting) depending on how reliable they believe the environment to be (Kidd et al. 2013). A different perspective on these cognitive biases has assumed stability within individuals, but emphasised how they are shaped during development by environmental and cultural factors such as socioeconomic uncertainty (Belsky et al. 1991; Griskevicius et al. 2011; Amir et al. 2018). Future studies could examine the flexibility of goal commitment biases in the context of relevant variables such as the predictability of the environment or the cost of deliberation.

## 7.3 VmPFC and maintaining goals

In **chapters 4** and **5**, I presented various lines of evidence identifying vmPFC with goal commitment. First, vmPFC carries sustained contextual information about current goals which persists between trials. Second, baseline activity predicts individual differences in commitment and attention to the current goal. Finally, vmPFC damage reduces goal commitment, suggesting it is causally involved in maintaining persistence with the current goal.

How exactly do we explain the role of vmPFC in supporting current goals? After choosing to pursue a goal, subsequent decisions must be guided by goal-relevant variables. As discussed in **chapter 2**, vmPFC supports context-dependent behaviour across many settings: inferring the variables which are relevant for guiding behaviour given the current state of the world (Hampton et al. 2006; Wilson et al. 2014; Schuck et al. 2016; Mante et al. 2013; Castegnetti et al. 2021; Frömer et al. 2019). I have argued that analogies can be drawn between vmPFC's role in guiding behaviour when the experimenter manipulates the current context, and when the subject has chosen their own context by deciding to pursue a particular goal.

How could vmPFC orchestrate context-dependent computation across the brain? Medial PFC activity has been shown to modulate dopaminergic prediction errors in light of the current state (Starkweather et al. 2018). Taking insight from different artificial neural network architectures, Blanco-Pozo and colleagues modelled these dynamics using a recurrent neural network where information about the current context was continually fed back into the network and thus sustained across time (Blanco-Pozo et al. 2024). In contrast, the basal ganglia was modelled as a feed-forward network where the input was partly defined by the context signal from PFC. In other words, within this model recurrent medial PFC dynamics influence processing in other areas through a sustained representation of the current state. This is consistent with many studies showing that mPFC activity shows these sustained representations of contextual variables across decisions, which have an influence on behaviour (Abitbol et al. 2015; Mehta et al. 2019; Bari et al. 2019).

An interesting line of investigation for future studies relates to how wider neural processing adapts to the context of current goals. For example, dopamine responds to reward proximity as well as reward value (Howe et al. 2013; Hamid et al. 2016b; Collins et al. 2016), and some have argued that prolonged dopamine signalling could sustain motivation towards goal completion by invigourating behaviour as goals are approached (Howe et al. 2013; Berke 2018). It remains controversial and unclear how these 'ramping' to goal signals relate to traditional theories of dopamine in reward learning (Guru et al. 2020). However one question concerns how these proposed goal tracking signals respond to interference from sources of value which are irrelevant for the current goal (for example if alternative competing goals are presented during goal pursuit).

## 7.4 Dorsal ACC and re-deliberation

What mechanisms support re-deliberation of goals? In **chapter 4**, a network of regions - including dACC and dlPFC - were more active when goals were abandoned and when the value of abandoning was higher. This is consistent with many studies implicating these areas in decisions to abandon a behavioural strategy (Kaiser et al. 2021; Trudel et al. 2021), or go and explore other options (Tervo et al. 2021). However the precise contribution of these regions, and dACC in particular, is still debated. One school of thought has proposed the role of this region in exerting cognitive effort and allocating cognitive control (Shenhav et al. 2013). For example, dACC activity often correlates with reaction times (including in our task) and difficulty. Another school of thought has proposed dACC to be at the top of a decision hierarchy in selecting and motivating extended behaviours (Holroyd and Yeung 2012). A related theory has emphasised role of dACC in making particular kinds of longer-term decisions to switch between states or behavioural strategies for which decisions about timing become critical (Kolling and O'Reilly 2018). While our task tends to favour these latter interpretations, it remains an interesting challenge how to reconcile these different perspectives.

One recent proposal building considers a role for dACC in tracking metacognitive variables, for example about the value of deliberation itself (Clairis and Pessiglione 2022). This builds on a framework which considers different levels of evaluation guiding behaviour (and thus reflected in neural activity): an automatic valuation which provides initial input to a meta-cognitive decision about whether to evaluate further; followed by a slower more deliberate valuation which continues until a confidence threshold is met (Lee and Daunizeau 2021). This provides a very different framework for considering the results of our study. Rather than reflecting the value of abandonment persay, could it be that dACC is actually representing the value of re-consideration – the metacognitive value associated with opening up goal deliberation once again?

This account is consistent with the idea of goals as 'defaults' – that is, options which will continue to guide choice unless further deliberation is triggered (Lopez-Persem et al. 2016). It also chimes closely with the philosopher's notion of intentions which I introduced in the very first chapter: action-guiding mental states which are "stable" meaning they will persist over time unless explicitly revised (Bratman 1987). To determine whether dACC is reflecting the value of goal re-selection or the general cost of deliberation within the setting of goal pursuit would require a different paradigm. Specifically, we would need to introduce a greater variety of decisions within the task. For example, harder decisions about how to implement chosen goals which do not affect the value of the goal itself may elicit higher dACC activity in the meta-cognitive account, but not in the state switching account. While some studies have already focussed on the role of dACC in hierarchical decision-making tasks (Ribas-Fernandes et al. 2011; Ribas-Fernandes et al. 2019), further studies could investigate the extent of this distinction between 'goal selection' and 'goal implementation' (OReilly 2020). Evidence that this distinction has real biological relevance in terms of mapping onto neural architectures has been found elsewhere, for example in the discovery of different cell types in superior colliculus supporting orientation during prey selection versus prey pursuit (Hoy et al. 2019).

## 7.5 Planning and medial PFC

Pursuing goals requires planning how to get there. In complex environments, this is a composite process resting on the capacities to simulate transitions, use good heuristics to make efficient choices, and select relevant cues among large quantities of information. The results from **chapter 6** suggest this third element causes particular problems for medial prefrontal lesion patients in planning settings. Building on the earlier ideas of this thesis, regions which play a role in guiding attention in light of current goals are likely to be critical in complex environments where subjects are faced with numerous competing options. This is the other side of the coin to the paradox of goal commitment: selecting goal-relevant information will reduce flexibility to good alternative goals, but may be critical for pursuing chosen goals in busy environments with overwhelming amounts of information to select from.

The exact contributions of medial PFC to model-based planning remain an ongoing topic of research. How does mPFC interact with hippocampal areas during planning, and do their respective roles change over the course of learning as predicted by some accounts of consolidation (Bradfield et al. 2020)? What exactly are the distinct contributions of lateral versus medial prefrontal areas to model-based planning, given the diversity of pre-frontal signals involved in planning tasks (Daw et al. 2011; Smittenaar et al. 2013; Hampton et al. 2006; Bartolo and Averbeck 2020)? And to what extent can vmPFC be considered to coordinate or signal the outcome of computations performed elsewhere, versus constructing models itself? These questions I will leave to future studies and PhD candidates to answer.

# Appendices

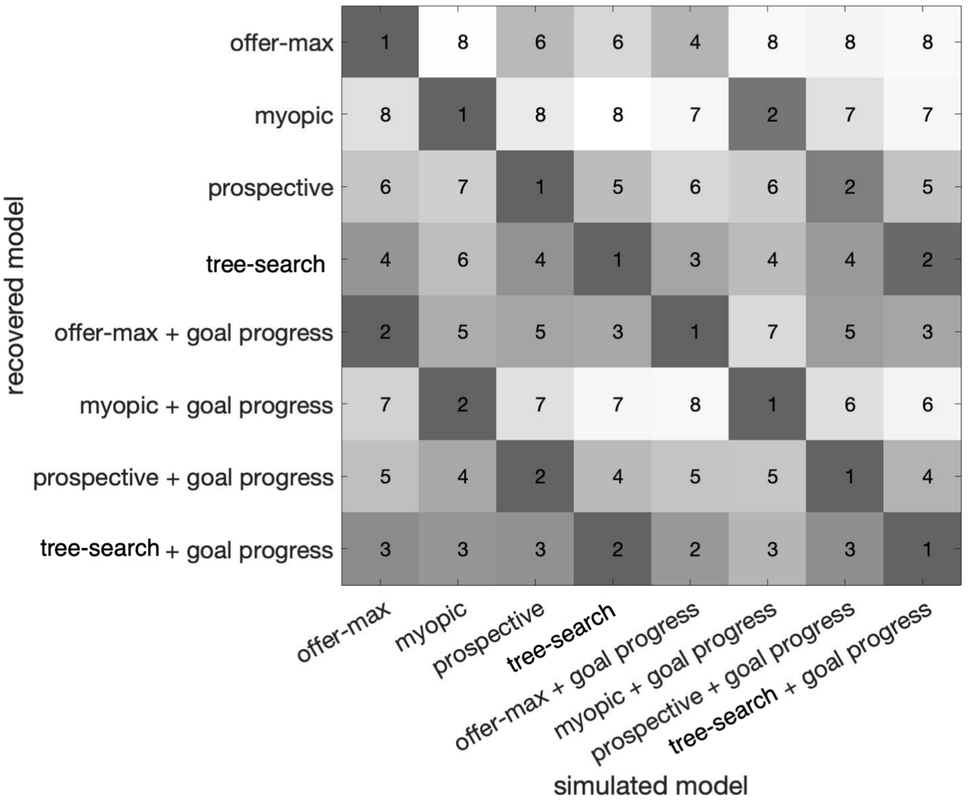
# A

# Supplementary materials for chapter 3

This appendix includes analyses validating the model fitting procedure described in Chapter 3. Model recovery (between competing models), parameter recovery (for the winning tree-search model parameters), and comparison between tree-search and model free parameters are shown below.

## A.1 Model recovery

Fig.A.1 shows the results of a model recovery analysis for the competing models capturing different strategies people could use in the goal pursuit task. The model recovery procedure consisted of simulating data in the empirical range, on the same schedules experienced by participants in our study. To simulate data for each model, we used the empirical parameters from logistic regression models which were fit separately to each participant. A soft-max function was then used to simulate choices from the subjective value (Eq. (3.5)). Subsequently, all models were fitted to all simulated data-sets (using the empirical model fitting procedure described above). To account for stochasticity resulting from the soft-max function, we repeated the simulation process 100 times for each of the 30 participants (resulting in 3000 simulated data-sets per model). We then assessed whether each simulated data set was best fit by its generative model. The averaged confusion matrix is



| | offer-max | myopic | prospective | tree-search | offer-max + goal progress | myopic + goal progress | prospective + goal progress | tree-search + goal progress |
|---|---|---|---|---|---|---|---|---|
| offer-max | 1 | 8 | 6 | 6 | 4 | 8 | 8 | 8 |
| myopic | 8 | 1 | 8 | 8 | 7 | 2 | 7 | 7 |
| prospective | 6 | 7 | 1 | 5 | 6 | 6 | 2 | 5 |
| tree-search | 4 | 6 | 4 | 1 | 3 | 4 | 4 | 2 |
| offer-max + goal progress | 2 | 5 | 5 | 3 | 1 | 7 | 5 | 3 |
| myopic + goal progress | 7 | 2 | 7 | 7 | 8 | 1 | 6 | 6 |
| prospective + goal progress | 5 | 4 | 2 | 4 | 5 | 5 | 1 | 4 |
| tree-search + goal progress | 3 | 3 | 3 | 2 | 2 | 3 | 3 | 1 |

**Figure A.1:** Model recoveries. Confusion matrix resulting from model recovery procedure. Each column corresponds to a model used to simulate the dataset. Each row corresponds to the model used to recover the dataset. Within a column, shading corresponds to the BIC of each competing model relative to the winning model. Lower BICs corresponding to better fits are displayed in darker shades. Numbers indicate the rank of the model in the model comparison per column (where 1 is the winning model, and 8 is the worst fitting model). In all cases, simulated behaviour is best fit by the true generative model. As well as simulating and recovering each model alone, we also confirmed that the models could be recovered with an additional parameter capturing additional effects of goal progress (since this was the best fitting empirical model).

displayed in A.1, showing that the four final simulated models can be correctly identified. Importantly, we find that in the empirical parameter range and across 100 repetitions, there are no cases of more simple models being confused for the empirically best-fitting model (tree-search model).

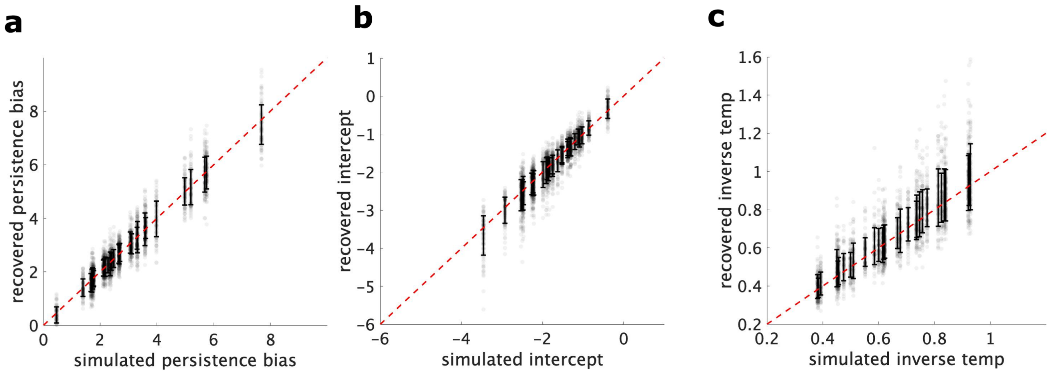


**Figure A.2:** Parameter recovery. Mean recovered parameter (error bars indicate SD) are plotted, with dots showing recovered parameters for individual iterations. Red dotted line indicates the identity line (perfect recovery). Two-sided Pearsons correlations were performed. **(a)** Persistence bias for each individual was defined as their indifference point to abandonment when predicting abandonment choices using the tree search value of abandonment. This is equal to $-beta_0/beta_1$ from the logistic regression (see Persistence bias section in Methods). Recovered persistence biases correlated with the simulated biases with a Pearsons correlation of 0.96 ($p < 0.001$). **(b)** Recovery of the intercept parameter ($beta_0$). The simulated intercepts can be recovered with a Pearsons correlation of 0.92 ($p < 0.001$). **(c)** Recovery of the inverse temperature parameter ($beta_1$). The simulated inverse temperature can be recovered with a Pearsons correlation of 0.84 ($p < 0.001$).

## A.2 Parameter recovery

Fig.A.2 shows parameter recovery for the critical parameters in the goal pursuit tree-search model. Empirical parameters from decision data aggregated across both sessions were used to simulate behaviour (100 iterations per participant yielding 300 simulations total). Parameters were recovered for each simulation by fitting a logistic regression using the same procedure used for the empirical data (described in Chapter 3, Model fitting).

## A.3 Model comparison by trial type

Tree-search model predicted choices best across all trials (Fig.A.3). To determine whether tree-search model was the best description of behaviour across both trial types (abandonment choices and persist choices), below we plot the mean predictive accuracy split by trial type. In both cases, tree-search model predicts

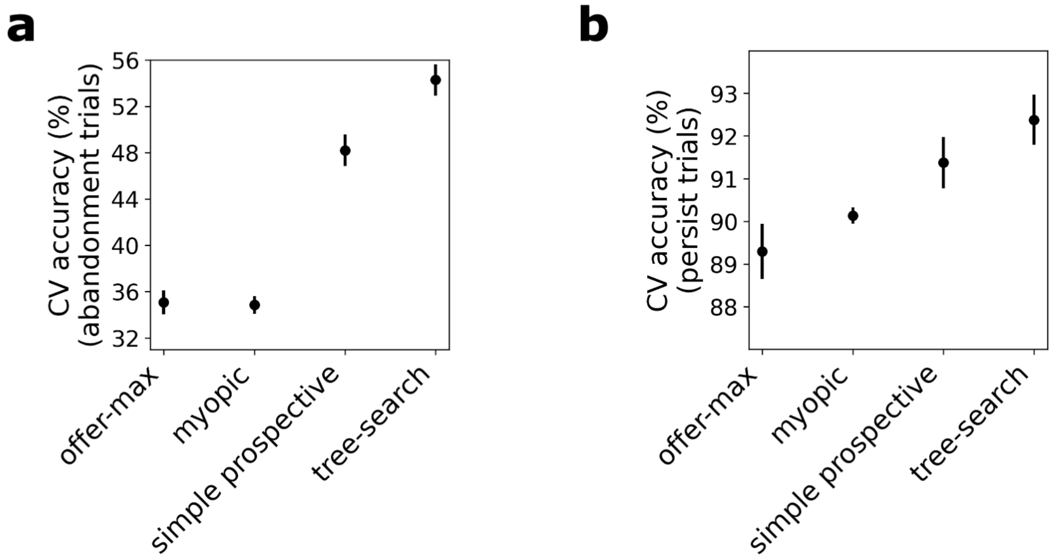


**Figure A.3:** Model comparison by trial type. **(a)** Cross validation accuracy of each model predicting only abandonment trials. A leave one out procedure was used. For each participant, we fit each of the mixed effects model to the choices of all other participants ($n = 29$). Predictive accuracy was computed from the fixed effects on the left out participant. Mean cross validated performance across participants is plotted, with error bars depicting SEM **(b)** Cross validation accuracy of each model predicting only persistence trials. Mean cross validated performance across participants is plotted, with error bars depicting SEM. As for (a), the tree search model describes behaviour best.

choices best. This suggests that rather than capturing just the total number of abandonment choices better than other models (as shown in 3.4), it also captures the *timing* of abandonment choices better. That is, on trials where participants choose to persist, tree-search model also predicts choices to keep persisting better than the other models.

## A.4 Parameter test-retest reliability

Persistence bias and both of the sub-parameters (intercept and beta weight for value) had good test-retest reliability (Fig.A.4). Below are plots of parameters fit separately to the two behavioural testing sessions (300 trials of 'scanner session', 100 trials of 'post-scan session').

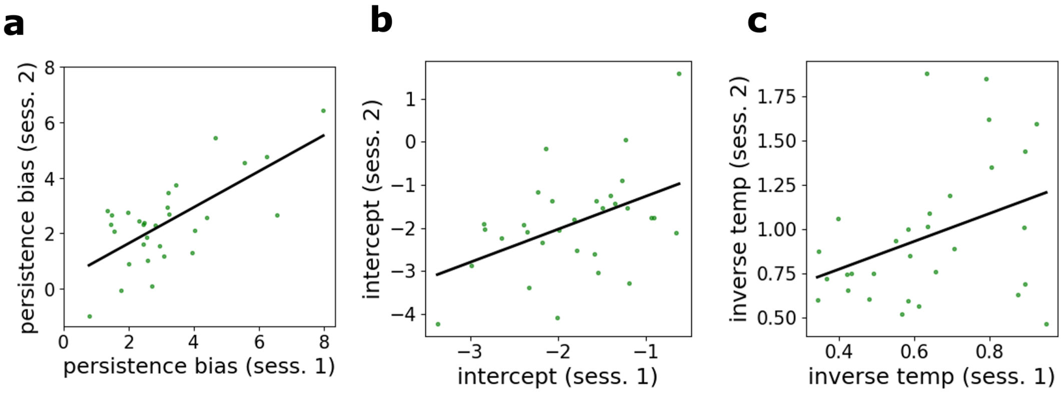


**Figure A.4:** Test retest reliability of parameters across the two behavioural testing sessions. Parameters were separately fitted to the decision task inside the scanner and to the decision task outside the scanner. Two sided Pearsons correlations are reported. All parameters show significant test retest reliability. **(a)** Test retest reliability for persistence biases across the two sessions (Pearsons $r = 0.69$, $p < 0.001$). Persistence bias was defined as the indifference point to abandonment when predicting abandonment choices using the tree search value of abandonment. This is equal to $-beta_0/beta_1$ from the logistic regression (see Persistence bias section in Methods). **(b)** Test retest reliability for the intercept parameter ($beta_0$) across the two sessions (Pearsons $r = 0.46$, $p = 0.010$) **(c)** Test retest reliability of inverse temperature ($beta_1$) across the two sessions (Pearsons $r = 0.38$, $p = 0.040$).

## A.5 Model free persistence metric

In this analysis (A.5), we take the total number of abandonment choices as a model-free metric of individual persistence (whereby the fewer the number of abandonment choices across the study, the greater the persistence). Note that this is an imperfect measure as it will also depend on differences between the schedules completed by each participant. Model-based persistence bias (deviation from the tree-search model) is highly correlated with model-free abandonment. Intuitively, this shows people who make more choices to abandon the current goal (model-free) are less biased towards persisting (model-based). It also shows that our model-based metric is capturing the variability in the raw data.

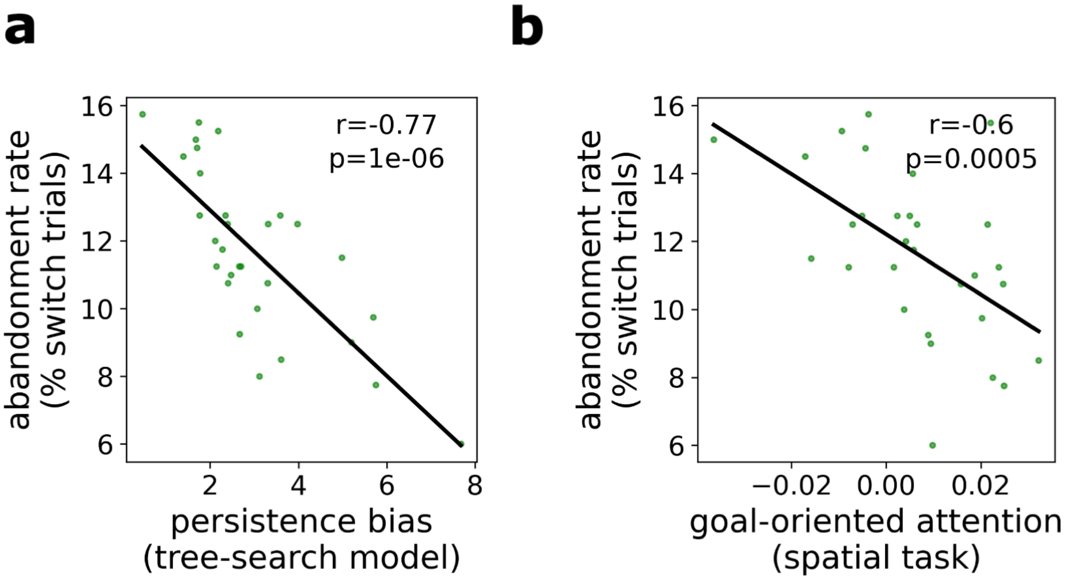


**Figure A.5:** **(a)** Correlation between model-based persistence bias (from the tree-search model) and a model-free metric of persistence (i.e. total number of abandonment choices). **(b)** Correlation between goal-oriented attention in the spatial task, and total number of abandonment choices.

## A.6 Persistence bias is associated with worse performance

In (Fig.A.6), we show that people with higher persistence biases perform worse at the task. Since the tree-search model provides an approximation of optimal choices, peoples' tendency to over-persist beyond this model will lead to worse performance. Across all analyses of this task, we operationalize performance in terms of how quickly people complete goals i.e. the average number of trials to complete a goal, where lower values correspond to better performance (faster goal completion).

## A.7 Model-free replication of analyses

Below (Fig.A.7) we include analyses showing the impact of the current goal offer and best alternative goal offer on behaviour. This analysis uses model free magnitude of offers rather than tree-search value.

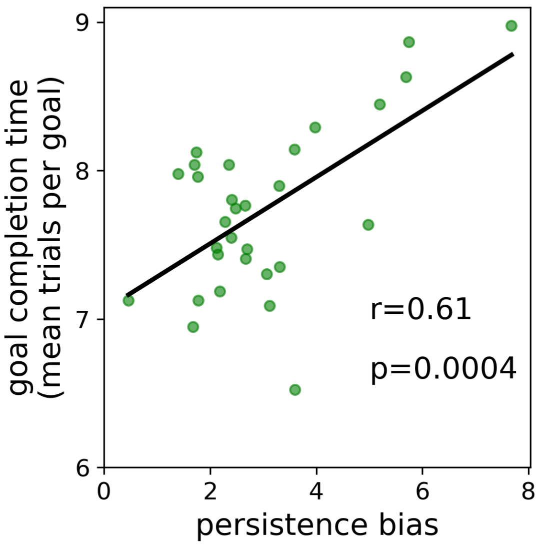


**Figure A.6:** Persistence bias is associated with lower task performance. Performance corresponds to the average goal completion time, in terms of trials. Lower values mean goals are completed faster (better performance). Persistence bias captures how much individuals persist with the goal more than the tree-search model. Statistic shows pearson correlation.

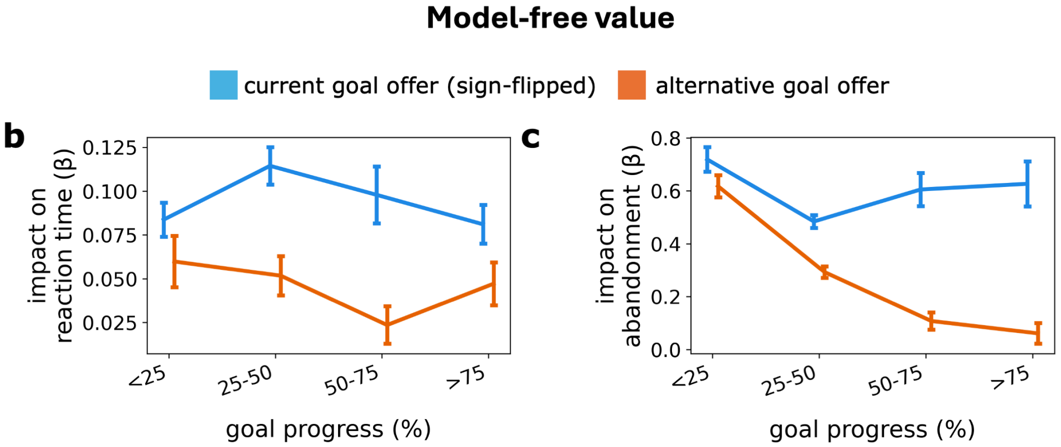


**Figure A.7:** Illustration of the impact of raw offer magnitudes on reaction times and choices, over the course of goal pursuit. Data is binned by quartile of goal progress (proportion of the net filled).

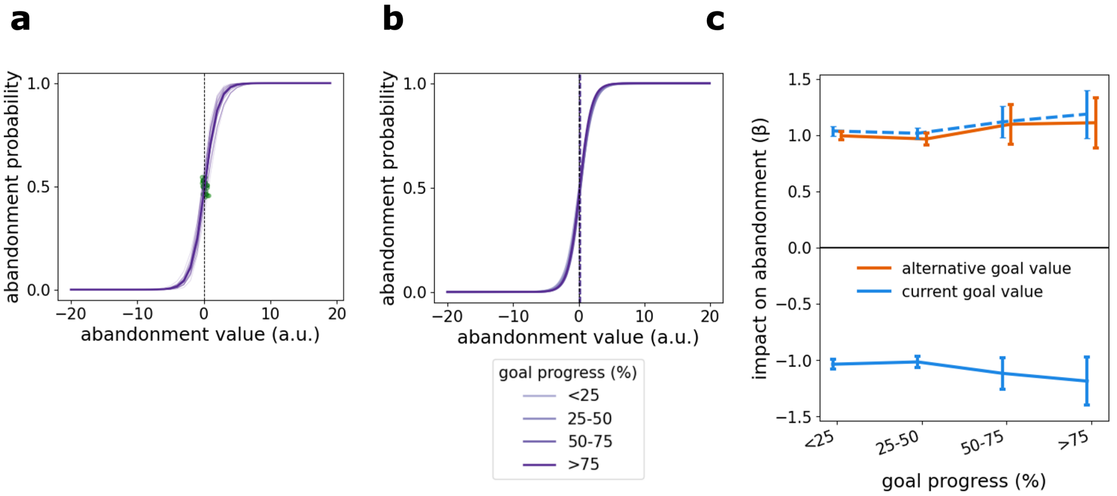


**Figure A.8:** **(a)** Individual fits to simulated data sets show persistence biases of zero, demonstrating that the empirically found biases are not an inherent feature of the schedules used. **(b)** Quartile fits demonstrating that we accurately recover no difference in persistence biases across quartiles when simulating with the normative model. **(c)** Effects of temptation and frustration on abandonment choices. In the normative model simulations, we accurately recover no difference in how sensitivity to these two sources of value develop across goal pursuit.

## A.8 Simulations of tree-search behaviour

In this analysis (Fig.A.8), we perform the empirical behavioural analyses on simulated data from the normative tree-search model, demonstrating the empirical biases are not an artefact of the testing schedules. Here, behaviour from the tree-search model is simulated on the 30 participant schedules. In the figure below, simulated results in (a) can be compared to empirical results Fig.3.7, simulated results in (b) can be compared to empirical results Fig.3.8, and simulated results in (c) can be compared to empirical results Fig.3.8.

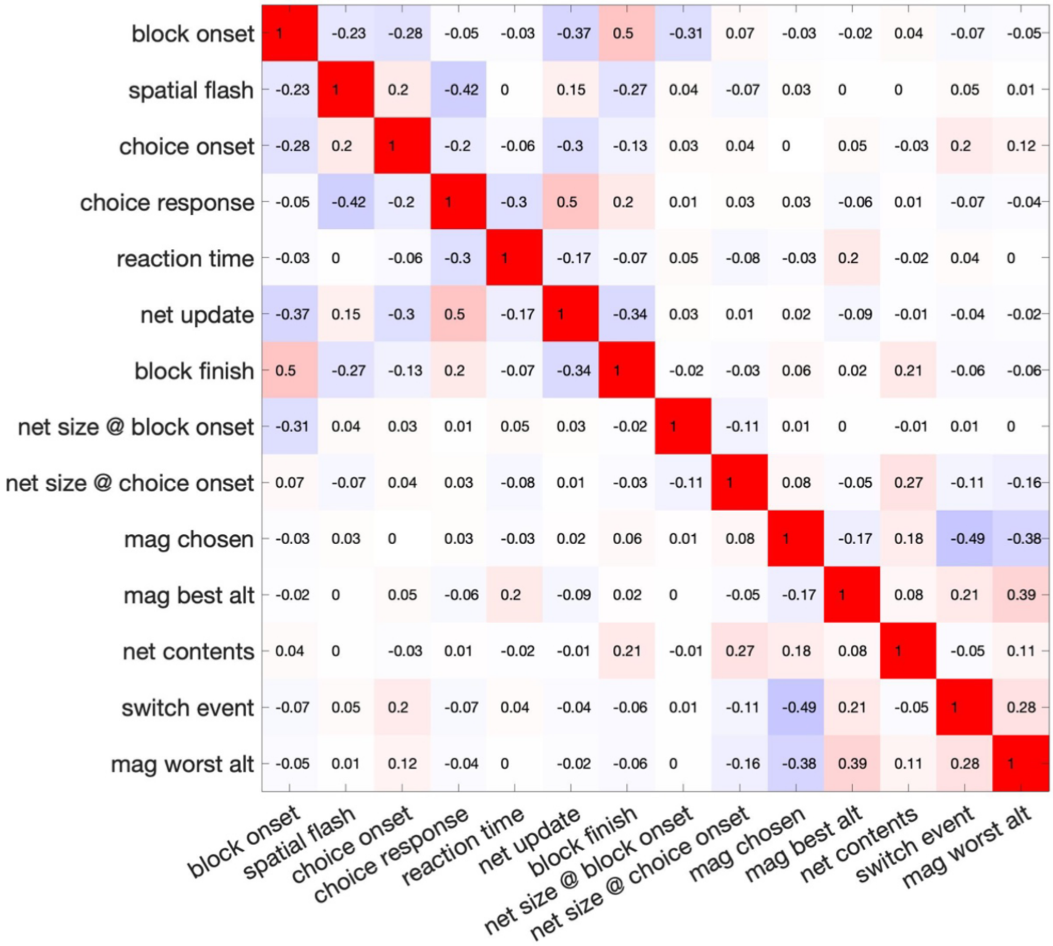
# B

# Supplementary materials for chapter 4

## B.1 Whole-brain analysis with basic (model-free) regressors

We first performed an analysis with no regressors derived from the modelling, as a sanity check. Six regressors of interest were included in the main GLM, predicting BOLD activity at the onset of the decision period (all modelled as stick functions). These regressors included the three offers (current goal, best alternative, worst alternative), the accumulated goods (contents of the net), the size of the net, and a binary regressor capturing whether the participant chose to persist or abandon. See Fig.B.1 for correlation between regressors.

The results of this cluster-corrected analysis can be seen in Fig.B.2. The analysis revealed activity positively correlating with the best alternative offer and negatively correlating with current goal offers in the dorsolateral PFC (dlPFC), dACC, and insula. In contrast, vmPFC, medial frontal pole and striatum showed positive correlation with the current goal offer. In addition, we found activity negatively correlating with the best alternative offer in dorsal medial frontal pole and in a lateral orbital frontal area (Brodmann area 47). We found no neural areas correlating with the worst alternative, consistent with our behavioural findings that this had no influence on reaction times or choices. Finally, we found widespread



| | block onset | spatial flash | choice onset | choice response | reaction time | net update | block finish | net size @ block onset | net size @ choice onset | mag chosen | mag best alt | net contents | switch event | mag worst alt |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| block onset | 1 | -0.23 | -0.28 | -0.05 | -0.03 | -0.37 | 0.5 | -0.31 | 0.07 | -0.03 | -0.02 | 0.04 | -0.07 | -0.05 |
| spatial flash | -0.23 | 1 | 0.2 | -0.42 | 0 | 0.15 | -0.27 | 0.04 | -0.07 | 0.03 | 0 | 0 | 0.05 | 0.01 |
| choice onset | -0.28 | 0.2 | 1 | -0.2 | -0.06 | -0.3 | -0.13 | 0.03 | 0.04 | 0 | 0.05 | -0.03 | 0.2 | 0.12 |
| choice response | -0.05 | -0.42 | -0.2 | 1 | -0.3 | 0.5 | 0.2 | 0.01 | 0.03 | 0.03 | -0.06 | 0.01 | -0.07 | -0.04 |
| reaction time | -0.03 | 0 | -0.06 | -0.3 | 1 | -0.17 | -0.07 | 0.05 | -0.08 | -0.03 | 0.2 | -0.02 | 0.04 | 0 |
| net update | -0.37 | 0.15 | -0.3 | 0.5 | -0.17 | 1 | -0.34 | 0.03 | 0.01 | 0.02 | -0.09 | -0.01 | -0.04 | -0.02 |
| block finish | 0.5 | -0.27 | -0.13 | 0.2 | -0.07 | -0.34 | 1 | -0.02 | -0.03 | 0.06 | 0.02 | 0.21 | -0.06 | -0.06 |
| net size @ block onset | -0.31 | 0.04 | 0.03 | 0.01 | 0.05 | 0.03 | -0.02 | 1 | -0.11 | 0.01 | 0 | -0.01 | 0.01 | 0 |
| net size @ choice onset | 0.07 | -0.07 | 0.04 | 0.03 | -0.08 | 0.01 | -0.03 | -0.11 | 1 | 0.08 | -0.05 | 0.27 | -0.11 | -0.16 |
| mag chosen | -0.03 | 0.03 | 0 | 0.03 | -0.03 | 0.02 | 0.06 | 0.01 | 0.08 | 1 | -0.17 | 0.18 | -0.49 | -0.38 |
| mag best alt | -0.02 | 0 | 0.05 | -0.06 | 0.2 | -0.09 | 0.02 | 0 | -0.05 | -0.17 | 1 | 0.08 | 0.21 | 0.39 |
| net contents | 0.04 | 0 | -0.03 | 0.01 | -0.02 | -0.01 | 0.21 | -0.01 | 0.27 | 0.18 | 0.08 | 1 | -0.05 | 0.11 |
| switch event | -0.07 | 0.05 | 0.2 | -0.07 | 0.04 | -0.04 | -0.06 | 0.01 | -0.11 | -0.49 | 0.21 | -0.05 | 1 | 0.28 |
| mag worst alt | -0.05 | 0.01 | 0.12 | -0.04 | 0 | -0.02 | -0.06 | 0 | -0.16 | -0.38 | 0.39 | 0.11 | 0.28 | 1 |

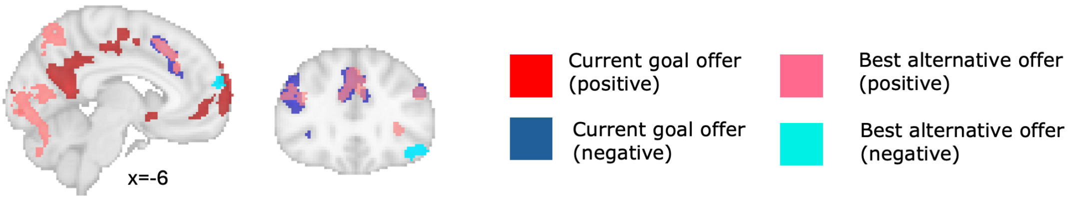
**Figure B.1:** Correlation between regressors in the model-free whole-brain analysis. Note the events between decisions (stick regressors) are also displayed in the correlation matrix.

activity positively correlating with accumulated goods including in dorsal and rostral ACC, 32pl, striatum, hippocampus, and vmPFC. No frontal regions correlated negatively with accumulated goods.

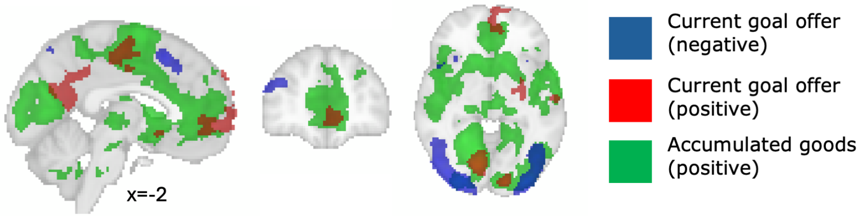
## B.2 fMRI peak activity table

Peaks of activity for the whole-brain analyses (using tree-search value). See Fig.B.4.

**(a)** Overlapping opposing activity for current goal and alternative goal offers in dACC, frontal pole, and dlPFC



**(b)** Distinct areas in mPFC track accumulated goods



**(c)** Reaction times

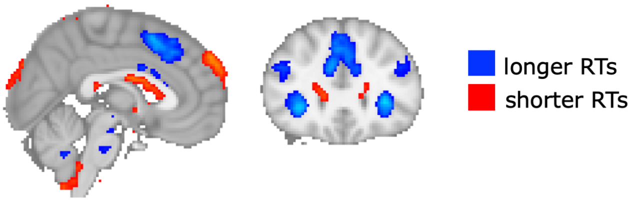


**Figure B.2:** Model-free whole brain analysis activity. **(a)** Activity correlating with the best alternative and current goal offers. **(b)** Activity correlating with accumulated goods in the net (shown alongside activity for the current goal offer). **(c)** Activity correlating with reaction times.

## B.3 Time courses by decision outcome

Below we include illustrations of the time course of value-related activity in our three ROIs on persistence versus abandonment trials (Fig.B.5)

## B.4 Baseline vmPFC control analyses

We performed two control analyses for our finding that vmPFC baseline activity predicted persistence biases.

First, we controlled for the location, by performing the identical analysis in our two other neural ROIs (striatum and dACC). Fig.B.6a,b shows the relationship between behavioural persistence and baseline goal-related activity in these control ROIs. The fact that neither of the other ROIs (which also contained

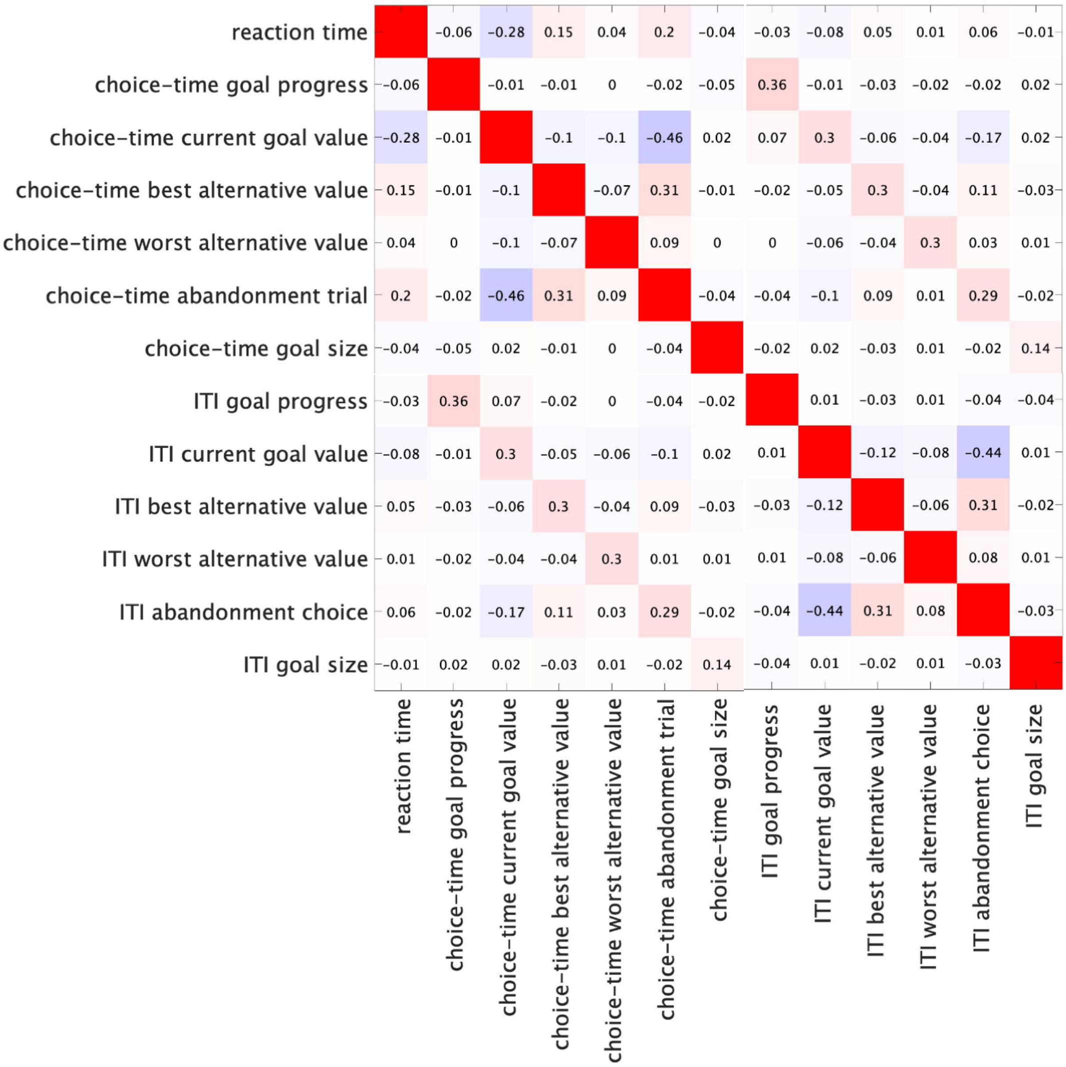


**Figure B.3:** Correlation between regressors in the inter-trial analysis.

value-related activity in our whole-brain analysis) predicts behaviour suggests it is specific to vmPFC.

Second, we investigated whether decision-related activity in vmPFC correlated with our behavioural metrics (Fig.B.6c). We quantified the decision-related activity in vmPFC in response to goal progress using the HRF function. We took the fitted beta coefficients for goal progress at each time-point (from stimulus onset) and multiplied these by the double gamma HRF function, and summed the products to produce a coefficient for each participant. Decision-related vmPFC activity did not significantly predict behavioural persistence (Fig.B.6c).

| Contrast | Region | Peak coordinates (x,y,z in mm MNI space) | Z Value |
|---|---|---|---|
| **Whole-brain fMRI analysis time-locked to decision onset.** | | | |
| **Persistence value**: current goal value–best alternative. value | Ventromedial prefrontal cortex (area 11m) | -2, 48, -8* | 5.2 |
| | Ventral striatum | 8, 8, -10* | 5.16 |
| | Lateral frontal pole | -38, 30, -16 | 5.26 |
| | Supplementary motor cortex | 0, -6, 58 | 5.8 |
| | Parietal operculum cortex | Left:- 58, -38, 26; Right: 54, -28, 24 | Left: 5.98;<br>Right: 5.62 |
| | Intracalcarine cortex | 8, -84, 4 | 6.94 |
| | Precuneous cortex | -6, -54, 12 | 5.55 |
| | Superior parietal lobule | 28, -40, 64 | 5.04 |
| | Parahippocampal gyrus | -24, -36, -14 | 4.8 |
| | Precentral gyrus | Left: -60, 0, 30 ; Right: 42, -18, 62; | Left: 4.69;<br>Right: 4.6 |
| | Cerebellum area VIIa | 14, -64, -46 | 4.36 |
| | Cerebral crus | 26, -80, -32 | 4.56 |
| **Abandonment value**: best alternative value–current goal value | Dorsolateral prefrontal cortex | Left: -46, 44, 16;<br>Right: 48, 36, 26 | Left: 5.74;<br>Right: 6.27 |
| | preSMA extending into dorsal ACC | -4, 20, 46<br>dACC subpeak:<br>8, 28, 30* | 6.27<br>dACC subpeak:<br>5.4 |
| | Insular cortex | Left: -32, 20, 4;<br>Right: 36, 18, 6 | Left: 5.77;<br>Right: 5.49 |
| | Lateral frontal pole | 42, 44, 0 | 4.41 |
| | Inferior frontal gyrus / precentral gyrus | -44, 8, -32 | 5.74; |
| | Supramarginal gyrus | 50, -48, 48 | 6.21 |
| | Superior frontal gyrus | -26, 2, 62 | 4.76 |
| | Intracalcarine cortex | -8, -74, 8 | 6.03 |
| | Cerebellum area VI | -8, -74, -22 | 5.53 |
| | Cerebellum area VIIb | -34, -68, -56 | 4.93 |
| **Persistence choice:** Persistence trials–abandonment trials | Area 11m / Frontal medial pole | 0, 64, -8 | 4.95 |
| | Area 25 | 0, 20, -8 | 4.45 |
| | Orbitofrontal cortex (area 47) | -42, 28, -10 | 4.48 |
| | Occipital Pole | Left: -22,-104, -2;<br>Right: 4, -88, 36 | Left: 4.3;<br>Right: 4.6 |
| | Lingual gyrus | 10, -90, -18 | 4.49 |
| | Superior parietal lobule | -22, -46, 62 | 4.46 |
| | Central opercular cortex | Left: -56, 6, 0; Right: 62, 2, 8 | Left: 4.85;<br>Right: 5.54 |
| | Precentral gyrus | 26, -22, 74 | 4.62 |
| | Lateral ventricle | Left: -20, -48, 16; Right: 36, -42, 4 | Left: 4.45;<br>Right: 4.63 |
| | Cerebral crus | 18, -80, -40 | 3.95 |
| | Inferior temporal gyrus | 54, -58, -24 | 4.33 |
| | Lateral occipital cortex | -48, -76, 34 | 4.15 |
| **Abandonment choice:**<br>Abandonment choice– persistence choice | Caudate | -12, 16, 4 | 5.57 |
| | preSMA extending into dorsal ACC | 6, 20, 44 | Right: 5.69 |
| | Dorsolateral frontal pole | -42, 42, 22 | 4.4 |
| | Lateral frontal pole (left) | -22, 62, 2 | 4.52 |
| | Posterior cingulate gyrus | -4, -20, 30 | 4.83 |
| | Insular cortex extending into frontal operculum cortex | Left: -32, 18, -6;<br>Right: 36, 24, 4 | Left: 7.04;<br>Right: 6.64 |
| | Occipital pole | -12, -94, -4 | 5.58 |
| | Superior parietal lobule | -38, -52, 50 | 4.41 |
| | Occipital fusiform gyrus | 24, -74, -4 | 4.81 |
| | Inferior frontal gyrus / precentral gyrus | -46, 2, 40 | 5.38 |
| | Superior frontal gyrus | -22, -2, 58 | 4.61 |
| | Lateral occipital cortex | -26, -84, 18 | 5.14 |
| **Goal progress** (i.e. proportion of net completed) | Non-region specific cluster (74958 voxels), encompassing areas stretching from lateral occipital cortex, temporal gyrus, insular cortex, striatum, cingulate cortex, and medial prefrontal areas | 10, -92, 6 | 8.1 |
| | Middle / superior frontal gyrus | -28, 34, 48 | 5.38 |
| | Brain stem | 2, -30, -42 | 4.6 |
| **Negative goal progress** (i.e. activity related to proportion of net *remaining* to be filled) | Lateral occipital cortex | -42, -64, 50 | 6.39 |
| Contrast | Region | Peak coordinates (x,y,z in mm MNI space) | Z Value |
| **Whole-brain analysis time-locked to the inter-trial fixation cross.** | | | |
| **Inter-trial goal progress** (i.e. goal progress time-locked to the ITI fixation cross) | Frontal medial pole (Area 10 stretching to area 11 and area 14) | 4, 58, -6* | 5.17 |
| | Right hippocampus | 26, -20, -14 | 4.69 |
| | Temporal pole | Left: -40, 22, -30; Right: 48, 18, -30 | Left: 4.34;<br>Right: 4.47 |
| | Area 8m | -20, 36, 48 | 4.73 |
| | Precuneus cortex stretching to posterior cingulate gyrus | -6, -56, 22 | 4.48 |
| | Middle temporal gyrus | -64, -12, -14 | 4.46 |
| | Caudate | 14, 20, 6 | 4.08 |
| | Postcentral gyrus | 36, -24, 56 and second cluster at 54, -22, 54 | 4.45, 4.58 |
| Family-wise error cluster corrected, $z > 2.3$, $p < 0.05$<br>*Coordinates used for region of interest analyses | | | |

**Figure B.4:** fMRI peaks from main whole-brain analysis. Stars indicate ROI peaks.

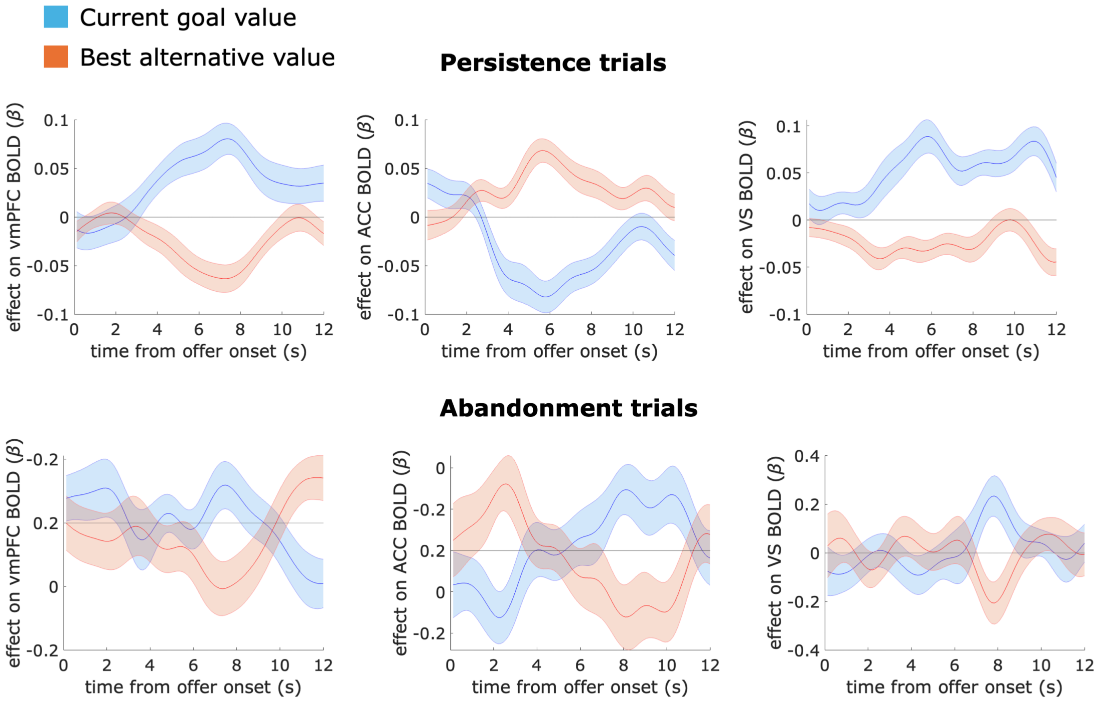


**Figure B.5:** Illustration of neural time-courses split by trial type. The top panel shows value-related activity in vmPFC, dACC and ventral striatum on trials where the participant chose to persist. The bottom panel shows the same except for trials where the participant chose to abandon the current goal for the best alternative. Speculatively, the dACC appears to show a reversal in its value-related signal over the time-course of abandonment trials. However, we only attained a small number of abandonment trials in our study as the task was not designed to answer this question. This observation is therefore speculative and more data would be needed to establish the activity profile on these trials.

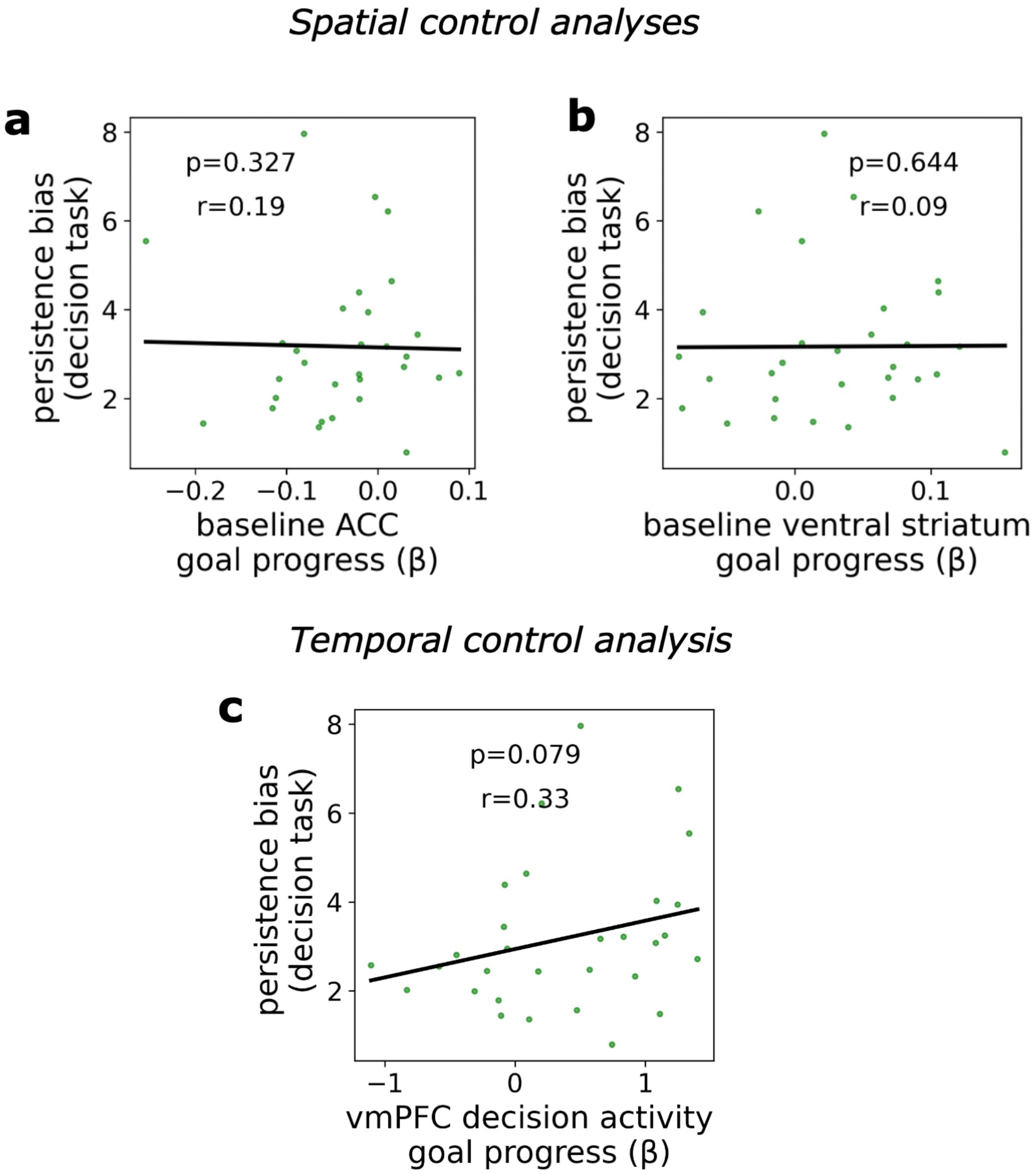


**Figure B.6:** Control analyses for baseline vmPFC finding. **(a,b)** We found no relationship between behaviour and baseline goal-related activity in the two other ROIs (ACC or striatum). Statistics refer to spearman correlations. Dots show individual data points, line shows regression. **(c)** We found no significant relationship between behaviour and decision-related activity in the vmPFC ROI.

# C
# Supplementary materials for chapter 5

## C.1 Model fitting in patients and age-matched controls

Fig.C.1 shows a model comparison between the four behavioural models when fit to all lesion patient data and age-matched control data. In both the lesion patient group and age-matched controls, the tree-search model was the best description of behaviour in terms of predictive accuracy and when quantified as the most frequent across participants.

## C.2 Lesion maps for vmPFC patients

Fig.C.2 shows lesion maps for the four patients with damage in the vmPFC ROI identified by the fMRI study.

## C.3 Performance in voxel-wise analysis cohort

Fig.C.3 shows a comparison of performance between patients with lower persistence identified in the voxelwise analysis, to healthy controls and other patients. This is the same as Fig.5.3 except the fifth participant from the voxelwise analysis is also included.

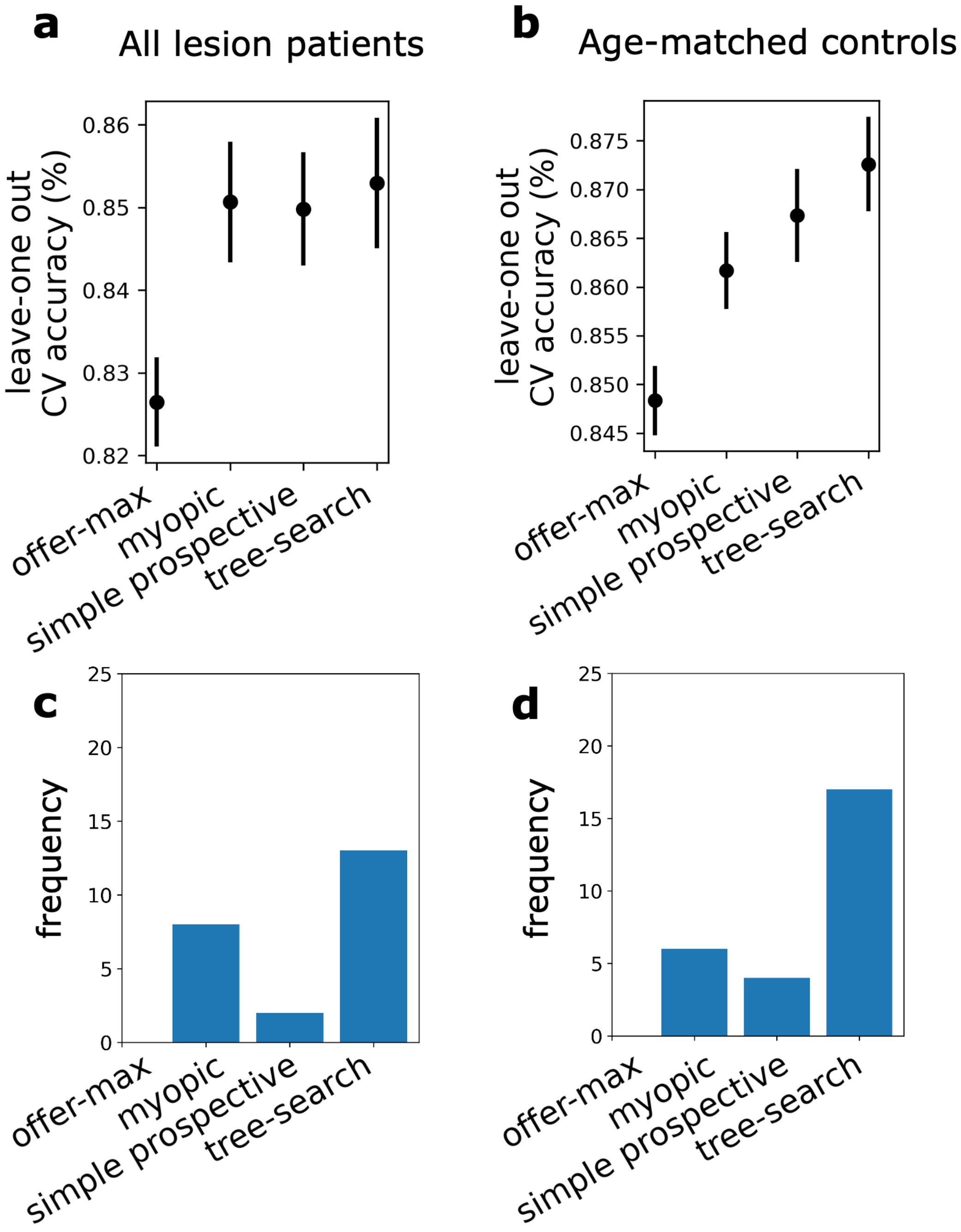


**Figure C.1:** Comparison of behavioural models among all lesion patients, and among age-matched healthy control subjects.

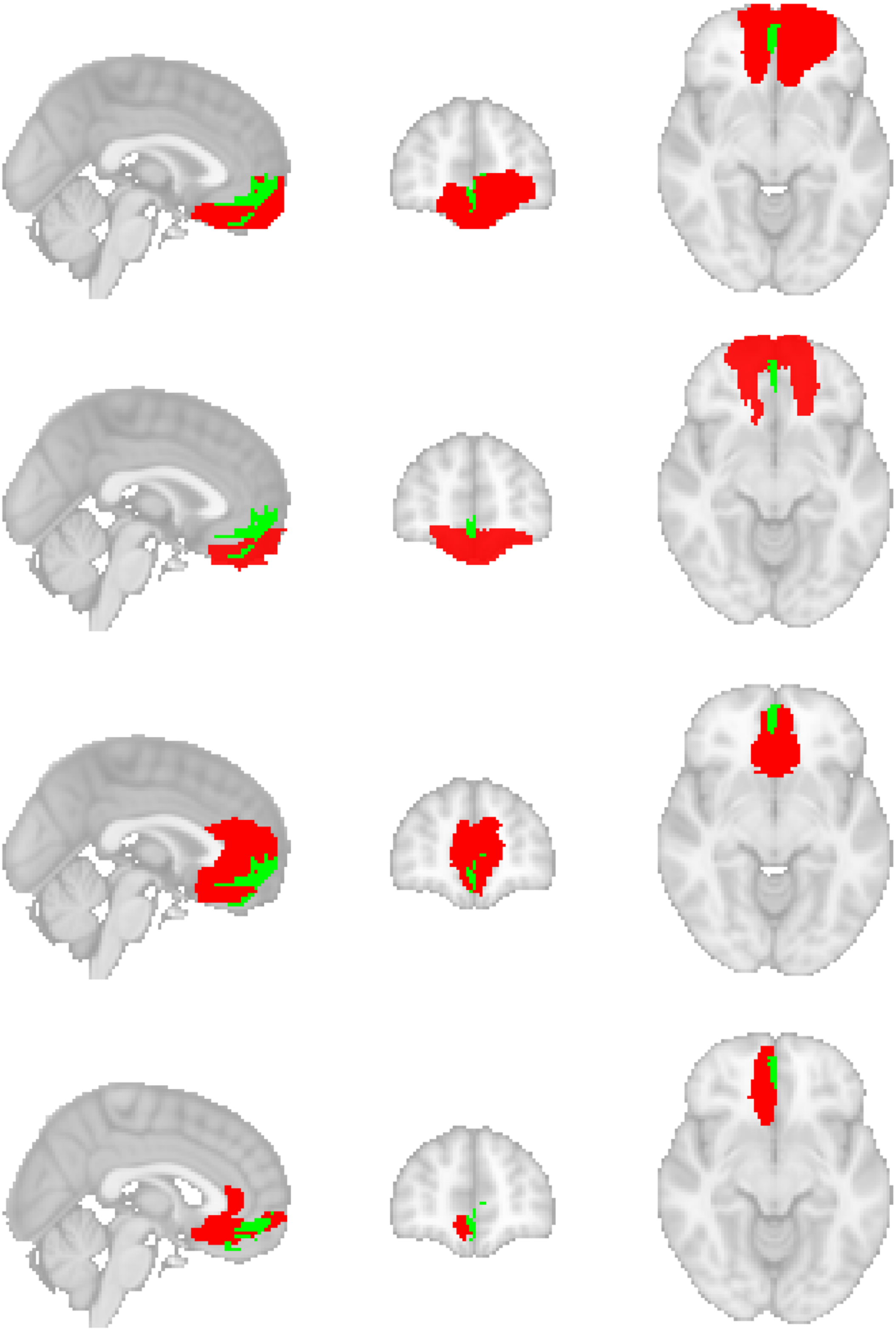


**Figure C.2:** Lesion maps for the four patients with damage in the fMRI vmPFC ROI. Red shows damaged area. Green shows map of area independently identified by voxel-wise analysis (areas where damage predicts reduced persistence with the goal).

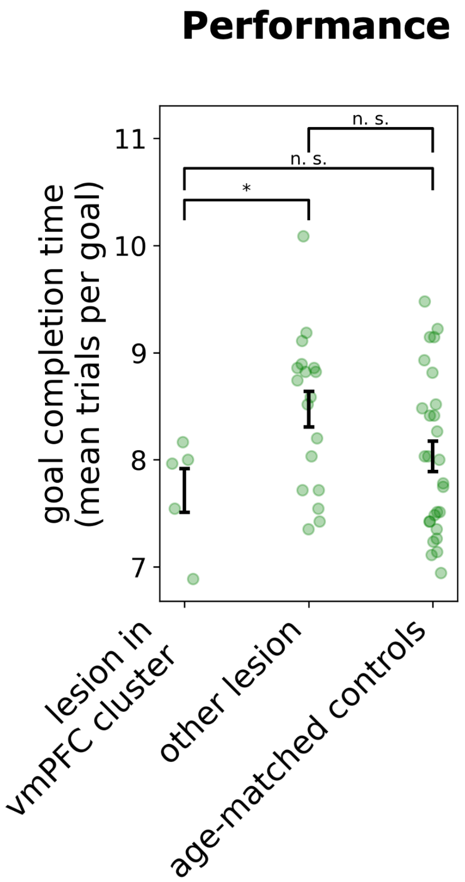


**Figure C.3:** Comparison of performance in patients with damage to vmPFC area identified in voxelwise analysis.

## C.4 Reaction times in spatial attention task

Fig.C.4 shows reaction times in the spatial attention task for all our cohorts (fMRI cohort, age-matched controls, lesion controls, and vmPFC lesion patients).

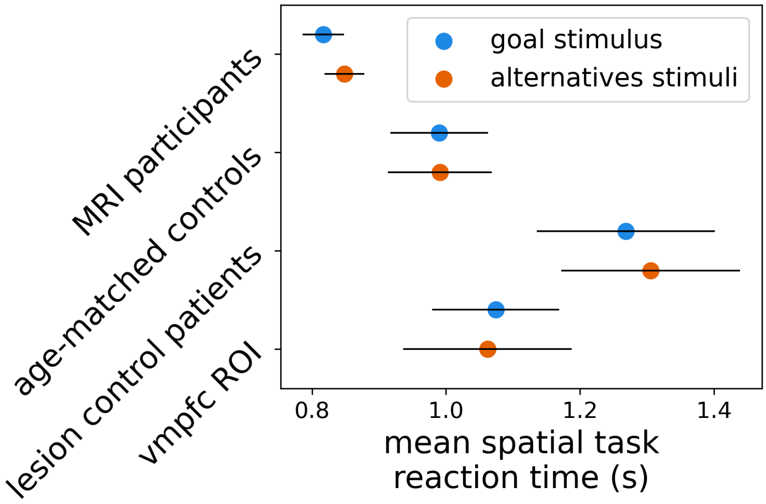


**Figure C.4:** Reaction times in spatial attention task across all cohorts. Blue dots show mean RT for reporting the current goal stimulus location. Orange dots show mean RT for reporting the two alternative stimuli locations. Groups correspond to the fMRI cohort (Chapter 3 and 4), the age-matched controls in the lesion study (Chapter 5), the lesion control patients, and the patients with damage to the vmPFC ROI. Error bars show s.e.m. across each group.

# D

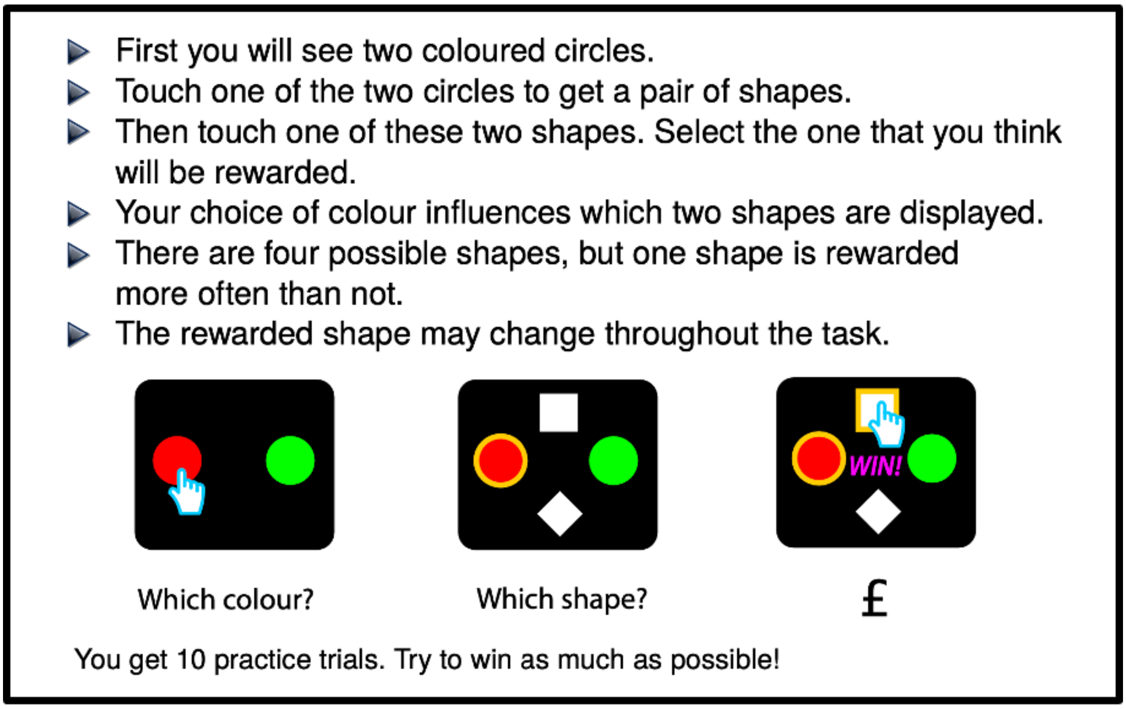
# Supplementary materials for chapter 6

## D.1 Two-step task instructions

Task instructions can be seen in Fig.D.1. Participants were told about contingencies between the two steps (“Your choice of colour influences which two shapes are displayed”), and about the basic reward structure (“There are four possible shapes, but one shape is rewarded more often than not. The rewarded shape may change throughout the task”). However, unlike in some studies e.g. Castro-Rodrigues et al. (2022), they were not explicitly informed of the transition probabilities.

## D.2 RL Model Fitting

A Bayesian hierarchical modelling framework was used to fit the reinforcement learning models to behaviour, allowing us to pool data across participants to improve individual parameter estimates. To aid model fitting in stan, we used reparameterization to sample parameters as centred standard normal distributions (which facilitate gradient calculations in stan), which were then transformed into the appropriate prior distributions. Group-level parameters for means and variances were sampled from the following distributions:



- First you will see two coloured circles.
- Touch one of the two circles to get a pair of shapes.
- Then touch one of these two shapes. Select the one that you think will be rewarded.
- Your choice of colour influences which two shapes are displayed.
- There are four possible shapes, but one shape is rewarded more often than not.
- The rewarded shape may change throughout the task.

Which colour?

Which shape?

£

You get 10 practice trials. Try to win as much as possible!

**Figure D.1:** Two-step task instructions.

$$\mu_{\beta_{mf}}, \mu_{\beta_{mb}}, \mu_{\beta_{step-two}}, \mu_{\alpha}, \mu_{\lambda} \sim \mathcal{N}(0,\, 1) \tag{D.1}$$

$$\mu_p \sim \mathcal{N}(0,\, 10) \tag{D.2}$$

$$\sigma^2_{\beta_{mf}}, \sigma^2_{\beta_{mb}}, \sigma^2_{\beta_{step-two}}, \sigma^2_{\alpha}, \sigma^2_{\lambda}, \sigma^2_{p} \sim ln\mathcal{N}(0,\, 1) \tag{D.3}$$

Group-level variance was defined as a lognormal distribution to ensure only positive values.

For all parameters with the exception of p (for which the appropriate prior distribution is a centred normal distribution) parameter transformations were used to enforce constraints and impose uniform prior distributions across the appropriate ranges. Parameters were transformed using an approximation of the phi function (i.e. normal cumulative density function), which leads to a uniform prior over the constrained range when applying the cumulative density function to a normal distribution:

$$\alpha = \Phi(\mu_\alpha) \tag{D.4}$$

$$\alpha = \Phi(\mu_\lambda) \tag{D.5}$$

$$\alpha = \Phi(\mu_{\beta_{mb}}) \tag{D.6}$$

$$\alpha = \Phi(\mu_{\beta mf}) \tag{D.7}$$

$$\alpha = \Phi(\mu_{\beta step-two}) \tag{D.8}$$

This constrains $\alpha$ and $\lambda$ to have a uniform prior on (0,1), and constrains $\beta_{mb}$, $\beta_{mf}$, and $\beta_{step-two}$ to have a uniform prior on (0,10). The individual-level parameters for the $i_{th}$ participant ($p^i$, $\alpha^i$,$\lambda^i$,$\beta^i_{mb}$,$\beta^i_{mf}$,$\beta^i_{step-two}$) were given a normal distribution with the mean as the prior on group mean, and variance as the prior on group variance. The individual-level parameters were then also transformed using the phi function to enforce constraints. For example, in the case of $\alpha^i$:

$$\alpha^i \sim \Phi\mathcal{N}(\mu_\alpha,\, \sigma^2_\alpha) \tag{D.9}$$

Models were coded in the Stan modelling language (Carpenter et al. 2017), and fitted to each dataset using the Cmdstanpy interface. Datasets were fit with 4 chains, using 1000 samples per chains (warmup 500). R-hat values $\leq 1.1$ indicated convergence across all parameters. Following previous studies (Decker et al. 2016; Potter et al. 2017), we did not include the first 9 choice trials in the analysis.

## D.3 RL Model Validation

While the reinforcement learning model in this paper has been widely used in the literature (Daw et al. 2011; Decker et al. 2016; Potter et al. 2017), this dataset was collected using an adapted version of the two-step task with stationary rather than drifting reward probabilities. For this reason, it was important to check that critical parameter values could still be recovered in the new task variant, where reward probabilities were stationary with abrupt shifts.

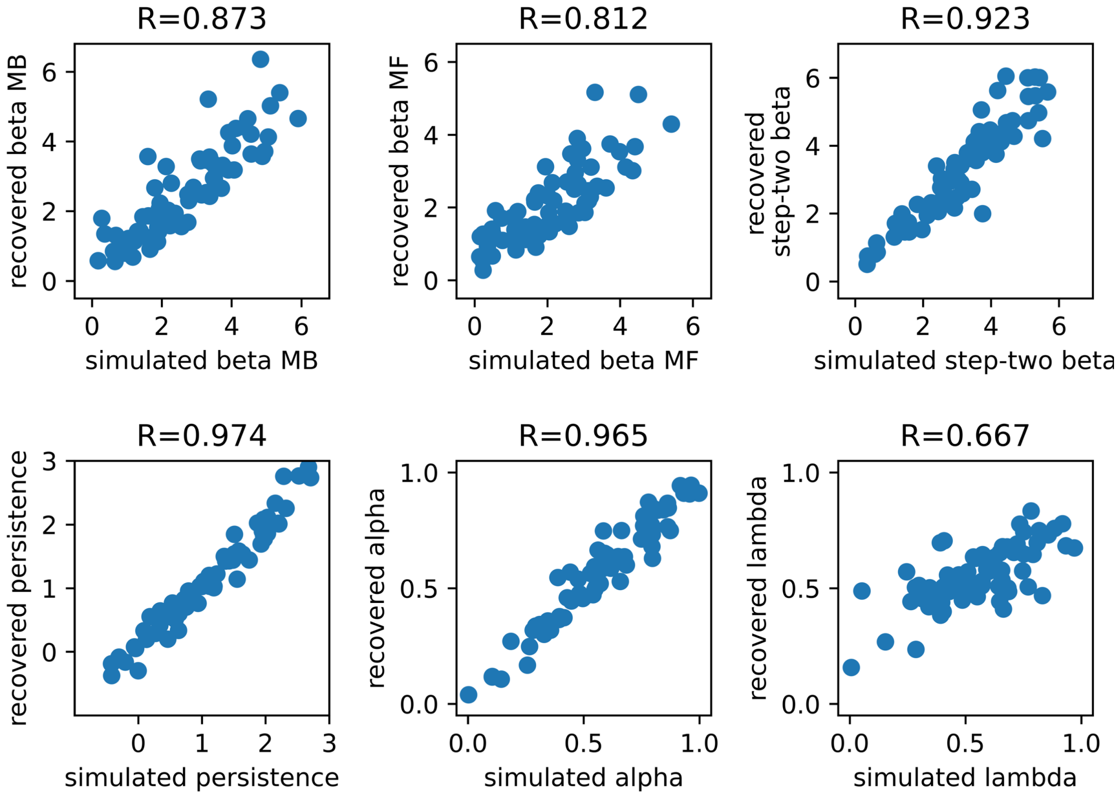


**Figure D.2:** Parameter recoverability. The six reinforcement learning parameters were simulated using the range reported in Decker et al. 2016. Dots show individual simulated data sets for the 70 empirical schedules. Titles refer to the Pearsons R between simulated and recovered parameters in each case.

Behaviour was simulated for 70 participants using the actual transitions (rare/common) and reward probability structure experienced by the 70 participants in our dataset. For each simulated participant, the six parameters were sampled from a normal distribution based on the range reported in Decker et al. (2016) (mean and standard deviation from this range). The participant data was then fit using the same procedure described in 'Model Fitting'. The parameters were recoverable in the adapted version of the task (Fig.D.2).

## D.4 Four-in-a-row lapse rate parameter

Fig. D.3 shows a sub-parameter of the Four-in-a-row model, namely 'Lapse Rate', as a function fo group. We include this because this sub-parameter, like feature

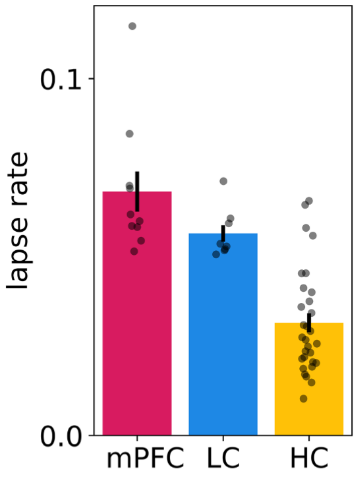


**Figure D.3:** Four-in-a-row lapse rate. Lapse rate is one of the sub-parameters in the planning model, which corresponds to the the probability of choosing a completely random move. This provides a control for a different source of noise to feature drop rate (which is the probability of missing a particular feature). While lapse rate was generally higher for lesion patients than healthy age-matched controls, the effect of lesion location (lesion in mPFC versus sparing mPFC) was not significant.

drop rate, contributes a source of noise. Lapse rate captures the probability of making a random move on any trial (not influenced by the value of the board state or planning algorithm). As shown in Fig. D.3, patients with lesion damage in general make more lapses than healthy age-matched controls (mean lapse rate for lesion patients=0.063, std=0.015; mean lapse rate for age-matched controls=0.037, std=0.015; two-sided Mann-Whitney $U$ test comparing lapse rates: $n_1$=18, $n_2$=30, $U$=500, $p < 0.001$). However, unlike for feature drop rate, mPFC lesion patients did not have higher lapse rate than other lesion patients (mean lapse rate for lesion patients=0.068, std=0.018; mean lapse rate for age-matched controls=0.057, std=0.007; two-sided Mann-Whitney $U$ test comparing lapse rates: $n_1$=10, $n_2$=8, $U$=59, $p = 0.101$, $n.s.$)

## D.5 Feature drop rate illustration

Below we include analyses reprinted with permission from van Opheusden et al. (2023), showing cases of when a model with feature drop rate captures participant

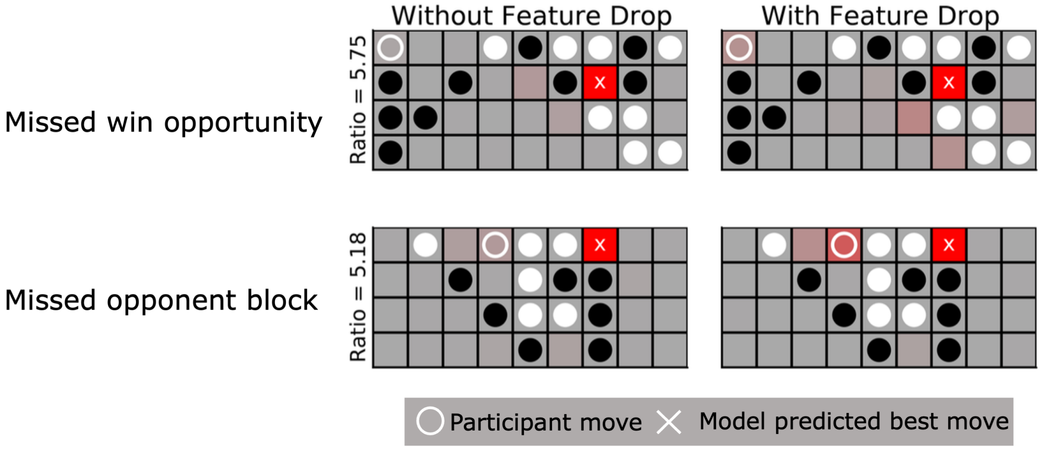


**Figure D.4:** Illustration of scenarios where feature drop rate captures human behaviour. Reprinted with permission from van Opheusden et al. (2023). Feature drop rate corresponds to the probability of dropping particular features on the board. The model preferred move is indicated with an x, while the actual participant move is indicated with an open circle. Red shading shows divergence between model with and without the feature drop parameter, as quantified by Jensen-Shannon divergence.

data better than a model without this parameter (Fig.D.4). As described in van Opheusden et al. (2023), ‘The feature drop mechanism is primarily necessary to account for people’s tendency to overlook possibilities to immediately make four-in-a-row, or block immediate four-in-a-row threats by the opponent.’

# References

Abitbol, Raphaëlle et al. (Feb. 4, 2015). “Neural Mechanisms Underlying Contextual Dependency of Subjective Values: Converging Evidence from Monkeys and Humans”. en. In: *Journal of Neuroscience* 35.5, pp. 2308–2320. DOI: `10.1523/JNEUROSCI.1878-14.2015`. pmid: `25653384`.

Achtziger, Anja and Peter Gollwitzer (Feb. 18, 2008). “Motivation and Volition in the Course of Action”. In: *Publ. in: Motivation and action / ed. by Jutta Heckhausen. New York : Cambridge University Press, 2010, 2nd ed., pp. 275-299.* DOI: `10.1017/CBO9780511499821.012`.

Adams, Christopher D. and Anthony Dickinson (May 1, 1981). “Instrumental Responding Following Reinforcer Devaluation”. en. In: *The Quarterly Journal of Experimental Psychology Section B* 33 (2b), pp. 109–121. DOI: `10.1080/14640748108400816`.

Addis, Donna Rose and Daniel Schacter (Jan. 4, 2012). “The Hippocampus and Imagining the Future: Where Do We Stand?” English. In: *Frontiers in Human Neuroscience* 5. DOI: `10.3389/fnhum.2011.00173`.

Akam, Thomas, Rui Costa, and Peter Dayan (Dec. 11, 2015). “Simple Plans or Sophisticated Habits? State, Transition and Learning Interactions in the Two-Step Task”. en. In: *PLOS Computational Biology* 11.12, e1004648. DOI: `10.1371/journal.pcbi.1004648`.

Amir, Dorsa, Matthew Jordan, and David G. Rand (July 28, 2018). *An Uncertainty Management Perspective on Long-Run Impacts of Adversity: The Influence of Childhood Socioeconomic Status on Risk, Time, and Social Preferences.* en. DOI: `10.2139/ssrn.3037019`. URL: `https://papers.ssrn.com/abstract=3037019` (visited on 04/11/2024). preprint.

Amir, Nadav, Yael Niv, and Angela Langdon (Jan. 3, 2024). *States as Goal-Directed Concepts: An Epistemic Approach to State-Representation Learning.* arXiv: `2312.02367 [cs, math, q-bio]`. URL: `http://arxiv.org/abs/2312.02367` (visited on 03/15/2024). preprint.

Anderson, Brian A., Patryk A. Laurent, and Steven Yantis (June 21, 2011). “Value-Driven Attentional Capture”. eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.25, pp. 10367–10371. DOI: `10.1073/pnas.1104047108`. pmid: `21646524`.

Anderson, Brian A. and Steven Yantis (2013). “Persistence of Value-Driven Attentional Capture”. In: *Journal of Experimental Psychology: Human Perception and Performance* 39.1, pp. 6–9. DOI: `10.1037/a0030860`.

Ariely, Dan and Michael I. Norton (Jan. 2008). “How Actions Create–Not Just Reveal–Preferences”. eng. In: *Trends in Cognitive Sciences* 12.1, pp. 13–16. DOI: `10.1016/j.tics.2007.10.008`. pmid: `18063405`.

Arkes, Hal R. and Catherine Blumer (1985). “The Psychology of Sunk Cost”. In: *Organizational Behavior and Human Decision Processes* 35, pp. 124–140. DOI: 10.1016/0749-5978(85)90049-4.

Armel, K. Carrie, Aurelie Beaumel, and Antonio Rangel (2008). “Biasing Simple Choices by Manipulating Relative Visual Attention”. In: *Judgment and Decision Making* 3.5, pp. 396–403. DOI: 10.1017/S1930297500000413.

Balleine, Bernard (Dec. 15, 2005). “Neural Bases of Food-Seeking: Affect, Arousal and Reward in Corticostriatolimbic Circuits”. In: *Physiology & Behavior*. Purdue University Ingestive Behavior Research Center Symposium. Dietary Influences on Obesity: Environment, Behavior and Biology 86.5, pp. 717–730. DOI: 10.1016/j.physbeh.2005.08.061.

Balleine, Bernard and Anthony Dickinson (Apr. 5, 1998). “Goal-Directed Instrumental Action: Contingency and Incentive Learning and Their Cortical Substrates”. In: *Neuropharmacology* 37.4, pp. 407–419. DOI: 10.1016/S0028-3908(98)00033-1.

Bari, Bilal A. et al. (Sept. 4, 2019). “Stable Representations of Decision Variables for Flexible Behavior”. eng. In: *Neuron* 103.5, 922–933.e7. DOI: 10.1016/j.neuron.2019.06.001. pmid: 31280924.

Baron, Reuben M. and David A. Kenny (1986). “The ModeratorMediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations”. In: *Journal of Personality and Social Psychology* 51.6, pp. 1173–1182. DOI: 10.1037/0022-3514.51.6.1173.

Barron, Helen C., Raymond J. Dolan, and Timothy E. J. Behrens (Oct. 2013). “Online Evaluation of Novel Choices by Simultaneous Representation of Multiple Memories”. en. In: *Nature Neuroscience* 16.10, pp. 1492–1498. DOI: 10.1038/nn.3515.

Barron, Helen C. et al. (Oct. 1, 2020). “Neuronal Computation Underlying Inferential Reasoning in Humans and Mice”. In: *Cell* 183.1, 228–243.e21. DOI: 10.1016/j.cell.2020.08.035.

Bartlett, Frederic Charles (1932). *Remembering: A Study in Experimental and Social Psychology*. Remembering: A Study in Experimental and Social Psychology. New York, NY, US: Cambridge University Press, pp. xix, 317. xix, 317.

Bartolo, Ramon and Bruno B. Averbeck (June 2020). “Prefrontal Cortex Predicts State Switches during Reversal Learning”. en. In: *Neuron* 106.6, 1044–1054.e4. DOI: 10.1016/j.neuron.2020.03.024.

Bartra, Oscar, Joseph T. McGuire, and Joseph W. Kable (Aug. 1, 2013). “The Valuation System: A Coordinate-Based Meta-Analysis of BOLD fMRI Experiments Examining Neural Correlates of Subjective Value”. eng. In: *NeuroImage* 76, pp. 412–427. DOI: 10.1016/j.neuroimage.2013.02.063. pmid: 23507394.

Bates, Elizabeth et al. (May 2003). “Voxel-Based LesionSymptom Mapping”. en. In: *Nature Neuroscience* 6.5, pp. 448–450. DOI: 10.1038/nn1050.

Beer, Randall (Jan. 1, 1995). “A Dynamical Systems Perspective on Agent-Environment Interaction”. In: *Artificial Intelligence* 72.1, pp. 173–215. DOI: 10.1016/0004-3702(94)00005-L.

Behrens, Timothy E. J. et al. (Oct. 24, 2018). “What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior”. eng. In: *Neuron* 100.2, pp. 490–509. DOI: 10.1016/j.neuron.2018.10.002. pmid: 30359611.

Belsky, Jay, Laurence Steinberg, and Patricia Draper (1991). “Childhood Experience, Interpersonal Development, and Reproductive Strategy: An Evolutionary Theory of Socialization”. In: *Child Development* 62.4, pp. 647–670. DOI: 10.2307/1131166.

Bengio, Yoshua, Aaron Courville, and Pascal Vincent (Apr. 23, 2014). *Representation Learning: A Review and New Perspectives.* DOI: `10.48550/arXiv.1206.5538`. arXiv: `1206.5538 [cs]`. URL: `http://arxiv.org/abs/1206.5538` (visited on 03/13/2024). preprint.

Berke, Joshua D. (June 2018). "What Does Dopamine Mean?" en. In: *Nature Neuroscience* 21.6, pp. 787–793. DOI: `10.1038/s41593-018-0152-y`.

Berman, Benjamin A and Robert Dismukes (Dec. 2006). "Pressing the Approach: A NASA Study of 19 Recent Accidents Yields a New Perspective on Pilot Error". In: *Aviation Safety World*, pp. 28–33.

Bertossi, Elena et al. (Jan. 29, 2016a). "Stuck in the Here and Now: Construction of Fictitious and Future Experiences Following Ventromedial Prefrontal Damage". eng. In: *Neuropsychologia* 81, pp. 107–116. DOI: `10.1016/j.neuropsychologia.2015.12.015`. pmid: `26707714`.

Bertossi, Elena et al. (Sept. 2016b). "Ventromedial Prefrontal Damage Causes a Pervasive Impairment of Episodic Memory and Future Thinking". eng. In: *Neuropsychologia* 90, pp. 12–24. DOI: `10.1016/j.neuropsychologia.2016.01.034`. pmid: `26827916`.

Bertossi, Elena et al. (Mar. 2017). "Episodic Future Thinking Following vmPFC Damage: Impaired Event Construction, Maintenance, or Narration?" eng. In: *Neuropsychology* 31.3, pp. 337–348. DOI: `10.1037/neu0000345`. pmid: `28054822`.

Blanchard, Tommy C. and Benjamin Y. Hayden (Jan. 8, 2014). "Neurons in Dorsal Anterior Cingulate Cortex Signal Postdecisional Variables in a Foraging Task". en. In: *Journal of Neuroscience* 34.2, pp. 646–655. DOI: `10.1523/JNEUROSCI.3151-13.2014`. pmid: `24403162`.

— (Feb. 11, 2015). "Monkeys Are More Patient in a Foraging Task than in a Standard Intertemporal Choice Task". en. In: *PLOS ONE* 10.2, e0117057. DOI: `10.1371/journal.pone.0117057`.

Blanco-Pozo, Marta, Thomas Akam, and Mark E. Walton (Feb. 2024). "Dopamine-Independent Effect of Rewards on Choices through Hidden-State Inference". en. In: *Nature Neuroscience* 27.2, pp. 286–297. DOI: `10.1038/s41593-023-01542-x`.

Boorman, Erie D., Matthew F. Rushworth, and Tim E. Behrens (Feb. 6, 2013). "Ventromedial Prefrontal and Anterior Cingulate Cortex Adopt Choice and Default Reference Frames during Sequential Multi-Alternative Choice". en. In: *Journal of Neuroscience* 33.6, pp. 2242–2253. DOI: `10.1523/JNEUROSCI.3022-12.2013`. pmid: `23392656`.

Bouret, Sebastien and Barry J. Richmond (June 23, 2010). "Ventromedial and Orbital Prefrontal Neurons Differentially Encode Internally and Externally Driven Motivational Values in Monkeys". en. In: *Journal of Neuroscience* 30.25, pp. 8591–8601. DOI: `10.1523/JNEUROSCI.0049-10.2010`. pmid: `20573905`.

Bowren, Mark D. et al. (2018). "Choosing Spouses and Houses: Impaired Congruence between Preference and Choice Following Damage to the Ventromedial Prefrontal Cortex". In: *Neuropsychology* 32.3, pp. 280–303. DOI: `10.1037/neu0000421`.

Bradfield, Laura A. et al. (Dec. 16, 2015). "Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations". In: *Neuron* 88.6, pp. 1268–1280. DOI: `10.1016/j.neuron.2015.10.044`.

Bradfield, Laura A. et al. (Oct. 2020). “Goal-Directed Actions Transiently Depend on Dorsal Hippocampus”. en. In: *Nature Neuroscience* 23.10, pp. 1194–1197. DOI: `10.1038/s41593-020-0693-8`.

Bratman, Michael (1987). *Intention, Plans, and Practical Reason.* Cambridge: Cambridge, MA: Harvard University Press.

— (May 4, 2014). “Temptation and the Agents Standpoint”. In: *Inquiry* 57.3, pp. 293–310. DOI: `10.1080/0020174X.2014.894271`.

Brehm, Jack W. (1956). “Postdecision Changes in the Desirability of Alternatives”. In: *The Journal of Abnormal and Social Psychology* 52.3, pp. 384–389. DOI: `10.1037/h0041006`.

Brooks, Jonathan C. W. et al. (Jan. 15, 2008). “Physiological Noise Modelling for Spinal Functional Magnetic Resonance Imaging Studies”. en. In: *NeuroImage* 39.2, pp. 680–692. DOI: `10.1016/j.neuroimage.2007.09.018`.

Brooks, Rodney (Jan. 1, 1991). “Intelligence without Representation”. In: *Artificial Intelligence* 47.1, pp. 139–159. DOI: `10.1016/0004-3702(91)90053-M`.

Browne, Cameron et al. (Mar. 1, 2012). “A Survey of Monte Carlo Tree Search Methods”. In: *IEEE Transactions on Computational Intelligence and AI in Games* 4:1, pp. 1–43. DOI: `10.1109/TCIAIG.2012.2186810`.

Burgess, P. W. et al. (2000). “The Cognitive and Neuroanatomical Correlates of Multitasking”. eng. In: *Neuropsychologia* 38.6, pp. 848–863. DOI: `10.1016/s0028-3932(99)00134-7`. pmid: `10689059`.

Callaway, Frederick et al. (Aug. 2022). “Rational Use of Cognitive Resources in Human Planning”. en. In: *Nature Human Behaviour* 6.8, pp. 1112–1125. DOI: `10.1038/s41562-022-01332-8`.

Camille, Nathalie et al. (May 18, 2011). “Ventromedial Frontal Lobe Damage Disrupts Value Maximization in Humans”. en. In: *Journal of Neuroscience* 31.20, pp. 7527–7532. DOI: `10.1523/JNEUROSCI.6527-10.2011`. pmid: `21593337`.

Caraco, Thomas (June 1, 1981). “Energy Budgets, Risk and Foraging Preferences in Dark-Eyed Juncos (Junco Hyemalis)”. en. In: *Behavioral Ecology and Sociobiology* 8.3, pp. 213–217. DOI: `10.1007/BF00299833`.

Carmichael, S.T. and J.I. Price (1996). “Connectional Networks within the Orbital and Medial Prefrontal Cortex of Macaque Monkeys”. en. In: *Journal of Comparative Neurology* 371.2, pp. 179–207. DOI: `10.1002/(SICI)1096-9861(19960722)371:2<179::AID-CNE1>3.0.CO;2-#`.

Carmichael, S.T. and J.L. Price (Dec. 25, 1995a). “Limbic Connections of the Orbital and Medial Prefrontal Cortex in Macaque Monkeys”. eng. In: *The Journal of Comparative Neurology* 363.4, pp. 615–641. DOI: `10.1002/cne.903630408`. pmid: `8847421`.

— (Dec. 25, 1995b). “Sensory and Premotor Connections of the Orbital and Medial Prefrontal Cortex of Macaque Monkeys”. eng. In: *The Journal of Comparative Neurology* 363.4, pp. 642–664. DOI: `10.1002/cne.903630409`. pmid: `8847422`.

Carpenter, Bob et al. (2017). “Stan: A Probabilistic Programming Language”. eng. In: *Journal of Statistical Software* 76, p. 1. DOI: `10.18637/jss.v076.i01`. pmid: `36568334`.

Castegnetti, G., M. Zurita, and B. De Martino (2021). “How Usefulness Shapes Neural Representations during Goal-Directed Behavior”. In: *Science advances* 7.15, eabd5363.

Castro-Rodrigues, Pedro et al. (Aug. 2022). “Explicit Knowledge of Task Structure Is a Primary Determinant of Human Model-Based Action”. en. In: *Nature Human Behaviour* 6.8, pp. 1126–1141. DOI: `10.1038/s41562-022-01346-2`.

Charnov, Eric L. (Apr. 1, 1976). “Optimal Foraging, the Marginal Value Theorem”. In: *Theoretical Population Biology* 9.2, pp. 129–136. DOI: `10.1016/0040-5809(76)90040-X`.

Chen, M. Keith (July 15, 2008). *Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?* en. URL: `https://papers.ssrn.com/abstract=1160268` (visited on 03/17/2024). preprint.

Chen, M. Keith and Jane L. Risen (Feb. 1, 2009). “Is Choice a Reliable Predictor of Choice? A Comment on Sagarin and Skowronski”. In: *Journal of Experimental Social Psychology* 45.2, pp. 425–427. DOI: `10.1016/j.jesp.2008.08.026`.

Cheng, Shaozhe et al. (Sept. 2023). “Intention beyond Desire: Spontaneous Intentional Commitment Regulates Conflicting Desires”. eng. In: *Cognition* 238, p. 105513. DOI: `10.1016/j.cognition.2023.105513`. pmid: `37331323`.

Chib, Vikram S. et al. (Sept. 30, 2009). “Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex”. en. In: *Journal of Neuroscience* 29.39, pp. 12315–12320. DOI: `10.1523/JNEUROSCI.2575-09.2009`. pmid: `19793990`.

Chu, Junyi and Laura Schulz (2022). “"Because I Want to": Valuing Goals for Their Own Sake”. en. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* 44.44.

Cisek, Paul (Dec. 2012). “Making Decisions through a Distributed Consensus”. eng. In: *Current Opinion in Neurobiology* 22.6, pp. 927–936. DOI: `10.1016/j.conb.2012.05.007`. pmid: `22683275`.

Clairis, Nicolas and Mathias Pessiglione (July 13, 2022). “Value, Confidence, Deliberation: A Functional Partition of the Medial Prefrontal Cortex Demonstrated across Rating and Choice Tasks”. en. In: *Journal of Neuroscience* 42.28, pp. 5580–5592. DOI: `10.1523/JNEUROSCI.1795-21.2022`. pmid: `35654606`.

Coe, Malcolm (1984). “Primates: Their Niche Structure and Habitats”. en. In: *Food Acquisition and Processing in Primates*. Ed. by David J. Chivers, Bernard A. Wood, and Alan Bilsborough. Boston, MA: Springer US, pp. 1–32. DOI: `10.1007/978-1-4757-5244-1_1`.

Collins, Anne L. et al. (Feb. 12, 2016). “Dynamic Mesolimbic Dopamine Signaling during Action Sequence Learning and Expectation Violation”. eng. In: *Scientific Reports* 6, p. 20231. DOI: `10.1038/srep20231`. pmid: `26869075`.

Constantinescu, Alexandra O., Jill X. OReilly, and Timothy E. J. Behrens (June 17, 2016). “Organizing Conceptual Knowledge in Humans with a Gridlike Code”. In: *Science* 352.6292, pp. 1464–1468. DOI: `10.1126/science.aaf0941`.

Corbetta, Maurizio and Gordon L. Shulman (Mar. 2002). “Control of Goal-Directed and Stimulus-Driven Attention in the Brain”. en. In: *Nature Reviews Neuroscience* 3.3, pp. 201–215. DOI: `10.1038/nrn755`.

Czerlinski, Jean, Gerd Gigerenzer, and Daniel G. Goldstein (1999). “How Good Are Simple Heuristics?” In: *Simple Heuristics That Make Us Smart*. Evolution and Cognition. New York, NY, US: Oxford University Press, pp. 97–118.

Dalley, Jeffrey W. and Trevor W. Robbins (Mar. 2017). “Fractionating Impulsivity: Neuropsychiatric Implications”. en. In: *Nature Reviews Neuroscience* 18.3 (3), pp. 158–171. DOI: `10.1038/nrn.2017.8`.

Daw, Nathaniel D., Yael Niv, and Peter Dayan (Dec. 2005). “Uncertainty-Based Competition between Prefrontal and Dorsolateral Striatal Systems for Behavioral Control”. en. In: *Nature Neuroscience* 8.12, pp. 1704–1711. DOI: 10.1038/nn1560.

Daw, Nathaniel D. et al. (Mar. 24, 2011). “Model-Based Influences on Humans Choices and Striatal Prediction Errors”. In: *Neuron* 69.6, pp. 1204–1215. DOI: 10.1016/j.neuron.2011.02.027. pmid: 21435563.

De Martino, Benedetto and Aurelio Cortese (Jan. 2023). “Goals, Usefulness and Abstraction in Value-Based Choice”. eng. In: *Trends in Cognitive Sciences* 27.1, pp. 65–80. DOI: 10.1016/j.tics.2022.11.001. pmid: 36446707.

De Martino, Benedetto et al. (Jan. 2013). “Confidence in Value-Based Choice”. eng. In: *Nature Neuroscience* 16.1, pp. 105–110. DOI: 10.1038/nn.3279. pmid: 23222911.

Dechter, Rina and Judea Pearl (July 1, 1985). “Generalized Best-First Search Strategies and the Optimality of A*”. In: *Journal of the ACM* 32.3, pp. 505–536. DOI: 10.1145/3828.3830.

Decker, Johannes H. et al. (June 2016). “From Creatures of Habit to Goal-Directed Learners: Tracking the Developmental Emergence of Model-Based Reinforcement Learning”. eng. In: *Psychological Science* 27.6, pp. 848–858. DOI: 10.1177/0956797616639301. pmid: 27084852.

Deichmann, R et al. (June 1, 2003). “Optimized EPI for fMRI Studies of the Orbitofrontal Cortex”. en. In: *NeuroImage* 19.2, pp. 430–441. DOI: 10.1016/S1053-8119(03)00073-9.

Deserno, Lorenz et al. (Feb. 3, 2015). “Ventral Striatal Dopamine Reflects Behavioral and Neural Signatures of Model-Based Control during Sequential Decision Making”. eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 112.5, pp. 1595–1600. DOI: 10.1073/pnas.1417219112. pmid: 25605941.

Desimone, R. and J. Duncan (1995). “Neural Mechanisms of Selective Visual Attention”. eng. In: *Annual Review of Neuroscience* 18, pp. 193–222. DOI: 10.1146/annurev.ne.18.030195.001205. pmid: 7605061.

Dickinson, Anthony (1985). “Actions and Habits: The Development of Behavioural Autonomy”. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 308.1135, pp. 67–78. JSTOR: 2396284.

Dickinson, Anthony and Bernard Balleine (2002). “The Role of Learning in the Operation of Motivational Systems”. en. In: *Stevens’ Handbook of Experimental Psychology*. John Wiley & Sons, Ltd. DOI: 10.1002/0471214426.pas0312.

Doody, Max, Maaike M. H. Van Swieten, and Sanjay G. Manohar (Feb. 11, 2022). “Model-Based Learning Retrospectively Updates Model-Free Values”. eng. In: *Scientific Reports* 12.1, p. 2358. DOI: 10.1038/s41598-022-05567-3. pmid: 35149713.

Doya, Kenji et al. (June 1, 2002). “Multiple Model-Based Reinforcement Learning”. In: *Neural Computation* 14.6, pp. 1347–1369. DOI: 10.1162/089976602753712972.

Duncan, John (1984). “Selective Attention and the Organization of Visual Information”. In: *Journal of Experimental Psychology: General* 113.4, pp. 501–517. DOI: 10.1037/0096-3445.113.4.501.

Egan, Louisa C., Laurie R. Santos, and Paul Bloom (Nov. 2007). “The Origins of Cognitive Dissonance: Evidence from Children and Monkeys”. eng. In: *Psychological Science* 18.11, pp. 978–983. DOI: 10.1111/j.1467-9280.2007.02012.x. pmid: 17958712.

Eichenbaum, Howard and Neal J. Cohen (Aug. 20, 2014). “Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function?” eng. In: *Neuron* 83.4, pp. 764–770. DOI: `10.1016/j.neuron.2014.07.032`. pmid: `25144874`.

Einhorn, Hillel J. and Robin M. Hogarth (Apr. 1, 1975). “Unit Weighting Schemes for Decision Making”. In: *Organizational Behavior and Human Performance* 13.2, pp. 171–192. DOI: `10.1016/0030-5073(75)90044-6`.

Ekstrom, Arne D. et al. (Sept. 2003). “Cellular Networks Underlying Human Spatial Navigation”. en. In: *Nature* 425.6954, pp. 184–188. DOI: `10.1038/nature01964`.

Elo, Arpad E. (1978). *The Rating of Chessplayers, Past and Present.* en. Arco Pub. 216 pp. Google Books: `8pMnAQAAMAAJ`.

Eppinger, Ben, Maik Walter, and Hauke R. Heekeren (Dec. 24, 2013). “Of Goals and Habits: Age-Related and Individual Differences in Goal-Directed Decision-Making”. English. In: *Frontiers in Neuroscience* 7. DOI: `10.3389/fnins.2013.00253`.

Eslinger, P. J. and A. R. Damasio (Dec. 1985). “Severe Disturbance of Higher Cognition after Bilateral Frontal Lobe Ablation: Patient EVR”. eng. In: *Neurology* 35.12, pp. 1731–1741. DOI: `10.1212/wnl.35.12.1731`. pmid: `4069365`.

Eysenbach, Benjamin et al. (2022). “Contrastive Learning as Goal-Conditioned Reinforcement Learning”. In: *Advances in Neural Information Processing Systems* 35, pp. 35603–35620.

Farovik, Anja et al. (May 27, 2015). “Orbitofrontal Cortex Encodes Memories within Value-Based Schemas and Represents Contexts That Guide Memory Retrieval”. eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 35.21, pp. 8333–8344. DOI: `10.1523/JNEUROSCI.0134-15.2015`. pmid: `26019346`.

Feher da Silva, Carolina and Todd A. Hare (Oct. 2020). “Humans Primarily Use Model-Based Inference in the Two-Stage Task”. en. In: *Nature Human Behaviour* 4.10, pp. 1053–1066. DOI: `10.1038/s41562-020-0905-y`.

Fellows, Lesley K. (2018). “The Neuroscience of Human Decision-Making Through the Lens of Learning and Memory”. eng. In: *Current Topics in Behavioral Neurosciences* 37, pp. 231–251. DOI: `10.1007/7854_2016_468`. pmid: `28213812`.

Fellows, Lesley K. and Martha J. Farah (2005). “Dissociable Elements of Human Foresight: A Role for the Ventromedial Frontal Lobes in Framing the Future, but Not in Discounting Future Rewards”. In: *Neuropsychologia* 43.8, pp. 1214–1221. DOI: `10.1016/j.neuropsychologia.2004.07.018`.

— (Nov. 2007). “The Role of Ventromedial Prefrontal Cortex in Decision Making: Judgment under Uncertainty or Judgment per Se?” eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 17.11, pp. 2669–2674. DOI: `10.1093/cercor/bhl176`. pmid: `17259643`.

Fermin, Alan S. R. et al. (Aug. 19, 2016). “Model-Based Action Planning Involves Cortico-Cerebellar and Basal Ganglia Networks”. en. In: *Scientific Reports* 6.1, p. 31378. DOI: `10.1038/srep31378`.

Festinger, Leon (1957). *A Theory of Cognitive Dissonance.* A Theory of Cognitive Dissonance. Stanford University Press, pp. xi, 291. xi, 291.

Fouragnan, Elsa F. et al. (May 2019). “The Macaque Anterior Cingulate Cortex Translates Counterfactual Choice Value into Actual Behavioral Change”. en. In: *Nature Neuroscience* 22.5 (5), pp. 797–808. DOI: `10.1038/s41593-019-0375-6`.

Friedel, Eva et al. (Aug. 4, 2014). “Devaluation and Sequential Decisions: Linking Goal-Directed and Model-Based Behavior”. English. In: *Frontiers in Human Neuroscience* 8. DOI: 10.3389/fnhum.2014.00587.

Frömer, Romy, Carolyn K. Dean Wolf, and Amitai Shenhav (2019). “Goal Congruency Dominates Reward Value in Accounting for Behavioral and Neural Correlates of Value-Based Decision-Making”. In: *Nature communications* 10.1, p. 4926.

Genovese, Christopher R., Nicole A. Lazar, and Thomas Nichols (Apr. 2002). “Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate”. eng. In: *NeuroImage* 15.4, pp. 870–878. DOI: 10.1006/nimg.2001.1037. pmid: 11906227.

Gershman, Samuel J., Eric J. Horvitz, and Joshua B. Tenenbaum (July 17, 2015). “Computational Rationality: A Converging Paradigm for Intelligence in Brains, Minds, and Machines”. In: *Science* 349.6245, pp. 273–278. DOI: 10.1126/science.aac6076.

Ghosh, Vanessa E. et al. (Sept. 3, 2014). “Schema Representation in Patients with Ventromedial PFC Lesions”. eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34.36, pp. 12057–12070. DOI: 10.1523/JNEUROSCI.0740-14.2014. pmid: 25186751.

Gigerenzer, Gerd and Wolfgang Gaissmaier (2011). “Heuristic Decision Making”. In: *Annual Review of Psychology* 62.1, pp. 451–482. DOI: 10.1146/annurev-psych-120709-145346. pmid: 21126183.

Gigerenzer, Gerd and Peter M. Todd (1999). *Simple Heuristics That Make Us Smart.* Simple Heuristics That Make Us Smart. New York, NY, US: Oxford University Press, pp. xv, 416. xv, 416.

Giguère, Gyslain and Bradley C. Love (May 7, 2013). “Limits in Decision Making Arise from Limits in Memory Retrieval”. In: *Proceedings of the National Academy of Sciences* 110.19, pp. 7613–7618. DOI: 10.1073/pnas.1219674110.

Gläscher, Jan et al. (May 27, 2010). “States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning”. In: *Neuron* 66.4, pp. 585–595. DOI: 10.1016/j.neuron.2010.04.016. pmid: 20510862.

Glasser, Matthew F. et al. (Aug. 2016). “A Multi-Modal Parcellation of Human Cerebral Cortex”. en. In: *Nature* 536.7615, pp. 171–178. DOI: 10.1038/nature18933.

Gluth, Sebastian, Mikhail S. Spektor, and Jörg Rieskamp (2018). “Value-Based Attentional Capture Affects Multi-Alternative Decision Making”. In: *Elife* 7, e39659.

Gluth, Sebastian et al. (June 2020a). “Value-Based Attention but Not Divisive Normalization Influences Decisions with Multiple Alternatives”. en. In: *Nature Human Behaviour* 4.6 (6), pp. 634–645. DOI: 10.1038/s41562-020-0822-0.

— (June 2020b). “Value-Based Attention but Not Divisive Normalization Influences Decisions with Multiple Alternatives”. en. In: *Nature Human Behaviour* 4.6, pp. 634–645. DOI: 10.1038/s41562-020-0822-0.

Griskevicius, Vladas et al. (June 2011). “The Influence of Mortality and Socioeconomic Status on Risk and Delayed Rewards: A Life History Theory Approach”. eng. In: *Journal of Personality and Social Psychology* 100.6, pp. 1015–1026. DOI: 10.1037/a0022403. pmid: 21299312.

Grueschow, Marcus et al. (Feb. 18, 2015). “Automatic versus Choice-Dependent Value Representations in the Human Brain”. en. In: *Neuron* 85.4, pp. 874–885. DOI: 10.1016/j.neuron.2014.12.054.

Günseli, Eren and Mariam Aly (Apr. 7, 2020). “Preparation for Upcoming Attentional States in the Hippocampus and Medial Prefrontal Cortex”. In: *eLife* 9. Ed. by Morgan Barense, Timothy E Behrens, and Kia Nobre, e53191. DOI: 10.7554/eLife.53191.

Gür, Ezgi, Yalçn Akn Duyan, and Fuat Balc (Jan. 1, 2018). “Spontaneous Integration of Temporal Information: Implications for Representational/Computational Capacity of Animals”. en. In: *Animal Cognition* 21.1, pp. 3–19. DOI: 10.1007/s10071-017-1137-z.

Guru, Akash et al. (May 22, 2020). *Ramping Activity in Midbrain Dopamine Neurons Signifies the Use of a Cognitive Map.* en. DOI: 10.1101/2020.05.21.108886. URL: https://www.biorxiv.org/content/10.1101/2020.05.21.108886v1 (visited on 03/18/2024). preprint.

Hafting, Torkel et al. (Aug. 2005). “Microstructure of a Spatial Map in the Entorhinal Cortex”. en. In: *Nature* 436.7052, pp. 801–806. DOI: 10.1038/nature03721.

Hamid, Arif A. et al. (Jan. 2016a). “Mesolimbic Dopamine Signals the Value of Work”. en. In: *Nature Neuroscience* 19.1 (1), pp. 117–126. DOI: 10.1038/nn.4173.

— (Jan. 2016b). “Mesolimbic Dopamine Signals the Value of Work”. en. In: *Nature Neuroscience* 19.1 (1), pp. 117–126. DOI: 10.1038/nn.4173.

Hampton, Alan N., Peter Bossaerts, and John P. ODoherty (Aug. 9, 2006). “The Role of the Ventromedial Prefrontal Cortex in Abstract State-Based Inference during Decision Making in Humans”. en. In: *Journal of Neuroscience* 26.32, pp. 8360–8367. DOI: 10.1523/JNEUROSCI.1010-06.2006. pmid: 16899731.

Hare, Todd A., Jonathan Malmaud, and Antonio Rangel (July 27, 2011). “Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice”. en. In: *Journal of Neuroscience* 31.30, pp. 11077–11087. DOI: 10.1523/JNEUROSCI.6383-10.2011. pmid: 21795556.

Harlow, John Martyn (1868). *Recovery from the Passage of an Iron Bar through the Head - Digital Collections - National Library of Medicine.* 20th ed. Vol. 39. Boston: Publications of the Massachusetts Medical Society. 327-47.

Hayden, Benjamin Y (Dec. 1, 2018). “Economic Choice: The Foraging Perspective”. In: *Current Opinion in Behavioral Sciences.* Survival Circuits 24, pp. 1–6. DOI: 10.1016/j.cobeha.2017.12.002.

Hayden, Benjamin Y. and Rubén Moreno-Bote (Jan. 1, 2018). “A Neuronal Theory of Sequential Economic Choice”. en. In: *Brain and Neuroscience Advances* 2, p. 2398212818766675. DOI: 10.1177/2398212818766675.

Hayden, Benjamin Y., John M. Pearson, and Michael L. Platt (2011). “Neuronal Basis of Sequential Foraging Decisions in a Patchy Environment”. In: *Nature neuroscience* 14.7, pp. 933–939.

Heberlein, Andrea S. et al. (Apr. 2008). “Ventromedial Frontal Lobe Plays a Critical Role in Facial Emotion Recognition”. eng. In: *Journal of Cognitive Neuroscience* 20.4, pp. 721–733. DOI: 10.1162/jocn.2008.20049. pmid: 18052791.

Heckhausen, Heinz and Peter M. Gollwitzer (June 1, 1987). “Thought Contents and Cognitive Functioning in Motivational versus Volitional States of Mind”. en. In: *Motivation and Emotion* 11.2, pp. 101–120. DOI: 10.1007/BF00992338.

Heilbronner, Sarah R. and Benjamin Y. Hayden (2016). “Dorsal Anterior Cingulate Cortex: A Bottom-Up View”. In: *Annual Review of Neuroscience* 39.1, pp. 149–170. DOI: `10.1146/annurev-neuro-070815-013952`. pmid: `27090954`.

Heron, Campbell Le et al. (Mar. 1, 2019). “Brain Mechanisms Underlying Apathy”. en. In: *Journal of Neurology, Neurosurgery & Psychiatry* 90.3, pp. 302–312. DOI: `10.1136/jnnp-2018-318265`. pmid: `30366958`.

Holroyd, Clay B. and Tom Verguts (Apr. 2021). “The Best Laid Plans: Computational Principles of Anterior Cingulate Cortex”. eng. In: *Trends in Cognitive Sciences* 25.4, pp. 316–329. DOI: `10.1016/j.tics.2021.01.008`. pmid: `33593641`.

Holroyd, Clay B. and Nick Yeung (2012). “Motivation of Extended Behaviors by Anterior Cingulate Cortex”. In: *Trends in Cognitive Sciences* 16.2, pp. 122–128. DOI: `10.1016/j.tics.2011.12.008`.

Hornak, J., E. T. Rolls, and D. Wade (Apr. 1996). “Face and Voice Expression Identification in Patients with Emotional and Behavioural Changes Following Ventral Frontal Lobe Damage”. eng. In: *Neuropsychologia* 34.4, pp. 247–261. DOI: `10.1016/0028-3932(95)00106-9`. pmid: `8657356`.

Howe, Mark W. et al. (Aug. 2013). “Prolonged Dopamine Signalling in Striatum Signals Proximity and Value of Distant Rewards”. en. In: *Nature* 500.7464 (7464), pp. 575–579. DOI: `10.1038/nature12475`.

Hoy, Jennifer L., Hannah I. Bishop, and Cristopher M. Niell (Dec. 2, 2019). “Defined Cell Types in Superior Colliculus Make Distinct Contributions to Prey Capture Behavior in the Mouse”. eng. In: *Current biology: CB* 29.23, 4130–4138.e5. DOI: `10.1016/j.cub.2019.10.017`. pmid: `31761701`.

Hunt, L. T. et al. (Aug. 2021). “Formalizing Planning and Information Search in Naturalistic Decision-Making”. en. In: *Nature Neuroscience* 24.8, pp. 1051–1064. DOI: `10.1038/s41593-021-00866-w`.

Hunt, Laurence T et al. (Dec. 11, 2015). “Capturing the Temporal Evolution of Choice across Prefrontal Cortex”. In: *eLife* 4. Ed. by Michael J Frank, e11945. DOI: `10.7554/eLife.11945`.

Hunt, Laurence T. (June 2021). “Frontal Circuit Specialisations for Decision Making”. eng. In: *The European Journal of Neuroscience* 53.11, pp. 3654–3671. DOI: `10.1111/ejn.15236`. pmid: `33864305`.

Hunt, Laurence T. et al. (Mar. 2012). “Mechanisms Underlying Cortical Activity during Value-Guided Choice”. en. In: *Nature Neuroscience* 15.3, pp. 470–476. DOI: `10.1038/nn.3017`.

Hunt, Laurence T. et al. (Oct. 2018). “Triple Dissociation of Attention and Decision Computations across Prefrontal Cortex”. en. In: *Nature Neuroscience* 21.10, pp. 1471–1481. DOI: `10.1038/s41593-018-0239-5`.

Huys, Quentin J. M. et al. (Mar. 8, 2012). “Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees”. en. In: *PLOS Computational Biology* 8.3, e1002410. DOI: `10.1371/journal.pcbi.1002410`.

Jenkinson, M. and S. Smith (June 2001). “A Global Optimisation Method for Robust Affine Registration of Brain Images”. eng. In: *Medical Image Analysis* 5.2, pp. 143–156. DOI: `10.1016/s1361-8415(01)00036-6`. pmid: `11516708`.

Jenkinson, Mark et al. (Oct. 2002). “Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images”. eng. In: *NeuroImage* 17.2, pp. 825–841. DOI: `10.1016/s1053-8119(02)91132-8`. pmid: `12377157`.

Jenkinson, Mark et al. (Aug. 15, 2012). “FSL”. eng. In: *NeuroImage* 62.2, pp. 782–790. DOI: `10.1016/j.neuroimage.2011.09.015`. pmid: `21979382`.

Jevons, William Stanley (1866). “Brief Account of a General Mathematical Theory of Political Economy”. en. In: *History of Economic Thought Articles* 29, pp. 282–287.

Jocham, Gerhard, Tilmann A. Klein, and Markus Ullsperger (Feb. 2, 2011). “Dopamine-Mediated Reinforcement Learning Signals in the Striatum and Ventromedial Prefrontal Cortex Underlie Value-Based Choices”. eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31.5, pp. 1606–1613. DOI: `10.1523/JNEUROSCI.3904-10.2011`. pmid: `21289169`.

Johnson, Adam and David Redish (Nov. 7, 2007). “Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point”. en. In: *Journal of Neuroscience* 27.45, pp. 12176–12189. DOI: `10.1523/JNEUROSCI.3761-07.2007`. pmid: `17989284`.

Johnson, Dominic D. P. et al. (Aug. 1, 2013). “The Evolution of Error: Error Management, Cognitive Constraints, and Adaptive Decision-Making Biases”. English. In: *Trends in Ecology & Evolution* 28.8, pp. 474–481. DOI: `10.1016/j.tree.2013.05.014`.

Jones, Joshua et al. (Nov. 16, 2012). “Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values”. In: *Science* 338.6109, pp. 953–956. DOI: `10.1126/science.1227489`.

Jones, Matt and Fabian Canas (2010). “Integrating Reinforcement Learning with Models of Representation Learning”. en. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* 32.32.

Juechems, Keno and Christopher Summerfield (2019). “Where Does Value Come From?” In: *Trends in cognitive sciences* 23.10, pp. 836–850.

Kable, Joseph W. and Paul W. Glimcher (Dec. 2007). “The Neural Correlates of Subjective Value during Intertemporal Choice”. en. In: *Nature Neuroscience* 10.12, pp. 1625–1633. DOI: `10.1038/nn2007`.

Kaelbling, L. (1993). “Learning to Achieve Goals”. In: International Joint Conference on Artificial Intelligence.

Kaiser, Luca F. et al. (2021). “Dissociable Roles of Cortical Excitation-Inhibition Balance during Patch-Leaving versus Value-Guided Decisions”. In: *Nature Communications* 12.1, p. 904.

Karnath, Hans-Otto, Christoph Sperber, and Christopher Rorden (Jan. 15, 2018). “Mapping Human Brain Lesions and Their Functional Consequences”. In: *NeuroImage* 165, pp. 180–189. DOI: `10.1016/j.neuroimage.2017.10.028`. pmid: `29042216`.

Kearns, Michael, Yishay Mansour, and Andrew Y. Ng (Nov. 1, 2002). “A Sparse Sampling Algorithm for Near-Optimal Planning in Large Markov Decision Processes”. en. In: *Machine Learning* 49.2, pp. 193–208. DOI: `10.1023/A:1017932429737`.

Keramati, Mehdi, Amir Dezfouli, and Payam Piray (May 26, 2011). “Speed/Accuracy Trade-Off between the Habitual and the Goal-Directed Processes”. en. In: *PLOS Computational Biology* 7.5, e1002055. DOI: `10.1371/journal.pcbi.1002055`.

Keramati, Mehdi and Boris Gutkin (Dec. 2, 2014). “Homeostatic Reinforcement Learning for Integrating Reward Collection and Physiological Stability”. eng. In: *eLife* 3, e04811. DOI: `10.7554/eLife.04811`. pmid: `25457346`.

Keramati, Mehdi et al. (Nov. 8, 2016). "Adaptive Integration of Habits into Depth-Limited Planning Defines a Habitual-GoalDirected Spectrum". In: *Proceedings of the National Academy of Sciences* 113.45, pp. 12868–12873. DOI: `10.1073/pnas.1609094113`.

Kidd, Celeste, Holly Palmeri, and Richard N. Aslin (Jan. 1, 2013). "Rational Snacking: Young Childrens Decision-Making on the Marshmallow Task Is Moderated by Beliefs about Environmental Reliability". In: *Cognition* 126.1, pp. 109–114. DOI: `10.1016/j.cognition.2012.08.004`.

Killcross, Simon and Etienne Coutureau (Apr. 2003). "Coordination of Actions and Habits in the Medial Prefrontal Cortex of Rats". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 13.4, pp. 400–408. DOI: `10.1093/cercor/13.4.400`. pmid: `12631569`.

Klein-Flügge, Miriam C., Alessandro Bongioanni, and Matthew F. S. Rushworth (Sept. 7, 2022). "Medial and Orbital Frontal Cortex in Decision-Making and Flexible Behavior". eng. In: *Neuron* 110.17, pp. 2743–2770. DOI: `10.1016/j.neuron.2022.05.022`. pmid: `35705077`.

Klein-Flügge, Miriam C. et al. (Oct. 23, 2019). "Multiple Associative Structures Created by Reinforcement and Incidental Statistical Learning Mechanisms". en. In: *Nature Communications* 10.1, p. 4835. DOI: `10.1038/s41467-019-12557-z`.

Klinger, Eric (1975). "Consequences of Commitment to and Disengagement from Incentives". In: *Psychological Review* 82.1, pp. 1–25. DOI: `10.1037/h0076171`.

Koller, Daphne and Nir Friedman (July 2009). *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press. 1208 pp.

Kolling, Nils and Jill X O'Reilly (Aug. 1, 2018). "State-Change Decisions and Dorsomedial Prefrontal Cortex: The Importance of Time". In: *Current Opinion in Behavioral Sciences*. Apathy and Motivation 22, pp. 152–160. DOI: `10.1016/j.cobeha.2018.06.017`.

Kolling, Nils et al. (Apr. 6, 2012). "Neural Mechanisms of Foraging". eng. In: *Science (New York, N.Y.)* 336.6077, pp. 95–98. DOI: `10.1126/science.1216930`. pmid: `22491854`.

Kolling, Nils et al. (Oct. 2016). "Value, Search, Persistence and Model Updating in Anterior Cingulate Cortex". en. In: *Nature Neuroscience* 19.10, pp. 1280–1285. DOI: `10.1038/nn.4382`.

Koo, Terry K. and Mae Y. Li (June 2016). "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research". eng. In: *Journal of Chiropractic Medicine* 15.2, pp. 155–163. DOI: `10.1016/j.jcm.2016.02.012`. pmid: `27330520`.

Koscik, Timothy R. and Daniel Tranel (May 1, 2012). "The Human Ventromedial Prefrontal Cortex Is Critical for Transitive Inference". In: *Journal of Cognitive Neuroscience* 24.5, pp. 1191–1204. DOI: `10.1162/jocn_a_00203`.

Kouneiher, Frédérique, Sylvain Charron, and Etienne Koechlin (July 2009). "Motivation and Cognitive Control in the Human Prefrontal Cortex". en. In: *Nature Neuroscience* 12.7 (7), pp. 939–945. DOI: `10.1038/nn.2321`.

Krajbich, Ian (Oct. 2019). "Accounting for Attention in Sequential Sampling Models of Decision Making". eng. In: *Current Opinion in Psychology* 29, pp. 6–11. DOI: `10.1016/j.copsyc.2018.10.008`. pmid: `30368108`.

Krajbich, Ian, Carrie Armel, and Antonio Rangel (Oct. 2010). “Visual Fixations and the Computation and Comparison of Value in Simple Choice”. en. In: *Nature Neuroscience* 13.10, pp. 1292–1298. DOI: `10.1038/nn.2635`.

Kruglanski, Arie W. (1996). “Goals as Knowledge Structures”. In: *The Psychology of Action: Linking Cognition and Motivation to Behavior*. New York, NY, US: The Guilford Press, pp. 599–618.

Kumar, Ravin et al. (Jan. 15, 2019). “ArviZ a Unified Library for Exploratory Analysis of Bayesian Models in Python”. en. In: *Journal of Open Source Software* 4.33, p. 1143. DOI: `10.21105/joss.01143`.

Lavie, Nilli (Feb. 2005). “Distracted and Confused?: Selective Attention under Load”. eng. In: *Trends in Cognitive Sciences* 9.2, pp. 75–82. DOI: `10.1016/j.tics.2004.12.004`. pmid: `15668100`.

Le Pelley, Mike E. et al. (2015). “When Goals Conflict with Values: Counterproductive Attentional and Oculomotor Capture by Reward-Related Stimuli”. In: *Journal of Experimental Psychology: General* 144.1, pp. 158–171. DOI: `10.1037/xge0000037`.

Lebreton, Maël et al. (2009). “An Automatic Valuation System in the Human Brain: Evidence from Functional Neuroimaging”. In: *Neuron* 64.3, pp. 431–439.

Lee, Douglas and Jean Daunizeau (2020). “Choosing What We like vs Liking What We Choose: How Choice-Induced Preference Change Might Actually Be Instrumental to Decision-Making”. eng. In: *PloS One* 15.5, e0231081. DOI: `10.1371/journal.pone.0231081`. pmid: `32421699`.

— (Apr. 26, 2021). “Trading Mental Effort for Confidence in the Metacognitive Control of Value-Based Decision-Making”. In: *eLife* 10. Ed. by Tobias Donner, Michael Frank, and Andrew Westbook, e63282. DOI: `10.7554/eLife.63282`.

Lee, Sangil et al. (Aug. 15, 2021). “Subjective Value, Not a Gridlike Code, Describes Neural Activity in Ventromedial Prefrontal Cortex during Value-Based Decision-Making”. In: *NeuroImage* 237, p. 118159. DOI: `10.1016/j.neuroimage.2021.118159`.

Leong, Yuan Chang et al. (Jan. 18, 2017). “Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments”. In: *Neuron* 93.2, pp. 451–463. DOI: `10.1016/j.neuron.2016.12.040`.

Levy, Dino J and Paul W Glimcher (Dec. 1, 2012). “The Root of All Value: A Neural Common Currency for Choice”. In: *Current Opinion in Neurobiology*. Decision Making 22.6, pp. 1027–1038. DOI: `10.1016/j.conb.2012.06.001`.

— (Oct. 12, 2011). “Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain”. en. In: *Journal of Neuroscience* 31.41, pp. 14693–14707. DOI: `10.1523/JNEUROSCI.2218-11.2011`. pmid: `21994386`.

Li, Jiahui, Sabine Hügelschäfer, and Anja Achtziger (2019). “A Self-Regulatory Approach to Rational Decisions: The Implemental Mindset Optimizes Economic Decision Making in Situations Requiring Belief Updating”. en. In: *Journal of Theoretical Social Psychology* 3.2, pp. 115–126. DOI: `10.1002/jts5.38`.

Lim, Seung-Lark, John P. O’Doherty, and Antonio Rangel (Sept. 14, 2011). “The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That Is Guided by Visual Attention”. en. In: *Journal of Neuroscience* 31.37, pp. 13214–13223. DOI: `10.1523/JNEUROSCI.1246-11.2011`. pmid: `21917804`.

Liu, Minghuan, Menghui Zhu, and Weinan Zhang (Sept. 2, 2022). *Goal-Conditioned Reinforcement Learning: Problems and Solutions*. DOI:

10.48550/arXiv.2201.08299. arXiv: 2201.08299 [cs]. URL: http://arxiv.org/abs/2201.08299 (visited on 03/15/2024). preprint.

Loosen, Alisa M., Tricia Seow, and Tobias U. Hauser (Mar. 11, 2022). *Consistency within Change: Evaluating the Psychometric Properties of a Widely-Used Predictive-Inference Task*. en. DOI: 10.31234/osf.io/qkf7j. URL: https://osf.io/preprints/psyarxiv/qkf7j/ (visited on 10/20/2023). preprint.

Lopez-Persem, Alizée, Philippe Domenech, and Mathias Pessiglione (Nov. 19, 2016). "How Prior Preferences Determine Decision-Making Frames and Biases in the Human Brain". In: *eLife* 5. Ed. by Michael J Frank, e20317. DOI: 10.7554/eLife.20317.

Lopez-Persem, Alizée et al. (2020). "Four Core Properties of the Human Brain Valuation System Demonstrated in Intracranial Signals". In: *Nature Neuroscience* 23.5, pp. 664–675. DOI: 10.1038/s41593-020-0615-9.

Love, Bradley C., Douglas L. Medin, and Todd M. Gureckis (Apr. 2004). "SUSTAIN: A Network Model of Category Learning". eng. In: *Psychological Review* 111.2, pp. 309–332. DOI: 10.1037/0033-295X.111.2.309. pmid: 15065912.

Ludwig, Jonas, Alexander Jaudas, and Anja Achtziger (2020). "The Role of Motivation and Volition in Economic Decisions: Evidence from Eye Movements and Pupillometry". In: *Journal of Behavioral Decision Making* 33, pp. 180–195. DOI: 10.1002/bdm.2152.

Lyngs, Ulrik (2019). "Oxforddown: An Oxford University Thesis Template for R Markdown". In: *GitHub repository*. DOI: 10.5281/zenodo.3484681.

Ma, Ili et al. (2022). "The Component Processes of Complex Planning Follow Distinct Developmental Trajectories". In.

Mack, Michael L., Alison R. Preston, and Bradley C. Love (Jan. 7, 2020). "Ventromedial Prefrontal Cortex Compression during Concept Learning". en. In: *Nature Communications* 11.1 (1), p. 46. DOI: 10.1038/s41467-019-13930-8.

Mackey, Scott and Michael Petrides (Sept. 2014). "Architecture and Morphology of the Human Ventromedial Prefrontal Cortex". eng. In: *The European Journal of Neuroscience* 40.5, pp. 2777–2796. DOI: 10.1111/ejn.12654. pmid: 25123211.

Mackintosh, N. J. (1975). "A Theory of Attention: Variations in the Associability of Stimuli with Reinforcement". In: *Psychological Review* 82.4, pp. 276–298. DOI: 10.1037/h0076778.

MacLean, Mary and Barry Giesbrecht (Aug. 22, 2014). "Attention Capture by Task-Irrelevant Learned Value Interacts with Task-Relevant Top-down Factors". In: *Journal of Vision* 14.10, p. 501. DOI: 10.1167/14.10.501.

Mahmoodi, Ali et al. (2023). "Causal Role of a Neural System for Separating and Selecting Multidimensional Social Cognitive Information". In: *Neuron* 111.7, pp. 1152–1164.

Mante, Valerio et al. (Nov. 2013). "Context-Dependent Computation by Recurrent Dynamics in Prefrontal Cortex". en. In: *Nature* 503.7474 (7474), pp. 78–84. DOI: 10.1038/nature12742.

Mcafee, R. Preston, Hugo M. Mialon, and Sue H. Mialon (2010). "Do Sunk Costs Matter?" en. In: *Economic Inquiry* 48.2, pp. 323–336. DOI: 10.1111/j.1465-7295.2008.00184.x.

McNamara, John M. and Alasdair I. Houston (Mar. 1, 1992). "Risk-Sensitive Foraging: A Review of the Theory". en. In: *Bulletin of Mathematical Biology* 54.2, pp. 355–378. DOI: 10.1007/BF02464838.

Mehta, Priyanka S., Seng Bum Michael Yoo, and Benjamin Y. Hayden (Oct. 8, 2020). *Signatures of Processing Complexity during Global Cognitive States in Ventromedial Prefrontal Cortex*. en. DOI: `10.1101/2020.10.08.331579`. URL: `https://www.biorxiv.org/content/10.1101/2020.10.08.331579v1` (visited on 03/30/2024). preprint.

Mehta, Priyanka S. et al. (July 3, 2019). “Ventromedial Prefrontal Cortex Tracks Multiple Environmental Variables during Search”. en. In: *Journal of Neuroscience* 39.27, pp. 5336–5350. DOI: `10.1523/JNEUROSCI.2365-18.2019`. pmid: `31028117`.

Metropolis, Nicholas and S. Ulam (Sept. 1, 1949). “The Monte Carlo Method”. In: *Journal of the American Statistical Association* 44.247, pp. 335–341. DOI: `10.1080/01621459.1949.10483310`. pmid: `18139350`.

Miller, Kevin J., Matthew M. Botvinick, and Carlos D. Brody (Sept. 2017). “Dorsal Hippocampus Contributes to Model-Based Planning”. en. In: *Nature Neuroscience* 20.9, pp. 1269–1276. DOI: `10.1038/nn.4613`.

Mobbs, Dean et al. (July 2018). “Foraging for Foundations in Decision Neuroscience: Insights from Ethology”. en. In: *Nature Reviews Neuroscience* 19.7, pp. 419–427. DOI: `10.1038/s41583-018-0010-7`.

Molinaro, Gaia and Anne G. E. Collins (Dec. 1, 2023). “A Goal-Centric Outlook on Learning”. English. In: *Trends in Cognitive Sciences* 27.12, pp. 1150–1164. DOI: `10.1016/j.tics.2023.08.011`. pmid: `37696690`.

Moscovitch, Morris (1989). “Confabulation and the Frontal Systems: Strategic versus Associative Retrieval in Neuropsychological Theories of Memory”. In: *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc, pp. 133–160.

Muller, Timothy H. et al. (Feb. 28, 2019). “Control of Entropy in Neural Models of Environmental State”. eng. In: *eLife* 8, e39404. DOI: `10.7554/eLife.39404`. pmid: `30816090`.

Newell, A., J. Shaw, and H. Simon (1959). “Report on a General Problem-Solving Program”. In: IFIP Congress.

Niv, Yael (Oct. 2019). “Learning Task-State Representations”. en. In: *Nature Neuroscience* 22.10, pp. 1544–1553. DOI: `10.1038/s41593-019-0470-8`.

Noonan, MaryAnn P. et al. (Nov. 23, 2010). “Separate Value Comparison and Learning Mechanisms in Macaque Medial and Lateral Orbitofrontal Cortex”. eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 107.47, pp. 20547–20552. DOI: `10.1073/pnas.1012246107`. pmid: `21059901`.

Noonan, MaryAnn P. et al. (2012). “Re-Evaluating the Role of the Orbitofrontal Cortex in Reward and Reinforcement”. en. In: *European Journal of Neuroscience* 35.7, pp. 997–1010. DOI: `10.1111/j.1460-9568.2012.08023.x`.

Noonan, MaryAnn P. et al. (July 19, 2017). “Contrasting Effects of Medial and Lateral Orbitofrontal Cortex Lesions on Credit Assignment and Decision-Making in Humans”. en. In: *Journal of Neuroscience* 37.29, pp. 7023–7035. DOI: `10.1523/JNEUROSCI.0692-17.2017`. pmid: `28630257`.

Nussenbaum, Kate et al. (Nov. 23, 2020). “Moving Developmental Research Online: Comparing In-Lab and Web-Based Studies of Model-Based Reinforcement Learning”. In: *Collabra: Psychology* 6.1, p. 17213. DOI: `10.1525/collabra.17213`.

O’Keefe, J. and L. Nadel (1978). *The Hippocampus as a Cognitive Map*. eng. Oxford, UK: Oxford University Press. 570 pp.

O'Reilly, Randall C. et al. (Apr. 30, 2014). *Goal-Driven Cognition in the Brain: A Computational Framework*. DOI: `10.48550/arXiv.1404.7591`. arXiv: `1404.7591 [q-bio]`. URL: `http://arxiv.org/abs/1404.7591` (visited on 03/14/2024). preprint.

OReilly, Randall C. (June 1, 2020). “Unraveling the Mysteries of Motivation”. en. In: *Trends in Cognitive Sciences* 24.6, pp. 425–434. DOI: `10.1016/j.tics.2020.03.001`.

Ongür, D. and J. L. Price (Mar. 2000). “The Organization of Networks within the Orbital and Medial Prefrontal Cortex of Rats, Monkeys and Humans”. eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 10.3, pp. 206–219. DOI: `10.1093/cercor/10.3.206`. pmid: `10731217`.

Ongür, Dost, Amon T. Ferry, and Joseph L. Price (June 2, 2003). “Architectonic Subdivision of the Human Orbital and Medial Prefrontal Cortex”. eng. In: *The Journal of Comparative Neurology* 460.3, pp. 425–449. DOI: `10.1002/cne.10609`. pmid: `12692859`.

Orasanu, Judith, L Martin, and J Davison (2001). “Cognitive and Contextual Factors in Aviation Accidents: Decision Errors”. In: *Linking Expertise and Naturalistic Decision Making*. Lawrence Erlbaum Associates, Inc., pp. 209–225.

Orasanu, Judith et al. (Jan. 1, 1998). “Errors in Aviation Decision Making: Bad Decisions or Bad Luck?”

Otto, A. Ross et al. (May 1, 2013). “The Curse of Planning: Dissecting Multiple Reinforcement-Learning Systems by Taxing the Central Executive”. en. In: *Psychological Science* 24.5, pp. 751–761. DOI: `10.1177/0956797612463080`.

Owen, A. M. et al. (1990). “Planning and Spatial Working Memory Following Frontal Lobe Lesions in Man”. eng. In: *Neuropsychologia* 28.10, pp. 1021–1034. DOI: `10.1016/0028-3932(90)90137-d`. pmid: `2267054`.

Park, Seongmin A., Douglas S. Miller, and Erie D. Boorman (Sept. 2021). “Inferences on a Multidimensional Social Hierarchy Use a Grid-like Code”. en. In: *Nature Neuroscience* 24.9 (9), pp. 1292–1301. DOI: `10.1038/s41593-021-00916-3`.

Payne, John W., James R. Bettman, and Eric J. Johnson (May 28, 1993). *The Adaptive Decision Maker*. en. Cambridge University Press. 352 pp. Google Books: `QzXFqwrPLXkC`.

Pearl, Judea (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan kaufmann.

Pearson, Daniel et al. (Oct. 1, 2016). “Value-Modulated Oculomotor Capture by Task-Irrelevant Stimuli Is a Consequence of Early Competition on the Saccade Map”. en. In: *Attention, Perception, & Psychophysics* 78.7, pp. 2226–2240. DOI: `10.3758/s13414-016-1135-2`.

Pelletier, Gabriel and Lesley K. Fellows (May 22, 2019). “A Critical Role for Human Ventromedial Frontal Lobe in Value Comparison of Complex Objects Based on Attribute Configuration”. en. In: *Journal of Neuroscience* 39.21, pp. 4124–4132. DOI: `10.1523/JNEUROSCI.2969-18.2019`. pmid: `30867258`.

— (Apr. 2021). “Viewing Orbitofrontal Cortex Contributions to Decision-Making through the Lens of Object Recognition”. eng. In: *Behavioral Neuroscience* 135.2, pp. 182–191. DOI: `10.1037/bne0000447`. pmid: `33734731`.

Pelletier, Gabriel et al. (June 9, 2021). “A Preferential Role for Ventromedial Prefrontal Cortex in Assessing the Value of the Whole in Multiattribute Object Evaluation”.

en. In: *Journal of Neuroscience* 41.23, pp. 5056–5068. DOI: 10.1523/JNEUROSCI.0241-21.2021. pmid: 33906899.

Peters, Jan and Mark D'Esposito (Oct. 10, 2016). "Effects of Medial Orbitofrontal Cortex Lesions on Self-Control in Intertemporal Choice". eng. In: *Current biology: CB* 26.19, pp. 2625–2628. DOI: 10.1016/j.cub.2016.07.035. pmid: 27593380.

Piray, Payam, Ivan Toni, and Roshan Cools (Mar. 9, 2016). "Human Choice Strategy Varies with Anatomical Projections from Ventromedial Prefrontal Cortex to Medial Striatum". en. In: *Journal of Neuroscience* 36.10, pp. 2857–2867. DOI: 10.1523/JNEUROSCI.2033-15.2016. pmid: 26961942.

Plappert, Matthias et al. (Mar. 10, 2018). *Multi-Goal Reinforcement Learning: Challenging Robotics Environments and Request for Research.* DOI: 10.48550/arXiv.1802.09464. arXiv: 1802.09464 [cs]. URL: http://arxiv.org/abs/1802.09464 (visited on 03/15/2024). preprint.

Plassmann, Hilke, John O'Doherty, and Antonio Rangel (Sept. 12, 2007). "Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27.37, pp. 9984–9988. DOI: 10.1523/JNEUROSCI.2131-07.2007. pmid: 17855612.

Potter, Tracey C. S., Nessa V. Bryce, and Catherine A. Hartley (June 1, 2017). "Cognitive Components Underpinning the Development of Model-Based Learning". In: *Developmental Cognitive Neuroscience.* Sensitive Periods across Development 25, pp. 272–280. DOI: 10.1016/j.dcn.2016.10.005.

Redish, A. David (1999). *Beyond the Cognitive Map: From Place Cells to Episodic Memory.* Beyond the Cognitive Map: From Place Cells to Episodic Memory. Cambridge, MA, US: The MIT Press, pp. xviii, 420. xviii, 420.

Ribas-Fernandes, J. J. F. et al. (July 28, 2011). "A Neural Signature of Hierarchical Reinforcement Learning". In: *Neuron* 71.2, pp. 370–379. DOI: 10.1016/j.neuron.2011.05.042. pmid: 21791294.

Ribas-Fernandes, José JF et al. (2019). "Subgoal-and Goal-Related Reward Prediction Errors in Medial Prefrontal Cortex". In: *Journal of cognitive neuroscience* 31.1, pp. 8–23.

Richman, Ethan B. et al. (Nov. 2023). "Neural Landscape Diffusion Resolves Conflicts between Needs across Time". en. In: *Nature* 623.7987, pp. 571–579. DOI: 10.1038/s41586-023-06715-z.

Ringstrom, Thomas J. (2022). "Reward Is Not Necessary: How to Create a Compositional Self-Preserving Agent for Life-Long Learning." arXiv: 2211.10851.

Ronayne, David, Daniel Sgroi, and Anthony Tuckwell (June 1, 2021). "Evaluating the Sunk Cost Effect". en. In: *Journal of Economic Behavior & Organization* 186, pp. 318–327. DOI: 10.1016/j.jebo.2021.03.029.

Rudebeck, Peter H. et al. (Aug. 30, 2017). "Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes". eng. In: *Neuron* 95.5, 1208–1220.e5. DOI: 10.1016/j.neuron.2017.07.042. pmid: 28858621.

Rudorf, Sarah and Todd A. Hare (Nov. 26, 2014). "Interactions between Dorsolateral and Ventromedial Prefrontal Cortex Underlie Context-Dependent Stimulus Valuation in Goal-Directed Choice". en. In: *Journal of Neuroscience* 34.48, pp. 15988–15996. DOI: 10.1523/JNEUROSCI.3192-14.2014. pmid: 25429140.

Rummery, Gavin Adrian and M. Niranjan (1994). "On-Line Q-Learning Using Connectionist Systems". In.

Rushworth, Matthew F. S. et al. (Dec. 2012). “Valuation and Decision-Making in Frontal Cortex: One or Many Serial or Parallel Systems?” eng. In: *Current Opinion in Neurobiology* 22.6, pp. 946–955. DOI: `10.1016/j.conb.2012.04.011`. pmid: `22572389`.

Russek, Evan et al. (2017). “Predictive Representations Can Link Model-Based Reinforcement Learning to Model-Free Mechanisms”. In: *PLoS computational biology* 13.9, e1005768.

Russek, Evan et al. (2022). “Time Spent Thinking in Online Chess Reflects the Value of Computation”. In.

Russell, Stuart J. and Peter Norvig (2016). *Artificial Intelligence : A Modern Approach*. en. Pearson.

Salomon, Tom et al. (Feb. 26, 2018). “The Cue-Approach Task as a General Mechanism for Long-Term Non-Reinforced Behavioral Change”. en. In: *Scientific Reports* 8.1, p. 3614. DOI: `10.1038/s41598-018-21774-3`.

Samuelson, P. A. (1938). “A Note on the Pure Theory of Consumer’s Behaviour”. In: *Economica* 5.17, pp. 61–71. DOI: `10.2307/2548836`. JSTOR: `2548836`.

Schneider, Brett and Michael Koenigs (Dec. 2017). “Human Lesion Studies of Ventromedial Prefrontal Cortex”. eng. In: *Neuropsychologia* 107, pp. 84–93. DOI: `10.1016/j.neuropsychologia.2017.09.035`. pmid: `28966138`.

Schonberg, Tom et al. (Apr. 2014). “Changing Value through Cued Approach: An Automatic Mechanism of Behavior Change”. en. In: *Nature Neuroscience* 17.4 (4), pp. 625–630. DOI: `10.1038/nn.3673`.

Schuck, Nicolas W. et al. (Sept. 21, 2016). “Human Orbitofrontal Cortex Represents a Cognitive Map of State Space”. eng. In: *Neuron* 91.6, pp. 1402–1412. DOI: `10.1016/j.neuron.2016.08.019`. pmid: `27657452`.

Schultz, W., P. Dayan, and P. R. Montague (Mar. 14, 1997). “A Neural Substrate of Prediction and Reward”. eng. In: *Science (New York, N.Y.)* 275.5306, pp. 1593–1599. DOI: `10.1126/science.275.5306.1593`. pmid: `9054347`.

Searle, John R. (1980). “The Intentionality of Intention and Action*”. en. In: *Cognitive Science* 4.1, pp. 47–70. DOI: `10.1207/s15516709cog0401_3`.

Sellitto, Manuela, Elisa Ciaramelli, and Giuseppe di Pellegrino (Dec. 8, 2010). “Myopic Discounting of Future Rewards after Medial Orbitofrontal Damage in Humans”. eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 30.49, pp. 16429–16436. DOI: `10.1523/JNEUROSCI.2516-10.2010`. pmid: `21147982`.

Shadlen, Michael N. and Daphna Shohamy (June 1, 2016). “Decision Making and Sequential Sampling from Memory”. In: *Neuron* 90.5, pp. 927–939. DOI: `10.1016/j.neuron.2016.04.036`.

Shah, Anuj K. and Daniel M. Oppenheimer (2008). “Heuristics Made Easy: An Effort-Reduction Framework”. In: *Psychological Bulletin* 134.2, pp. 207–222. DOI: `10.1037/0033-2909.134.2.207`.

Shallice, T. and P. W. Burgess (Apr. 1991). “Deficits in Strategy Application Following Frontal Lobe Damage in Man”. eng. In: *Brain: A Journal of Neurology* 114 ( Pt 2), pp. 727–741. DOI: `10.1093/brain/114.2.727`. pmid: `2043945`.

Sharot, Tali, Benedetto De Martino, and Raymond J. Dolan (2009). “How Choice Reveals and Shapes Expected Hedonic Outcome”. In: *Journal of Neuroscience* 29.12, pp. 3760–3765.

Sharot, Tali, Cristina M. Velasquez, and Raymond J. Dolan (2010). “Do Decisions Shape Preference? Evidence from Blind Choice”. In: *Psychological Science* 21.9, pp. 1231–1235. DOI: `10.1177/0956797610379235`.

Shenhav, Amitai, Matthew M. Botvinick, and Jonathan D. Cohen (July 24, 2013). “The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function”. In: *Neuron* 79.2, pp. 217–240. DOI: `10.1016/j.neuron.2013.07.007`. pmid: `23889930`.

Shenhav, Amitai et al. (Sept. 2014). “Anterior Cingulate Engagement in a Foraging Context Reflects Choice Difficulty, Not Foraging Value”. en. In: *Nature Neuroscience* 17.9 (9), pp. 1249–1254. DOI: `10.1038/nn.3771`.

Shimojo, Shinsuke et al. (Dec. 2003). “Gaze Bias Both Reflects and Influences Preference”. en. In: *Nature Neuroscience* 6.12, pp. 1317–1322. DOI: `10.1038/nn1150`.

Shrout, P. E. and J. L. Fleiss (Mar. 1979). “Intraclass Correlations: Uses in Assessing Rater Reliability”. eng. In: *Psychological Bulletin* 86.2, pp. 420–428. DOI: `10.1037//0033-2909.86.2.420`. pmid: `18839484`.

Simon, Herbert A. (1955). “A Behavioral Model of Rational Choice”. In: *The Quarterly Journal of Economics* 69.1, pp. 99–118. DOI: `10.2307/1884852`. JSTOR: `1884852`.

— (1957). *Models of Man; Social and Rational.* Models of Man; Social and Rational. Oxford, England: Wiley, pp. xiv, 287. xiv, 287.

Sims, Christopher A. (Apr. 1, 2003). “Implications of Rational Inattention”. In: *Journal of Monetary Economics*. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information 50.3, pp. 665–690. DOI: `10.1016/S0304-3932(03)00029-1`.

Small, D. M et al. (Mar. 1, 2003). “The Posterior Cingulate and Medial Prefrontal Cortex Mediate the Anticipatory Allocation of Spatial Attention”. In: *NeuroImage* 18.3, pp. 633–641. DOI: `10.1016/S1053-8119(02)00012-5`.

Smittenaar, Peter et al. (Nov. 20, 2013). “Disruption of Dorsolateral Prefrontal Cortex Decreases Model-Based in Favor of Model-Free Control in Humans”. In: *Neuron* 80.4, pp. 914–919. DOI: `10.1016/j.neuron.2013.08.009`. pmid: `24206669`.

Spalding, Kelsey N. et al. (Nov. 25, 2015). “Investigating the Neural Correlates of Schemas: Ventromedial Prefrontal Cortex Is Necessary for Normal Schematic Influence on Memory”. en. In: *Journal of Neuroscience* 35.47, pp. 15746–15751. DOI: `10.1523/JNEUROSCI.2767-15.2015`. pmid: `26609165`.

Spalding, Kelsey N. et al. (Apr. 11, 2018). “Ventromedial Prefrontal Cortex Is Necessary for Normal Associative Inference and Memory Integration”. en. In: *Journal of Neuroscience* 38.15, pp. 3767–3775. DOI: `10.1523/JNEUROSCI.2501-17.2018`. pmid: `29555854`.

Starkweather, Clara Kwon, Samuel J. Gershman, and Naoshige Uchida (May 2, 2018). “The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty”. eng. In: *Neuron* 98.3, 616–629.e6. DOI: `10.1016/j.neuron.2018.03.036`. pmid: `29656872`.

Staw, Barry M. (June 1, 1976). “Knee-Deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action”. In: *Organizational Behavior and Human Performance* 16.1, pp. 27–44. DOI: `10.1016/0030-5073(76)90005-2`.

Stephens, David W. and John R. Krebs (1986). *Foraging Theory.* Vol. 1. Princeton University Press. DOI: `10.2307/j.ctvs32s6b`. JSTOR: `j.ctvs32s6b`.

Stoll, Frederic M., Vincent Fontanier, and Emmanuel Procyk (June 20, 2016). “Specific Frontal Neural Dynamics Contribute to Decisions to Check”. en. In: *Nature Communications* 7.1 (1), p. 11990. DOI: 10.1038/ncomms11990.

Strait, Caleb E., Tommy C. Blanchard, and Benjamin Y. Hayden (2014). “Reward Value Comparison via Mutual Inhibition in Ventromedial Prefrontal Cortex”. In: *Neuron* 82.6, pp. 1357–1366.

Strait, Caleb E., Brianna J. Sleezer, and Benjamin Y. Hayden (June 2015). “Signatures of Value Comparison in Ventral Striatum Neurons”. eng. In: *PLoS biology* 13.6, e1002173. DOI: 10.1371/journal.pbio.1002173. pmid: 26086735.

Sutton, Richard S. and Andrew G. Barto (1998). *Reinforcement Learning: An Introduction*. Cambridge: The MIT Press.

Sweis, Brian M. et al. (July 13, 2018). “Sensitivity to "Sunk Costs" in Mice, Rats, and Humans”. eng. In: *Science (New York, N.Y.)* 361.6398, pp. 178–181. DOI: 10.1126/science.aar8644. pmid: 30002252.

Talluri, Bharath Chandra et al. (Oct. 8, 2018). “Confirmation Bias through Selective Overweighting of Choice-Consistent Evidence”. In: *Current Biology* 28.19, 3128–3135.e8. DOI: 10.1016/j.cub.2018.07.052.

Tanji, Jun, Keisetsu Shima, and Hajime Mushiake (Dec. 1, 2007). “Concept-Based Behavioral Planning and the Lateral Prefrontal Cortex”. English. In: *Trends in Cognitive Sciences* 11.12, pp. 528–534. DOI: 10.1016/j.tics.2007.09.007. pmid: 18024183.

Tervo, D. Gowanlock R. et al. (June 2, 2021). “The Anterior Cingulate Cortex Directs Exploration of Alternative Strategies”. en. In: *Neuron* 109.11, 1876–1887.e6. DOI: 10.1016/j.neuron.2021.03.028.

Thorndike, Edward L. (1898). *Animal Intelligence: An Experimental Study of the Associative Processes in Animals.* en. New York: Columbia University Press. DOI: 10.1037/10780-000.

Thorngate, Warren (1980). “Efficient Decision Heuristics”. In: *Behavioral Science* 25.3, pp. 219–225. DOI: 10.1002/bs.3830250306.

Tolman, E. C. and C. H. Honzik (1930). “Degrees of Hunger, Reward and Non-Reward, and Maze Learning in Rats”. In: *University of California Publications in Psychology* 4, pp. 241–256.

Tolman, E. C., B. F. Ritchie, and D. Kalish (1946). “Studies in Spatial Learning. I. Orientation and the Short-Cut”. In: *Journal of Experimental Psychology* 36.1, pp. 13–24. DOI: 10.1037/h0053944.

Tom, Sabrina M. et al. (Jan. 26, 2007). “The Neural Basis of Loss Aversion in Decision-Making under Risk”. eng. In: *Science (New York, N.Y.)* 315.5811, pp. 515–518. DOI: 10.1126/science.1134239. pmid: 17255512.

Tranel, Daniel, Julie Hathaway-Nepple, and Steven W. Anderson (Apr. 2007). “Impaired Behavior on Real-World Tasks Following Damage to the Ventromedial Prefrontal Cortex”. In: *Journal of clinical and experimental neuropsychology* 29.3, pp. 319–332. DOI: 10.1080/13803390600701376. pmid: 17454352.

Treisman, Anne M. (1969). “Strategies and Models of Selective Attention”. In: *Psychological Review* 76.3, pp. 282–299. DOI: 10.1037/h0027242.

Trudel, Nadescha et al. (Jan. 2021). “Polarity of Uncertainty Representation during Exploration and Exploitation in Ventromedial Prefrontal Cortex”. en. In: *Nature Human Behaviour* 5.1 (1), pp. 83–98. DOI: 10.1038/s41562-020-0929-3.

Vaidya, A.R. and L.K. Fellows (2015a). “Testing Necessary Regional Frontal Contributions to Value Assessment and Fixation-Based Updating”. English. In: *Nature Communications* 6. DOI: 10.1038/ncomms10120.

Vaidya, Avinash R. and David Badre (June 1, 2022). “Abstract Task Representations for Inference and Control”. English. In: *Trends in Cognitive Sciences* 26.6, pp. 484–498. DOI: 10.1016/j.tics.2022.03.009. pmid: 35469725.

Vaidya, Avinash R. and Lesley K. Fellows (Sept. 16, 2015b). “Ventromedial Frontal Cortex Is Critical for Guiding Attention to Reward-Predictive Visual Features in Humans”. en. In: *Journal of Neuroscience* 35.37, pp. 12813–12823. DOI: 10.1523/JNEUROSCI.1607-15.2015. pmid: 26377468.

— (Sept. 21, 2016). “Necessary Contributions of Human Frontal Lobe Subregions to Reward Learning in a Dynamic, Multidimensional Environment”. In: *The Journal of Neuroscience* 36.38, pp. 9843–9858. DOI: 10.1523/JNEUROSCI.1337-16.2016. pmid: 27656023.

Vaidya, Avinash R., Marcus Sefranek, and Lesley K. Fellows (Nov. 1, 2018). “Ventromedial Frontal Lobe Damage Alters How Specific Attributes Are Weighed in Subjective Valuation”. eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 28.11, pp. 3857–3867. DOI: 10.1093/cercor/bhx246. pmid: 29069371.

Van Opheusden, Bas et al. (June 2023). “Expertise Increases Planning Depth in Human Gameplay”. en. In: *Nature* 618.7967, pp. 1000–1005. DOI: 10.1038/s41586-023-06124-2.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry (Sept. 2017). “Practical Bayesian Model Evaluation Using Leave-One-out Cross-Validation and WAIC”. In: *Statistics and Computing* 27.5, pp. 1413–1432. DOI: 10.1007/s11222-016-9696-4. arXiv: 1507.04544 [stat].

Veselic, Sebastijan et al. (Dec. 16, 2023). *A Cognitive Map for Value-Guided Choice in Ventromedial Prefrontal Cortex*. en. DOI: 10.1101/2023.12.15.571895. URL: https://www.biorxiv.org/content/10.1101/2023.12.15.571895v1 (visited on 03/12/2024). preprint.

Vikbladh, Oliver M. et al. (May 8, 2019). “Hippocampal Contributions to Model-Based Planning and Spatial Memory”. eng. In: *Neuron* 102.3, 683–693.e4. DOI: 10.1016/j.neuron.2019.02.014. pmid: 30871859.

Vinckier, Fabien et al. (Apr. 26, 2018). “Neuro-Computational Account of How Mood Fluctuations Arise and Affect Decision Making”. en. In: *Nature Communications* 9.1 (1), p. 1708. DOI: 10.1038/s41467-018-03774-z.

Von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Theory of Games and Economic Behavior. Princeton, NJ, US: Princeton University Press, pp. xviii, 625. xviii, 625.

Wajnberg, Eric et al. (2006). “Optimal Patch Time Allocation for Time-Limited Foragers”. In: *Behavioral Ecology and Sociobiology* 60.1, pp. 1–10. JSTOR: 25063777.

Walton, Mark E. et al. (Dec. 2011). “Giving Credit Where Credit Is Due: Orbitofrontal Cortex and Valuation in an Uncertain World”. eng. In: *Annals of the New York Academy of Sciences* 1239, pp. 14–24. DOI: 10.1111/j.1749-6632.2011.06257.x. pmid: 22145871.

Wang, Xiao-Jing (Dec. 5, 2002). “Probabilistic Decision Making by Slow Reverberation in Cortical Circuits”. eng. In: *Neuron* 36.5, pp. 955–968. DOI: 10.1016/s0896-6273(02)01092-9. pmid: 12467598.

Warren, David E. et al. (May 28, 2014). “False Recall Is Reduced by Damage to the Ventromedial Prefrontal Cortex: Implications for Understanding the Neural Correlates of Schematic Memory”. In: *The Journal of Neuroscience* 34.22, pp. 7677–7682. DOI: `10.1523/JNEUROSCI.0119-14.2014`. pmid: 24872571.

Wilson, Robert C. and Yael Niv (2011). “Inferring Relevance in a Changing World”. eng. In: *Frontiers in Human Neuroscience* 5, p. 189. DOI: `10.3389/fnhum.2011.00189`. pmid: 22291631.

Wilson, Robert C. et al. (Jan. 22, 2014). “Orbitofrontal Cortex as a Cognitive Map of Task Space”. en. In: *Neuron* 81.2, pp. 267–279. DOI: `10.1016/j.neuron.2013.11.005`.

Wing, Erik A. et al. (Aug. 1, 2021). “The Role of the Ventromedial Prefrontal Cortex and Basal Forebrain in Relational Memory and Inference”. eng. In: *Journal of Cognitive Neuroscience* 33.9, pp. 1976–1989. DOI: `10.1162/jocn_a_01722`. pmid: 34375419.

Wise, Steven P. (Dec. 2008). “Forward Frontal Fields: Phylogeny and Fundamental Function”. eng. In: *Trends in Neurosciences* 31.12, pp. 599–608. DOI: `10.1016/j.tins.2008.08.008`. pmid: 18835649.

Wittmann, Marco K. et al. (2016). “Predictive Decision Making Driven by Multiple Time-Linked Reward Representations in the Anterior Cingulate Cortex”. In: *Nature communications* 7.1, p. 12327.

Wolf, Richard C. et al. (June 1, 2014). “Ventromedial Prefrontal Cortex Mediates Visual Attention during Facial Emotion Recognition”. In: *Brain* 137.6, pp. 1772–1780. DOI: `10.1093/brain/awu063`.

Wolf, Richard C. et al. (Sept. 2016). “Emotion Recognition Deficits Associated with Ventromedial Prefrontal Cortex Lesions Are Improved by Gaze Manipulation”. eng. In: *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 82, pp. 255–262. DOI: `10.1016/j.cortex.2016.06.017`. pmid: 27423116.

Woolrich, M. W. et al. (Dec. 2001). “Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data”. eng. In: *NeuroImage* 14.6, pp. 1370–1386. DOI: `10.1006/nimg.2001.0931`. pmid: 11707093.

Woolrich, Mark W. et al. (Apr. 2004). “Multilevel Linear Modelling for FMRI Group Analysis Using Bayesian Inference”. eng. In: *NeuroImage* 21.4, pp. 1732–1747. DOI: `10.1016/j.neuroimage.2003.12.023`. pmid: 15050594.

Worbe, Y. et al. (May 2016). “Valence-Dependent Influence of Serotonin Depletion on Model-Based Choice Strategy”. en. In: *Molecular Psychiatry* 21.5, pp. 624–629. DOI: `10.1038/mp.2015.46`.

Young, Christina B. and Robin Nusslock (June 1, 2016). “Positive Mood Enhances Reward-Related Neural Activity”. In: *Social Cognitive and Affective Neuroscience* 11.6, pp. 934–944. DOI: `10.1093/scan/nsw012`.

Yu, Linda Q., Irene P. Kan, and Joseph W. Kable (2020). “Beyond a Rod through the Skull: A Systematic Review of Lesion Studies of the Human Ventromedial Frontal Lobe”. eng. In: *Cognitive Neuropsychology* 37.1-2, pp. 97–141. DOI: `10.1080/02643294.2019.1690981`. pmid: 31739752.

Zeithamova, Dagmar, April L. Dominick, and Alison R. Preston (July 12, 2012). “Hippocampal and Ventral Medial Prefrontal Activation during Retrieval-Mediated Learning Supports Novel Inference”. In: *Neuron* 75.1, pp. 168–179. DOI: `10.1016/j.neuron.2012.05.010`.

Zelazo, Philip David, Janet Wilde Astington, and David R. Olson (May 31, 2023). *Developing Theories of Intention: Social Understanding and Self-Control*. en. Psychology Press. 372 pp.